The Pearson Institute Discussion Paper No. 2023-3

Missing Discussions: Institutional Constraints in the Islamic Political Tradition

James A. Robinson
A. Arda Gitmez
Mehdi Shadmehr

THE PEARSON INSTITUTE
FOR THE STUDY AND RESOLUTION OF GLOBAL CONFLICTS

Missing Discussions: Institutional Constraints in the Islamic Political Tradition*

A. Arda Gitmez† James A. Robinson‡ Mehdi Shadmehr§

Abstract

Institutional constraints to counter potential abuses in the use of political power have been viewed as essential to well-functioning political institutions and good public policy outcomes in the Western World since the time of ancient Greece. A sophisticated intellectual tradition emerged to justify the need for such constraints. In this paper we identify a new puzzle: such an intellectual tradition did not exist in the Islamic world, even if the potential for abuse was recognized. We develop a model to explain why such ideas might not have emerged. We argue that this is due to the nature of Islamic law (the Sharia) being far more encompassing than Western law, making it easier for citizens to identify abuses of power and use collective action to discipline them. We study how the relative homogeneity and solidarity of Islamic society fortified this logic.

Keywords: Executive constraints, political traditions, Islam, collective action.
JEL classification: D7, P00, Z1.

*We are grateful for comments of Daron Acemoglu, Hassan Ansari, Alberto Bisin, Oeindrila Dube, Mohsen Kadivar, Timur Kuran, Gautam Nair, Josiah Ober, Şevket Pamuk, Steven Pincus, Michael Walzer, Alexander Wolitzky, seminar participants at Emory, UNC, University of Chile, Virtual Workshop in Formal Theory, 2022 Annual Meeting of American Political Science Association, 2023 North American Summer Meeting of Econometric Society, ASSA-NAWM 2024, and the discussion of Hannah Simpson at APSA. We have received excellent research assistance from David Kim.

†Department of Economics, Bilkent University & CEPR. E-mail: arda.gitmez@bilkent.edu.tr

‡University of Chicago, Harris School of Public Policy and Department of Political Science. E-mail: jamesrobinson@uchicago.edu

§Department of Public Policy, University of North Carolina at Chapel Hill. E-mail: mshadmehr@unc.edu

1 Introduction

Political institutions matter for public policy. A large political economy literature emphasizes the importance of institutional constraints on rulers to induce them to act in the collective interest (Persson et al., 1997; Aghion et al., 2004; Persson and Tabellini, 2004; Acemoglu et al., 2013). Some argue that the "rise of Europe" was based on the creation of such political institutions (North and Thomas, 1973; North and Weingast, 1989; Acemoglu et al., 2005). This modern literature reflects a deep Western intellectual tradition that includes the political thought of Madison, Montesquieu, Aquinas, Cicero, Polybius, and many others reaching as far back as Plato and Aristotle.

In this paper we identify a new puzzle: A similar intellectual tradition never arose in the Islamic world from the rise of Islam in the 7th century to the 18th century, prior to the emergence of broader modernization and Westernization currents in the Ottoman Empire and Iran. That is, Muslim thinkers, jurists, and philosophers did not develop, even in theory, ideas about the necessity of institutional mechanisms that aimed to constrain rulers.[1] This is despite the fact that Muslim thinkers were concerned about abuses of power and had access to much of the discussion of institutional political constraints by classical Greek philosophers (for example Plato's *Republic* and *Laws*, and Aristotle's *Ethics*, even if not his *Politics*). Moreover, the innovativeness of Islamic society generally in the Middle Ages is well documented, e.g., Ibn Haytham's (Alhazen) discoveries in optics—see Mokyr (1990) for other examples. How can we then make sense of the absence of discussions on institutional constraints on rulers in the Islamic political thought for over a millennium?[2]

We establish this puzzle and develop a model that formalizes a mechanism to make sense of it. There are two broad channels to hold rulers accountable and induce them to act in the public good: institutional constraints and collective action. In our model, the government chooses a policy, but citizens can overturn it by coordinating a revolt. Rulers may have private interests and citizens are uncertain about optimal government policies. There are

[1] We cannot claim that no Muslim scholar discussed institutional constraints in over a millennium—surely, there are undiscovered or lost works. Our aim is to establish the (likely) absence of a normative tradition.

[2] There is a literature on the absence of ruler-constraining institutions in the Islamic civilization (Lewis, 1982; Huntington, 1996; Blaydes and Chaney, 2013; Rubin, 2017; Stasavage, 2020; Bisin et al., 2023; Chaney, 2022; Kuran, 2012, 2023). What this research does not explain however is why notions of institutional constraints on rulers did not develop even in theory.

two categories of policies: those that fall under the scope of divine law, and those that do not. Divine law specifies the right policies under its scope. For example, divine law may specify a 10 percent tax on particular goods, but may not fully specify legitimate expenditures for the revenues. We compare citizen welfare in this environment with one in which there are also institutional constraints on rulers. Institutional constraints keep the government's transgression in check by dividing state authority. However, they also generate deadweight loss due to decision-making frictions and administrative costs.

We show that a broader scope of divine law facilitates collective action by making the rulers' deviations from the right policies observable. When the majority believe in divine law, this reduces their marginal benefits of institutional constraints. The scope of divine law is therefore a substitute for institutional constraints. This logic is strengthened when there is a larger majority (more homogeneity) and higher psychological benefits from participating in successful revolts that improve the majority's welfare (higher solidarity). The broader scope of divine law enables society to better evaluate government policy and hence the desirability of revolt; however, this knowledge helps the people only when they can mobilize, and their mobilization capacity depends on their homogeneity and solidarity. There is complementarity between the scope of divine law and societal homogeneity and solidarity.

Institutional constraints and political stability (revolt attempt and success) arise jointly in equilibrium. Assuming that institutions are adopted when they improve the majority's welfare, we show that institutional constraints and political stability are positively correlated. Institutional constraints are not adopted when collective action is an effective instrument to hold rulers accountable, e.g., in societies with comprehensive divine law and higher homogeneity and solidarity. Finally, we show that such societies have more institutional inertia: they are less likely to establish or dismantle institutional constraints in response to changes in their costs, e.g., reductions in costs due to the disappearance of foreign threats.

We argue that Islamic law in Islamic societies was viewed as divine law with a broad scope in public policy (Hallaq, 2009, 2014).[3] In contrast, Ancient Greek city-states and the Roman

[3]By Islamic law, we do not mean unified legal codes associated with modernity in the 19th and 20th century Islamic regions (Hallaq, 2009). Rather, broadly speaking, Islamic law refers to what God prescribes for human behavior and beliefs—for our purposes, according to Muslim scholars. It includes, among others, the laws of worship, contracts, inheritance, marriage, taxation, and wars. It existed since the time of Muhammad and was practiced long before the formation of Sunni schools.

Republic virtually had no divine law, and divine law had a limited scope in the subsequent Christian civilization.[4] Moreover, in the Islamic normative tradition, Islamic societies had a high degree of homogeneity and solidarity compared to the Western societies: all (were supposed to) contribute to the general welfare, "command right and forbid wrong" even in matters of government,[5] and all (were supposed to) seek the implementation of the God's law. These features were embodied in the notion of the Islamic *umma*.

These observations allow us to apply and interpret our formal results to compare Islamic and Western societies. We argue that the characteristics of Islamic civilization (in particular, a law with a broad scope stemming from the Quran, the Hadiths, and early traditions), combined with the nature of society, meant that it was less desirable to construct institutional constraints on rulers along the lines advocated in the West in various forms from Plato and Aristotle onwards. Given the costs of such institutions, revolt was a more effective disciplining device. We argue that this is a potential explanation for the lack of an intellectual tradition proposing institutional constraints. Islamic intellectuals and scholars were aware of the problem of tyranny, but saw the desirable solution as being different.

The normative political tradition in Jewish civilization is also consistent with these interpretations. Here, as in Islam, the scope of divine law was also broad, and the discussion of institutional constraints on rulers was absent throughout the first and second temple periods from the founding of the state up until its absorption into the Roman empire, including the Hasmoneans (Bickerman, 1962; Stern, 1968)—See Section D of the Online Appendix.

Our formalization focuses on showing that the marginal benefits of institutional constraints were lower in the Islamic and Jewish relative to the Greco-Roman and Christian ("Western") normative traditions, given their different core assumptions. Our broader argument about the "missing discussions" of institutional constraints in the Islamic normative tradition is based on our view of scholarly activities before the modern period (unmodelled here). This implicit "model of discussion" posits that scholars were more likely to write on an alternative theory to the status quo theory if they expected that societal welfare under

[4]As Bernard of Clairvaux wrote "True, thy [the pope's] palace is made to resound daily with noisy discussions relating to law, but it is not the law of the Lord, but the law of Justinian" (Tierney, 2010, p.92).

[5]This is reminiscent of Ibn Khaldûn (2015)'s *asabiyyah*, which "gives protection and makes possible mutual defense, the pressing of claims, and every other kind of social activity" (p.107).

that alternative arrangement would exceed welfare under the status quo arrangement by a higher margin. This reflects the scholars' limited resources (e.g., time and access to prior works) and their desire to work on relevant and useful topics among many potential topics of inquiry, especially when specialization was uncommon and resources were more scarce than the modern period—this resonates with modern academic readers who reflect on how we choose research topics. This approach is also familiar from the notion of paradigm shifts in science: "As in manufacturing so in science [and so in political theory]—retooling is an extravagance to be reserved for the occasion that demands it" (Kuhn, 1962, p.76). Our formal results, combined with this "model of discussion", imply that scholars in the Islamic tradition were less likely to write on institutional constraints than those in the Western tradition.

To our knowledge, this paper is the first that identifies the puzzle of missing discussions of institutional constraints on rulers in the Islamic normative tradition. Some came close, mentioning the problem in passing without offering an explanation. For example, Crone (2004, p.277) observes: "it was the scholars who formulated the law that the imam was meant to execute; by their own account, it was also they who elected and deposed him on behalf of the community. One would have thought that there was only a short step from all this to the view that the scholars should also monitor his performance, for example by forming independent councils authorized to signal when the rules had been breached, to strike out illegal decisions, and to block their execution. Small though the step may seem, however, there were few who took it."[6] Similarly, Roy (1994, p.61) notes that the "poverty of Islamist thought on political institutions is striking"; and Cook (2014, p.312) highlights that institutional constraints on rulers or "republics... were ignored by the normative tradition".

While the puzzle and explanation are new, the implications of our model are consistent with the literature documenting higher political instability in the Islamic versus the Western civilizations. For example, Cook (2001, p.161) argues that "In no other civilization was rebellion for conscience sake so widespread as it was in the early centuries of Islamic history"; Finer (1999, p.703) observes that between the 7th and the 10th centuries, the likelihood of

[6] Crone (2004) goes on to provide a few, short-lived, attempts on the eve of the Abbasid revolution to form councils that would rule along with the rulers. None of these attempts gained traction and they stand as exceptions proving the rule.

turmoil in Caliphate was higher compared to Byzantium: "On average, Byzantium suffered a violent incident every 16.7 years while the Caliphate did so every 7.4 years - *at more than double the rate*"; and Blaydes and Chaney (2013) provide quantitative evidence for this higher instability in the Islamic world from around 1000 to 1500. Our results are also consistent with the literature finding a negative correlation between institutional constraints and conflict (Besley and Persson, 2011).

Our paper is related to the origins of political institutions (Lizzeri and Persico, 2004; Acemoglu and Robinson, 2006) and the role of religion in their emergence, especially in the Islamic and Western civilizations. Blaydes and Chaney (2013) argue that executive constraints developed in Europe because rulers relied on local elites for military support. Because Muslim rulers relied on *mamluks* the society had less bargaining power to impose institutional constraints—see also Crone and Hinds (1986, p.106-7).[7] This, in turn, caused instability because rulers transgressed more often without institutional constraints—see also Besley and Persson (2011), in which institutional constraints reduce conflict by restricting in-group transfers. In our model, too, institutional constraints and stability are positively correlated, but they are determined jointly in equilibrium. While Blaydes and Chaney (2013) focus on why Muslims had relatively less *ability* to constrain rulers via institutions, we focus on the relatively lower *demand* for such institutional constraints and the puzzling absence of discussions of institutional constraints in Islamic (and Jewish) traditions.

An emerging literature explains "the long divergence" in economic performance or cultural dynamics by highlighting different aspects of Islamic law and its scholars. Kuran (2023) argues that Islamic law caused the weakening of civil society by contributing to preventing the emergence of powerful, flexible civil society organizations to counter the state's power, hindering political freedoms. He identifies multiple complementary self-enforcing causal channels, including egalitarian inheritance laws that hindered wealth accumulation, economic partnership laws that did not recognize corporations until modern times, and laws governing endowments (*waqf*), which did not allow for accountability of their managers to their constituents— see also Kuran (2012). Rubin (2017) argues that the rigidity of Islamic

[7] Stasavage (2020) offers a similar argument, but claims that the decisive factor was the adoption of the bureaucratic apparatus of the Persian Empire.

law compared to the law in the Western tradition caused stagnation in the long run. He further discusses how the origins of Islam caused Muslim rulers to follow Islamic law to acquire legitimacy and explores the self-enforcing dynamics of this legitimation path. Bisin et al. (2023) formalize and expand those dynamics to study the long-run evolution of culture and institutions. In their model, concessions to religious elites lead to a more religious culture, which, in turn, increases the returns from such concessions. The long-run outcomes of the resulting dynamics tend to be theocracy or a secular regime. Platteau (2017) and Auriol and Platteau (2017) argue that the decentralized clerical structure in Islam, compared to Christianity, made bargaining between rulers and the clergy less effective in avoiding conflict through the co-optation of the clergy who oppose the rulers' reforms. In Auriol, Platteau and Verdier (2023), the ruler balances concessions to the military and a decentralized clergy some of whom ideologically oppose the ruler's reforms.

We share with this literature the premise that the legitimacy and longevity of Islamic rulers depended on following Islamic law, which limited the range of legitimate policies compared to the West. However, our focus is on establishing and making sense of a novel puzzle: the missing discussions of institutional constraints in Islamic political thought. Our explanation is based on the direct influence of Islamic law on institutional constraints on rulers (in the normative tradition), unmediated by its effect on the relative power of various social groups. Critically, it is not the specific nature of the laws (e.g., contract law), but its expansive scope that underlies our explanation. Thus, our framework suggests that the development of institutional constraints may be more related to the scope of the divine law than to secular politics. Muslim thinkers in the Islamic normative tradition had very different policy preferences than their more secular counterparts in the West. However, they, too, did not want their rulers to deviate from those preferred policies. We provide an explanation for why it made sense to them to focus on revolt as the preferred means of accountability.

More broadly, our paper follows the tradition of Macpherson (1962), viewing political ideas as being embedded in the social context. For Macpherson, the arguments of Hobbes and Locke for political institutions presupposed "a certain model of society" that "did correspond in large measure to seventeenth century English society" (Macpherson, 1962, p.16)—see also Ashcraft (1986). He characterized this as a market-based society of possessive individualists,

where men "continually expected to be invaded by others" (p.46). This logic implies that Islamic society would come up with different ideas about desirable political institutions. For example, it did not make sense to imagine a state of nature where all law was removed, since God had provided the law. Moreover, Islamic society was not the hyper-individualistic society Hobbes and Locke considered. Our model helps to understand how these differences meant that Islamic thinkers did not innovate the types of ideas that Hobbes and Locke did.

Our model has two intertwined building blocks: institutional constraints on rulers, and revolt by citizens. The first is essentially an agency model, but we are agnostic about the exact mechanism through which institutional constraints improve accountability. In our baseline model, dividing power among multiple rulers check transgressors. This can be interpreted as lowering the threshold of successful revolt as in Aghion et al. (2004), or as a reduced-form way of capturing the idea that separation of powers reduces rent-seeking as in Persson et al. (1997) and Acemoglu et al. (2013). In our model in the Online Appendix (Section B), the equilibrium behavior of rulers reveals information about the policy, so that institutional constraints provide the information necessary for revolt. Our model of revolt takes a global games approach (Morris and Shin, 2003) to coordination for revolt, and can be transformed into the revolt model of Boleslavsky et al. (2021). That divine law facilitates collective action by making the ruler's transgressions more transparent is reminiscent of Weingast (1997)'s argument that constitutions and social consensus can act as focal points for coordination problems. While we abstract from the issue of conversion (Saleh and Tirole, 2021) or the role of leaders in instigating or controlling revolt (Chaney, 2013), we show how an extension of our model captures the rulers' attempts to co-opt jurists as reducing the scope of divine law and discuss the consequences—see the discussions after Proposition 7.

The paper proceeds as follows. Next, we develop our model and present the formal analysis. Section 3 presents the evidence on the absence of arguments for institutional constraints on rulers in the Islamic normative tradition. Section 4 examines the relationship between Islamic law, rebellion, and accountability. A discussion and conclusion follows, where we discuss alternative explanations. Section D of the Online Appendix explores institutional constraints on rulers in Jewish, Greco-Roman, and Christian normative traditions. The Online Appendix also examines the robustness of formal results to alternative modeling choices.

2 Model and Theoretical Analysis

Players There is a unit measure of citizens and a ruler. Each citizen belongs to one of the two groups: majority and minority, with conflicting policy preferences. The size of the majority is $M \in (1/2, 1]$ and the size of the minority is $1 - M$.

There are two types of rulers: majority-congruent and minority-congruent rulers. A majority-congruent ruler's preferences are aligned with the majority, while a minority-congruent ruler's preferences are aligned with the minority. A ruler is minority-congruent with probability $q \in (0, 1)$ and is majority-congruent with probability $1 - q$.

Actions The ruler chooses a binary action $a \in \{0, 1\}$, representing government policy. After observing the ruler's action, citizens simultaneously decide whether to revolt. The revolution succeeds if and only if the measure of revolters exceeds the regime's strength $T \in (1/2, M)$. If the revolt succeeds, denoted by $r = 1$, the ruler's action is reversed, so that government policy a becomes $1 - a$. If the revolt fails, the ruler's policy is maintained. Thus, the final government policy is $d(a, r) = a(1 - r) + (1 - a)r$.

Payoffs Citizen's policy payoffs depend on their group (majority or minority), the final government policy d, and the state of the world $s \in \{0, 1\}$. A minority citizen receives a policy payoff 1 if the final government policy is $d = 1$, and otherwise, a policy payoff of -1. In contrast, a majority citizen receives a policy payoff 1 if the final government policy matches the state of the world, and otherwise, a policy payoff -1.

Let k be the measure of revolters. Table 1 (Table 2) illustrates citizen i's payoffs when i is a member of the majority (minority), where $T \in (1/2, M)$ is the regime's strength, $\gamma \in (0, 1)$ is pleasure-in-agency rewards from revolting (Morris and Shadmehr, 2023), and $c_i \geq 0$ is citizen i's direct costs of revolting.

	If $a \neq s$		If $a = s$	
	$k > T$	$k \leq T$	$k > T$	$k \leq T$
revolt	$(1 + \gamma) - c_i$	$-1 - c_i$	$-(1 + \gamma) - c_i$	$1 - c_i$
no revolt	1	-1	-1	1

Table 1: Majority citizens' payoffs.

	If $a=0$		If $a=1$	
	$k > T$	$k \leq T$	$k > T$	$k \leq T$
revolt	$(1+\gamma) - c_i$	$-1 - c_i$	$-(1+\gamma) - c_i$	$1 - c_i$
no revolt	1	-1	-1	1

Table 2: Minority citizens' payoffs.

The majority-congruent ruler is nonstrategic, always taking the action that matches the state of the world. The minority-congruent ruler's payoff, u, depends on his action, the state of the world, and whether or not a successful revolution occurred against him. In particular,

$$u(a, r, s) = (a + (1-a)\delta_s)(1-r), \ \delta_s \in (0,1) \tag{1}$$

That is, the minority-congruent ruler receives 0 if there is a successful revolt against him. Otherwise, he receives a payoff 1 if he takes action 1, and a payoff of δ_s if he takes action 0 in state s. We set $0 = \delta_1 < \delta_0 < 1$, so that he has more incentives to take action 0 in state 0 than in state 1, and he has more incentives to take action 1 overall.

Information It is common knowledge that $Pr(s=1) = Pr(s=0) = 1/2$. The ruler always observes the state s. Citizens do not observe the ruler's type, and they observe the state with probability $p \in [0,1]$. For exposition, let $\hat{s}$ be a truth-or-noise public signal of the state:

$$\hat{s} = \begin{cases} s, & \text{with probability } p \\ \emptyset, & \text{with probability } 1-p \end{cases}$$

An interpretation is that there are various policy issues, ranging from criminal law (e.g., punishment for burglars) to public finance (e.g., the expenditure of revenue from conquests). A policy issue may be *preordained/canonical* or *non-preordained/secular*. When a policy issue is preordained/canonical, the majority has better information about the "right policy" for them. The probability that a preordained issue arises is $p \in [0,1]$. Thus, a higher p captures a larger *scope of the law*.

The costs of revolting c_i are correlated among citizens. In particular, $c_i = \bar{c} + \rho\epsilon_i$, $\rho > 0$,

where $\bar{c} \sim H = U[0,1]$ and $\epsilon_i \sim_{iid} F$, with $F(0) = 0$ and corresponding log-concave pdf f. A citizen i's cost c_i is citizen i's private information. We take a global games approach to equilibrium selection, focusing on the equilibrium outcomes in the limit when ρ is vanishingly small ($\rho \to 0$).

Timing The timing of the game is as follows.

1. The nature determines the realizations of ruler's type, the state of the world s, signal $\hat{s}$, the common value of costs $\bar{c}$, and idiosyncratic elements of costs ϵ_is.
2. The ruler observes his own type, the state s, and $\hat{s}$. Each citizen i observes $\hat{s}$ and her private cost c_is.
3. The ruler chooses government policy a, which the citizens observe.
4. Citizens simultaneously decide whether or not to revolt.
5. Success of the revolution is determined, payoffs are received, and the game ends.

Strategies and Equilibrium A majority-congruent ruler is a behavioral type who always chooses $a = s$. A minority-congruent ruler's strategy is a mapping from the state of the world s and signal $\hat{s}$ to a probability $\sigma(\hat{s}, s)$ of taking action 1: for every possible history $(\hat{s}, s) \in \{(0,0),(1,1),(\emptyset,0),(\emptyset,1)\}$, the minority-congruent ruler's strategy specifies $\sigma(\hat{s}, s) \in [0,1]$. A strategy for citizen i is a mapping from her group membership, signal $\hat{s}$, government policy a and her private costs c_i to a decision whether to revolt. We characterize Perfect Bayesian Equilibria in the limit when ρ approaches 0.

Preliminary Analysis Because a minority citizen prefers $d = 1$, she revolts only if $a = 0$. Moreover, because $T \in (1/2, M)$, she revolts only if she believes that some members of the majority revolt as well when $a = 0$. For this to be part of an equilibrium, the (minority-congruent) ruler must sometimes take action 0 when the state is 1, and receive a payoff of 0. If, instead, the ruler takes action $a = 1$, his payoff will be 0 if and only if the probability of successful revolution is 1. As we will show in Proposition 1 below, this probability is

always strictly less than one due to coordination and information frictions. Therefore, a ruler never takes $a = 0$ when the state is 1, and consequently a minority citizen never revolts in equilibrium.

We now focus on the decision of majority citizens; henceforth, by citizen, we mean a majority citizen. Let q' be a citizen's posterior belief that the ruler's action a does not match the state of the world s. The majority citizens' payoffs (Table 1) and information structure maps into the coordination problem analyzed in Proposition 1 of Boleslavsky, Shadmehr and Sonin (2021) by setting $u_r = u_0 = 1$, $\delta = \gamma$, $\theta = T$, and normalizing the size of citizens to M. Because the game has only one-sided limit dominance, there is always an equilibrium in which no one revolts. We follow the analysis of Boleslavsky, Shadmehr and Sonin (2021), and focus on symmetric cutoff strategy equilibria with cutoffs strictly greater than 0, when they exist.

To communicate the logic, suppose $\bar{c}$ has full support on $\mathbb{R}$. A citizen i's strategy in a symmetric equilibrium is described by a threshold c^*, so that she revolts if and only if her direct cost of revolt is below that threshold: $c_i < c^*$. When almost all citizens follow this strategy, for any $\bar{c}$, the size of revolters is $k(\bar{c}) = Pr(c_i < c^*|\bar{c})M$. The size of revolters $k(\bar{c})$ is decreasing, falling from $\lim_{\bar{c}\to-\infty} k(\bar{c}) = M$ to $\lim_{\bar{c}\to\infty} k(\bar{c}) = 0$, and crossing $T \in (0, M)$ at a unique $\bar{c}$. Let $\bar{c}^*$ be that threshold, so that

$$Pr(c_i < c^*|\bar{c} = \bar{c}^*) = T/M \qquad \text{(belief consistency)}$$

Thus, given the strategy of others c^*, the consistency of beliefs with strategies implies that a citizen will believe: if $\bar{c} < \bar{c}^*$, then $Pr(c_i < c^*|\bar{c} = \bar{c}^*) > T/M$ and the revolution succeeds; if $\bar{c} \geq \bar{c}^*$, then $Pr(c_i < c^*|\bar{c} = \bar{c}^*) \leq T/M$ and the revolution fails. A citizen with the critical threshold $c_i = c^*$ must be indifferent between revolting and not revolting:

$$Pr(\bar{c} < \bar{c}^*|c_i = c^*)(2q' - 1)\gamma = c^* \qquad \text{(individual rationality)}$$

where we recall that $q' = Pr(a \neq s)$, so that $Pr(a \neq s)\gamma - Pr(a = s)\gamma = (2q' - 1)\gamma$. As Morris and Shin (1998, 2003) show, when there is no common knowledge about the value of $\bar{c}$ or when the noise in private signals becomes vanishingly small ($\rho \to 0$), a citizen with

the threshold signal $c_i = c^*$ believes that the size of revolters is distributed uniformly in its range $[0, M]$, so that a revolution succeeds with probability $(1 - T/M)$.[8] Thus, individual rationality condition implies that in the limit as $\rho \to 0$: $\lim_{\rho \to 0} c^*(\rho) = \lim_{\rho \to 0} \bar{c}^*(\rho) = (1 - T/M)\gamma(2q' - 1)$. Along with $\bar{c} \sim H$, then, the probability of a successful revolution in the limit is $H\left((1 - T/M)\gamma(2q' - 1)\right)$.

Proposition 1. *In the limit when $\rho \to 0$, the likelihood of successful revolution is $\beta(q', M, \gamma) = H\left((1 - T/M)\gamma(2q' - 1)\right))$.*

The likelihood of successful revolution $\beta(q', M, \gamma)$ has natural properties: it is increasing in the size of the majority M and in pleasure-in-agency rewards γ, implying that the probability of successful revolt is higher in societies with higher levels of homogeneity or solidarity. It is also increasing in the citizen's posterior that the ruler chose a policy that did not match the state. In particular, citizens have a dominant strategy not to revolt when $q' < 1/2$; there is a successful revolution only if citizens believe that the ruler has likely taken a wrong policy.[9]

Equilibrium Characterization To characterize equilibrium outcomes, recall that (1) $\sigma(\hat{s}, s)$ is the minority-congruent ruler's strategy given signal $\hat{s}$ and state s, (2) $\hat{s} = s$ captures preordained policy issues and $\hat{s} = \emptyset$ captures non-preordained policy issues, and (3) $\beta(q', M, \gamma)$ is the probability of a successful revolt given a posterior belief q' that the ruler's action does not match the state ($a \neq s$). First, consider preordained policy issues, so that citizens observe the state s, and hence they know whether it matches the ruler's action ($q' \in \{0, 1\}$). When $s = 1$, so that there is no conflict of interest between the ruler and citizens, the ruler chooses action 1. When $s = 0$, so that there is conflict of interest, the ruler faces a trade-off. Take action 1 and risk revolution for a high payoff of 1, or take the

[8] This follows from the a statistical property that, in the limit, $Pr(\bar{c} < \bar{c}^* | c_i = c^*) = Pr(c_i \geq c^* | \bar{c} = \bar{c}^*)$. Alternatively, using this statistical property, one can substitute $1 - T/M$ from the belief consistency condition into the individual rationality condition to obtain the same result.

[9] The likelihood of successful revolt β is also decreasing in T, which captures, e.g., the ruler's strength relative to citizens. For instance, all else equal, a more vibrant civil society with more independent and flexible private organizations that control large resources corresponds to a lower T. Kuran (2016, 2023) argues that the institution of *waqf* contributed to locking private resources for public uses, relatively immune from government confiscations, but also with restrictions that hindered their employment in political actions.

safe action 0 and receive a low payoff of δ_0 with certainty.[10] From Proposition 1, the probability of a successful revolt following action 1 is $\beta(1, M, \gamma)$. Thus, the ruler takes action 1 whenever $\delta_0 < 1 - \beta(1, M, \gamma)$.

Next, consider non-preordained issues, so that $\hat{s} = \emptyset$, and hence citizens have to infer whether the state matches the ruler's action in equilibrium. Let $q'(a)$ be citizen posterior that the state does not match the ruler's action a:

$$q'(a) = Pr(s \neq a|a) = \frac{Pr(s \neq a, a)}{Pr(a)}.$$

Observe that $q'(a) \leq 1/2$ if and only if, in equilibrium, $Pr(s \neq a, a) \leq Pr(s = a, a)$. A majority-congruent ruler's action always matches the state. Thus, in equilibrium, a sufficient condition for $Pr(s \neq a, a) \leq Pr(s = a, a)$ is that the minority-congruent ruler takes action 1 with a weakly higher probability in state 1 than in state 0; that is, $\sigma(\emptyset, 1) \geq \sigma(\emptyset, 0)$. This is ensured by $\delta_0 > \delta_1$, so that the ruler has more incentives to take action 1 in state 1 than in state 0. Now, because $q'(\hat{s} = \emptyset, a) \leq 1/2$, no one revolts and the probability of successful revolution is 0 for any a (Proposition 1). Given that there is no risk of revolution in taking action 1, the minority-congruent ruler always takes action 1.

The majority citizen's expected policy payoff can then be calculated from this equilibrium characterization. Thus, we have proved the following Proposition.

Proposition 2. *In equilibrium,*

$$\sigma(\hat{s}, 1) = \sigma(\hat{s} = \emptyset, 0) = 1 \quad \textit{and} \quad \sigma(\hat{s} = s, 0) = \begin{cases} 0 & ; \ \beta(1, M, \gamma) > 1 - \delta_0 \\ 1 & ; \ \beta(1, M, \gamma) < 1 - \delta_0 \end{cases}$$

There is a revolt only if $\hat{s} = 0$ and the ruler takes action 1. This revolt succeeds with

[10] When the ruler takes $a = 1$, minority members do not take part in the revolution because they strictly prefer to keep $a = 1$. Therefore, the only citizens who may participate in a revolt are the majority citizens, and Proposition 1 applies. On the other hand, if the ruler takes $a = 0$ when $\hat{s} = 0$, majority members do not take part in the revolution. Anticipating this, minority members recognize that a revolution will not succeed, and so they do not revolt.

probability $\beta(1, M, \gamma)$*. Moreover, the expected policy payoff for a majority citizen is*

$$\begin{cases} 1 - q(1-p) & ; \ \beta(1, M, \gamma) > 1 - \delta_0 \\ 1 - q(1 - p\beta(1, M, \gamma)) & ; \ \beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

When $\beta(1, M, \gamma)$ is sufficiently large, the threat of successful revolt deters rulers from majority-incongruent policies that fall under the scope of preordained law. Thus, a citizen's policy payoff is 1 unless both the ruler is minority-congruent and the policy issue is secular, in which case the citizen receives -1. When $\beta(1, M, \gamma)$ is lower, e.g., due to low levels of homogeneity or solidarity, there is no deterrence effect, and bad policies (from the majority's perspective) are reversed when the policy issue is preordained (so that a revolt is attempted) and the revolt succeeds.

To simplify exposition, in the main text, we focus on the citizens' policy payoffs. In Section A of the Online Appendix, we show that our results go through if the costs of revolt are included in the payoff of majority citizens. In particular, we show that accounting for the expected costs of revolt in citizen payoffs amounts to substituting β with $\beta_c(\beta) = \beta - \beta^2/4$ when calculating expected payoffs. This allows us to extend our qualitative results and comparative statics when payoffs are inclusive of revolt costs.

It is worth highlighting that to improve their welfare the majority could also attempt to directly reduce q. One may call this approach the Deuteronomic or Platonic approach to good governance by attempting to install good rulers and controlling their temptations via education, advice, or prayer. Plato's ideas in *Republic* on the selection and training of the guardians and philosopher-kings was an early secular example. In the Jewish tradition, Deuteronomic editors insisted that the king must "write for himself a copy of this Law on a scroll... And it shall be with him, and he shall read it all the days of his life, so that he will learn to fear the LORD his God, by carefully following all the words of this Law and these statutes, that his heart may not be lifted up above his brothers, and that he may not turn aside from the commandment, either to the right hand or to the left" (Deut 17:18-19). The Mirrors for Princes in the Islamic tradition (e.g., Nizam al-Mulk's *Sīāsat Nāmih*) similarly advise the rulers to meet with religious scholars routinely to learn about Islamic law and

tradition.

Discussions on the institutions designed to select good rulers reflect this approach. Such institutions include *shūrā* in the Islamic tradition (see Section 3) and various electoral institutions in the Western tradition, including arguments about the effect of larger electoral districts on improving the quality of elected officials ("fit characters" with "enlightened views and virtuous sentiments") in the Federalist No. 10 (Hamilton et al., 2008). However, institutional arrangements to select rulers are distinct from those aimed to constrain rulers once they are in power. As Halpern (1981, p.222) puts it in the context of the ancient Jewish tradition, "The body negotiating the elevation of the monarch has the opportunity to impose conditions, to extract promises, and to level ultimata. Whether the king after his accession actually paid attention to them is, of course, another matter, about which our sources are too inadequate to permit speculation". Similarly, in the Islamic tradition, Crone (2004, p.277) argues that, "once elected, the caliph was free to ignore all the advice he received". The next section turns to institutional constraints on rulers once they are in power.

2.1 Institutional Constraint, Revolt, and the Scope of Law

We now introduce institutional constraints to the model. These institutional constraints are aimed to increase the likelihood of majority-congruent government policies. We consider a particular form of institutional constraints that divide decision-making power between multiple rulers. This approach is reminiscent of separation of powers, but we are not concerned with executive versus legislative or judicial powers per se. For example, the presence of two Roman consuls is an example of this power-sharing institutional setting among rulers.

Model The model is the same except that there are two rulers, ruler 1 and ruler 2, whose types (denoted by t_1 and t_2) are independent. Nature determines the state s, the signal $\hat{s}$, the rulers' types, the common value $\bar{c}$ and the idiosyncratic values ϵ_is of revolution costs. All the fundamentals and noises are independent of each other. The state is observed by both rulers, and the signal is observed by all. Moreover, a ruler observes his own type and the type of the

other ruler,[11] and a citizen i privately observes her own revolution cost $c_i = \bar{c} + \rho\epsilon_i$. Ruler 1 moves first, choosing $a_1 \in \{0, 1\}$. Then, ruler 2 observes a_1 and chooses $a_2 \in \{0, 1\}$. Absent revolt, the government's aggregate policy is a function of the rulers' actions, $A = y(a_1, a_2)$. Upon observing the rulers' actions a_1 and a_2, citizens simultaneously decide whether to revolt. If the revolution succeeds, denoted by $r = 1$, the government's aggregate policy is reversed. If the revolution fails, denoted by $r = 0$, the government's aggregate policy is maintained. Thus, the final government policy is $d(a_1, a_2, r) = A(1 - r) + (1 - A)r$. Payoffs are realized and the game ends.

A majority-congruent ruler is the same behavioral type as before. A minority-congruent ruler has the same payoffs as before, and if a revolution succeeds, both rulers receive 0. That is, the minority-congruent ruler's payoff u is:

$$u(a_1, a_2, r, s) = (A + (1 - A)\delta_s)(1 - r) \tag{2}$$

Citizens' payoffs are identical to the previous model, with the government's aggregate action A replacing the single ruler's action a. There is a deadweight loss $\mu \in [0, 1]$ due to institutional constraints, subtracted from the citizens' policy payoffs. The deadweight loss is associated with direct inefficiencies, delays, or administrative costs of institutional constraints such as power-sharing.

To proceed with the analysis, we must specify how the actions of two rulers (a_1, a_2) are combined into a government's aggregate policy A. Naturally, if both rulers take the same action, the aggregate government policy is the same as individual actions. When the rulers' actions differ, we take a certain stance. We assume that if one ruler takes action 0 and the other takes action 1, the government's aggregate policy will be 0: $A = \min\{a_1, a_2\}$. When actions differ, citizens will know that at least one ruler is minority-congruent, and that it is the minority-congruent ruler who has incentives to take action 1 in state 0. Moreover, the division in the government weakens the rulers' coercive power. Motivated by these observations, we assume that a majority-congruent ruler, backed by the majority, succeeds

[11] In Section C of the Online Appendix, we present a model where the rulers do not observe each others' types. That model has multiple equilibria, and the forward induction refinement yields a unique equilibrium outcome described in Proposition 3.

in making the aggregate government policy 0 even when his minority-congruent co-ruler attempts to set government policy 1. Another rationale is that when one of the rulers is majority-congruent, he will help the majority with government resources and ensure that a revolution attempt succeeds. In Section B of the Online Appendix, we present an alternative model of institutional constraints where $A = \max\{a_1, a_2\}$, and discuss which one of our main insights remain robust to the alternative specification.

Our modeling of institutional constraints aims to capture two key features. (i) Institutional constraints can directly reduce government power. While this may hinder the government's capacity to formulate and implement best policies (captured by μ), it also curbs the government's coercive power against citizens. (ii) Institutional constraints can facilitate the revelation of information to citizens about the government's private information (e.g., about the right policies). Our model in the text focuses on (i), while our model in Section B of the Online Appendix focuses on (ii).

Strategies and Equilibrium Let g denote the type of a majority-congruent ruler and b denote the type of a minority-congruent ruler, so that $t_j \in \{g, b\}$, for $j \in \{1, 2\}$. As before, a ruler j with type $t_j = g$ always chooses $a_j = s$. Let σ_1 be the strategy of ruler 1 with type $t_1 = b$, and σ_2 be the strategy of ruler 2 with type $t_2 = b$.

The strategy σ_1 is a mapping from the state of the world s, signal $\hat{s}$, and ruler 2's type t_2 to a probability of taking action 1: $\sigma_1(\hat{s}, s, t_2) \in [0, 1]$. The strategy σ_2 is a mapping from the state of the world s, signal $\hat{s}$, ruler 1's action a_1, and ruler 1's type t_1 to the probability of taking action 1. Given that ruler 2 observes a_1, ruler 1's type t_1 is not payoff-relevant, and hence we drop it from the arguments of σ_2, writing $\sigma_2(\hat{s}, s, a_1) \in [0, 1]$.[12] As before, a citizen i's strategy is a mapping from his group membership, signal $\hat{s}$, actions (a_1, a_2) and her private costs c_i to a decision whether to revolt; and we characterize Perfect Bayesian Equilibria in the limit when ρ approaches 0.

Equilibrium Characterization under Institutional Constraints First, consider preordained policy issues, so that $\hat{s} = s$. When at least one ruler is majority-congruent, or when

[12] When ruler 2 is indifferent between actions 0 and 1, he may condition his action on t_1, but as we will see, that will not matter for the equilibrium government policy A or citizen decisions.

$s = 1$ (so that there is no conflict of interest), the aggregate policy will match the state. When both rulers are minority-congruent and $s = 0$, the minority-congruent rulers face a trade-off. As in Proposition 2, both of them take action 1 whenever $\beta(1, M, \gamma) < 1 - \delta_0$.

Next, consider non-preordained issues, so that $\hat{s} = \emptyset$. Let $Pr_{(t_1,t_2)}(A)$ be the probability of A conditional on rulers' types (t_1, t_2), and let $q'(a_1, a_2)$ be the citizen posterior that the state does not match the aggregate policy A: $q'(a_1, a_2) = Pr(s \neq A|a_1, a_2)$. Suppose $s = 1$. The majority-congruent ruler takes action 1. If $a_1 = 1$, the minority-congruent ruler 1 takes action 1, because action 0 will yield a payoff of 0 whereas action 1 yields a strictly positive payoff; even if a revolution attempt follows, it fails with a non-zero probability. Due to the same reasoning, the minority-congruent ruler 1 also always chooses $a_1 = 1$ when $s = 1$. This implies that $Pr_{(t_1,t_2)}(A = 1|\hat{s} = \emptyset, s = 1) = 1$ in any equilibrium. Moreover,

$$q'(1,1) = \frac{Pr(a_1 = a_2 = 1, s = 0)}{\sum_s Pr(a_1 = a_2 = 1, s)} = \frac{Pr(a_1 = a_2 = 1|s = 0)}{1 + Pr(a_1 = a_2 = 1|s = 0)} \leq \frac{1}{2}.$$

That is, $(a_1, a_2) = (1, 1)$ does not provide sufficient information in favor of $s = 0$, and no revolts follow this action profile. This implies $\sigma_2(\emptyset, 0, 1) = 1$ in any equilibrium: upon observing $a_1 = 1$, if ruler 2 takes action 0, he at most gets δ_0; however, if he takes action 1, he will receive $1 > \delta_0$. Knowing that the minority-congruent ruler 2 will follow suit, the minority-congruent ruler 1 also takes action 1 even when $s = 0$: $\sigma_1(\emptyset, 0, b) = 1$. Thus, $Pr_{(b,b)}(A = 1|\hat{s} = \emptyset, s = 0) = 1$. This also implies $q'(0, 0) = 0$, therefore, majority citizens do not revolt following $(a_1, a_2) = (0, 0)$ and, because minority citizens cannot successfully revolt on their own, there are no revolts.

It remains to analyze what happens when $s = 0$ and rulers have different types. Suppose ruler 1 is majority-congruent, and ruler 2 is minority-congruent. If $(a_1, a_2) = (0, 1)$ is observed on the equilibrium path, majority citizens will deduce that $s = 0$: $q'(0, 1) = 0$, and they will not revolt because the aggregate action is 0. Because minority citizens cannot successfully revolt on their own, we conclude that there are no revolts following this action profile. Thus, ruler 2 is indifferent between the two actions, which is consistent with observing $(a_1, a_2) = (0, 1)$ on the equilibrium path: there is an equilibrium where $\sigma_2(\emptyset, 0, 0) > 0$. Alternatively, if $(a_1, a_2) = (0, 1)$ is never observed on the equilibrium path, Bayesian up-

dating does not restrict $q'(0,1)$. If $q'(0,1)$ is high enough, ruler 2 is deterred from taking action 1, which is consistent with never observing $(a_1, a_2) = (0,1)$ on the equilibrium path: there is an equilibrium where $\sigma_2(\emptyset, 0, 0) = 0$.[13] Regardless, in any equilibrium, the aggregate action is 0 and there are no revolts: the majority-congruent ruler 1 will discipline the minority-congruent ruler 2. The same logic applies when the order is reversed.

The following Proposition summarizes these results.

Proposition 3. *Recall that A is the aggregate government action, and $Pr_{(t_1,t_2)}(A)$ is the probability of A conditional on rulers' types (t_1, t_2). In equilibrium,*

$$Pr_{(t_1,t_2)}(A = s) = 1, \quad \textit{if } (t_1, t_2) \neq (b, b).$$

Otherwise,

$$Pr_{(b,b)}(A = 1|\hat{s}, s = 1) = Pr_{(b,b)}(A = 1|\hat{s} = \emptyset, s = 0) = 1$$

and

$$Pr_{(b,b)}(A = 1|\hat{s} = s, s = 0) = \begin{cases} 1 & ; \beta(1, M, \gamma) < 1 - \delta_0 \\ 0 & ; otherwise. \end{cases}$$

There is a revolt only if $\hat{s} = 0$ and both rulers take action 1. This revolt succeeds with probability $\beta(1, M, \gamma)$. Moreover, the expected policy payoff for a majority citizen is

$$\begin{cases} 1 - q^2(1-p) - \mu & ; \beta(1, M, \gamma) > 1 - \delta_0 \\ 1 - q^2(1 - p\beta(1, M, \gamma)) - \mu & ; \beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Proposition 3 shows that institutional constraints disciplines a bad ruler when he is matched with a good co-ruler, and does not change his behavior when his co-ruler is also bad.

[13] The equilibrium analysis simplifies even further if the minority-congruent ruler has a tie-breaking rule that favors $a_j = 0$ when he is indifferent. This can be microfounded by assuming that the minority-congruent ruler j obtains the payoffs associated with A if he takes action $a_j = A$, and 0 otherwise; or considering a infinitesimal positive payoff from taking $a_j = A$.

Propositions 2 and 3 enable us to compare the marginal change in a majority citizen's policy payoff from institutional constraints, and study how it varies with the environment.

Corollary 1. *The value of institutional constraints is:*

$$\begin{cases} (1-p)(q-q^2)-\mu & ;\beta(1,M,\gamma) > 1-\delta_0 \\ (1-p\beta(1,M,\gamma))(q-q^2)-\mu & ;\beta(1,M,\gamma) < 1-\delta_0. \end{cases}$$

When the policy issue is not preordained, which happens with probability $1-p$, both rulers must be minority-congruent for the final government policy to be $A=1$ in state $s=0$. Thus, the marginal benefit of institutional constraints is $(1-p)(q-q^2)$. When the policy issue is preordained, but the probability of successful revolt is sufficiently high ($\beta(1,M,\gamma) > 1-\delta_0$), the threat of revolt suffices to discipline the minority-congruent rulers and there is not marginal benefit to institutional constraints. However, when the probability of successful revolt is lower ($\beta(1,M,\gamma) < 1-\delta_0$), so that minority-congruent rulers risk revolt, again institutional constraints imply that both rulers must be minority-congruent and the revolution fails for the final government policy to be $A=1$ in state $s=0$. Thus, the marginal benefits of institutional constraint is $p(1-\beta(1,M,\gamma))(q-q^2)$.

Overall, institutional constraints benefit the majority both when the policy issues are preordained and when they are not, but with a higher margin for non-preordained policy issues: $(q-q^2)$ when the policy issue is not preordained and 0 or $(1-\beta(1,M,\gamma))(q-q^2)$ when it is preordained. Thus, a wider scope of the law (higher p) tend to reduce the added value of institutional constraints. We now state our main formal result.

Proposition 4. *There is a threshold $p^*(M,\gamma,q,\mu)$ such that a majority citizen's policy payoff is higher without institutional constraints if and only if the scope of the divine law $p > p^*$, where*

$$p^*(M,\gamma,q,\mu) = \begin{cases} 1-\frac{\mu}{q(1-q)} & ;\beta(1,M,\gamma) > 1-\delta_0 \\ \\ \frac{1}{\beta(1,M,\gamma)}\left(1-\frac{\mu}{q(1-q)}\right) & ;\beta(1,M,\gamma) < 1-\delta_0. \end{cases}$$

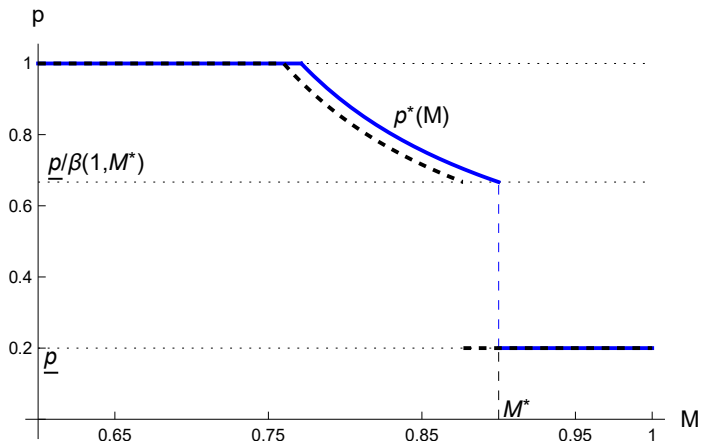

Figure 1: $p^*(M; q, \mu)$, where $\beta(1, M^*, \gamma) = 1 - \delta_0$ and $\underline{p} = 1 - \mu/(q(1-q))$. Parameters: $\mu = 0.2,\ q = 0.5,\ T = 0.6,\ \delta_0 = 0.7$, and $\gamma = 0.9$. The dashed curve corresponds to $\gamma = 0.95$.

Moreover,

1. *If $p^*(M, \gamma, q, \mu) > 0$, then $p^*(M, \gamma, q, \mu)$ is decreasing in M and γ; strictly so if and only if $\beta(1, M, \gamma) < 1 - \delta_0$.*

2. *$p^*(M, \gamma, q, \mu = 0) \geq 1$. For $\mu > 0$, $p^*(M, \gamma, q, \mu)$ has an inverted U-shape in q, with*

$$lim_{q \to 0^+} p^*(M, \gamma, q, \mu) = lim_{q \to 1^-} p^*(M, \gamma, q, \mu) = -\infty.$$

The threshold p^* follows from Corollary 1 and results 1 and 2 follow from the inspection of p^*. The majority can discipline the government to some extent solely by revolt or the threat of revolt. They can also combine this accountability instrument with institutional constraints at a cost. If these costs were negligible ($\mu \approx 0$), they would always do so. When the costs are higher, they must trade off the added benefits of institutional constraints against their costs. These benefits are higher when the scope of the law is narrower (p is lower), when the society is more heterogeneous (M is lower), or when the society has less "solidarity" (γ is lower). All these reduce the effectiveness of the revolt accountability channel by intensifying information and coordination frictions involved in collective action. That is, the marginal net gain from institutional constraints are lower when the revolt channel of disciplining rulers works more effectively. In this sense, institutional constraints and revolt are substitutes. Figure 1 illustrates p^* as a function of the degree of homogeneity in society

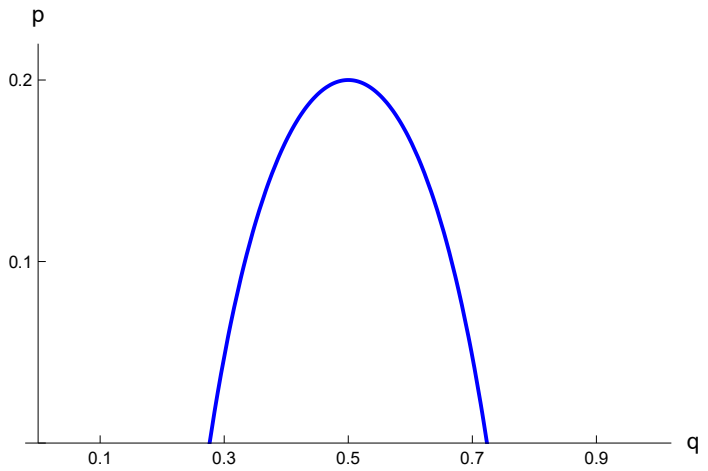

Figure 2: $p^*(q; M, \mu)$ when $\beta(1, M) > 1 - \delta_0$. The case of $\beta(1, M) < 1 - \delta_0$ differs only in scale. Parameters: $\mu = 0.2$.

M. The majority do not set up institutional constraints above the curve, where p and M are higher. Like the effect of homogeneity M, higher solidarity sentiments γ lower $p^*(M)$, as the dashed curve in Figure 1 illustrates.

Increases in the likelihood q that a ruler is bad first raise and then reduce the added value of institutional constraints. When rulers are almost surely good $q \approx 0$ or almost surely bad $q \approx 1$, institutional constraints have little marginal effect. The effect is maximized when there is also maximum uncertainty about the ruler's type $q = 1/2$. Figure 2 illustrates.

Moreover, it may be reasonable to posit that normative traditions view the typical range of q to be $(0, 1/2]$, perhaps closer to $1/2$. After all, citizens are supposed to do everything possible to install a good ruler. It then follows that p^* is increasing in q, so that societies that believe rulers are more likely to be bad, tend to value institutional constraints more. This observation points to another point of departure between the Christian tradition and the Jewish and Islamic traditions. The notion of the Original Sin (formulated by Augustine), prevalent in the Christian tradition, has no counterpart in the Jewish and Islamic traditions. That negative view of human nature may be reflected in having higher likelihood that a ruler is bad, thereby raising the marginal benefit of institutional constraints. The notion that self-interested rational people would, if they got the chance, take actions that were collectively undesirable is deeply embedded in western intellectual traditions (see Ober (2022) for an analysis of how this emerged in the Greek world).

In Proposition 4, we focused on the threshold of the scope of the law p. We can also focus on the threshold of the costs μ. From Corollary 1, we have:

Proposition 5. *There is a cost threshold such that the majority citizen's policy payoff is higher without institutional constraints if and only if $\mu > \mu^*$, where*

$$\mu^*(\beta, p, q) = \begin{cases} (1-p)(q-q^2) & ;\beta > 1-\delta_0 \\ \\ (1-p\beta)(q-q^2) & ;\beta < 1-\delta_0, \end{cases}$$

where $\beta = \beta(1, M, \gamma)$. Moreover,

1. *μ^* is strictly decreasing in p, and weakly decreasing in $\beta(1, M, \gamma)$ (and hence in M and γ); strictly so when $\beta < 1 - \delta_0$.*

2. *Suppose $\delta_0 < T/M$, so that there is sufficient conflict of interest that the threat of revolt does not deter the minority-congruent ruler ($\beta < 1 - \delta_0$). Then,*

$$\frac{\partial^2 \mu^*(\beta, p, q)}{\partial p \partial \beta} = -(q - q^2) < 0.$$

Higher scope of the law p, societal homogeneity M or solidarity γ all improve the majority's ability to control the ruler via revolt channel, thereby reducing the marginal value of institutional constraints, and hence the cost threshold below which they are adopted. Importantly, the second part of Proposition 5 highlights the complementarity between the scope of the law p on the one hand and homogeneity M and solidarity γ on the other (recall that β is increasing in both M and γ). Higher homogeneity and solidarity both increase the likelihood of successful revolution. Higher scope of the law enable majority citizens to better assess whether a successful revolution, which overturns the status quo, will be beneficial. These two channels complement each other: higher scope of the law is valuable because it enables citizens to better know when their rulers deviate from the right policy, but this knowledge helps them if only if they can mobilize, and their mobilization capacity depends on their homogeneity and solidarity.

Increases in the marginal costs of institutional constraints μ obviously tend to reduce their use. For example, in the Roman Republic, in which two consuls shared the highest executive office, the Senate, during times of crisis such as military defeats, sometimes authorized a "dictator" to reduce the costs of institutional constraints, including joint decision-making. Events such as foreign wars and natural disasters tend to raise μ and reduce institutional constraints on rulers. Which conditions are more conducive to the dismantlement of institutional checks due to such events? To glean insights, we introduce uncertainty about γ and discuss the probability of adopting institutional constraints due to an exogenous change in the cost of institutional constraints.

Proposition 6. *Suppose $\gamma \sim U[0,1]$. Let $Q = Pr_\gamma(\mu \leq \mu^*(\gamma))$ be the probability that institutional constraints improve the majority citizen's policy payoff. Suppose $\delta_0 < T/M$, so that there is sufficient conflict of interest that the threat of revolt does not deter the minority-congruent ruler ($\beta < 1 - \delta_0$). Then,*

$$Q(\mu'; M, p) = \begin{cases} 1 & ; \mu' \leq 1 - (1 - T/M)p \\ \frac{1-\mu'}{(1-T/M)p} & ; 1 - (1 - T/M)p \leq \mu' \leq 1 \\ 0 & ; 1 < \mu', \end{cases}$$

where $\mu' = \mu/(q - q^2)$. Moreover,

1. *Q is decreasing in p and in M; strictly so when $\mu' \in (1 - p(1 - T/M), 1)$.*
2. *$|Q(\mu'_2) - Q(\mu'_1)|$ is strictly decreasing in p and M for all $\mu'_2 > \mu'_1$, with $\mu'_1 \in (1 - p(1 - T/M), 1)$.*

Proposition 6 provides insights into the effect of changes in the costs of institutions. It may be more realistic to focus on settings where society's potential for collective action is high enough, so that the society may or may not adopt institutional constraints depending on the level of solidarity γ: $1 - p\beta(1, M, \gamma) < \mu/(q - q^2)$ for $\gamma = 1$. Thus, consider a reduction in the costs of institutional constraints from μ'_2 to $\mu'_1 > 1 - p(1 - T/M)$, e.g., due to peacetime. This drop in costs leads societies with lower levels of γ to adopt institutional constraints. But, as part 2 of the Proposition shows, this change tends to be smaller when the scope of

the law p is larger. This is because solidarity and the scope of the law are complements in disciplining the rulers: the disciplining value of higher γs are larger when the scope of the law p is higher, and hence the added-value (marginal benefits) of institutional constraints are smaller. Therefore, societies with high scope of law are less responsive to a decrease in μ. Conversely, an increase in costs from $\mu_1' > 1 - p(1 - T/M)$ to $\mu_2' > \mu_1'$ will cause the dismantling of institutional constraints by less when in societies with a larger scope of law p. That is, higher scope of law generates inertia in the institutional constraints that aim to control rulers. The same logic applies to the degree of homogeneity of the society M.

Taking the emergence of institutional constraints as exogenous, for a given set of parameter values, the probability of revolt attempt ($p\hat{q}/2$) and successful revolt ($p\hat{q}\beta/2$) are both lower with institutional constraints, where $\hat{q} = q^2$ and q, with and without institutional constraints, respectively. This observation captures Blaydes and Chaney (2013, p.24-5)'s argument that the development of feudalism (and hence some form of executive constraints) in Europe led to its higher political stability compared to the Islamic societies. However, our analysis highlights that the adoption of institutional constraints as means to hold rulers accountable may be the consequence, not the cause, of the ability of the society to mount revolts. More broadly, institutional constraints and political stability (the likelihoods of revolt attempts and successes) arise jointly in equilibrium. To see this, consider two societies W (for West) and E (for East), which are identical in all aspects except the scope of the law, with $p_W < p^* < p_E$. In this case, society W adopts institutional constraints, but society E does not. The likelihood of successful revolt in W is smaller than that in E: $q^2\beta p_W/2 < q\beta p_E/2$. The reason is twofold: (1) conditional on an incongruent government policy, revolt attempts are more likely in society E (revolt are attempted when deviations are observed, which happen with probabilities $p_W < p_E$); (2) the likelihood of deviations (i.e., incongruent government policies) are higher in society E, which has not adopted institutional constraints ($q^2 < q$). But as our analysis highlights, society E may forgo institutional constraints exactly because it is more effective at holding rulers accountable through collective action. The following proposition formalizes this logic. It highlights how the substitutability of revolt and institutional constraints can predict a negative correlation between institutional constraints and political stability (i.e., revolt attempts and successes), both of which are determined

jointly in equilibrium. In the proposition, we focus on the comparative statics with respect to our main variable of interest p, fixing all other parameters (e.g., q, M, γ).

Proposition 7. *Suppose that $p^* \in (0,1)$ and that $\delta_0 < T/M$, so that there is sufficient conflict of interest and the threat of revolt does not deter the minority-congruent ruler ($\beta < 1 - \delta_0$). Focusing on the scope of the law p as the only source of variation, the equilibrium probabilities of revolt attempts and successful revolt are both lower in societies with institutional constraints. Formally,*

$$\mathbb{E}[pq/2 \mid p > p^*] > \mathbb{E}[pq^2/2 \mid p < p^*] \quad \textit{and} \quad \mathbb{E}[pq\beta/2 \mid p > p^*] > \mathbb{E}[pq^2\beta/2 \mid p < p^*],$$

for a given q and $\beta = \beta(1, M, \gamma)$.

The endogeneity of institutional constraints is captured in the conditioning on the subset of parameters in which institutional constraints are or are not adopted (i.e., the added-value of institutional constraints exceeds their cost, $p > p^*$, or not, $p < p^*$). As we discussed in the Introduction, these results are consistent with the empirical observations in the literature on Islamic civilization (Blaydes and Chaney, 2013; Finer, 1999) and with the broader literature on the positive correlation between institutional constraints and political stability (Besley and Persson, 2011).

Of course, multiple causes can contribute to this systematic difference in stability. For example, Auriol and Platteau (2017) and Platteau (2017) argue that the decentralized structure of the clergy in Islam, compared to Christianity, made bargaining between rulers and the clergy less effective in avoiding conflict through the co-optation of the clergy who oppose the rulers' reforms. More broadly, because some form of co-option of the clerics also appear in Rubin (2017), Bisin et al. (2023), and Auriol et al. (2023), we next discuss how our setting can accommodate such actions.

To capture the rulers' attempts to co-opt the law or the jurists, suppose that rulers can pay a cost to reduce the probability p that citizens observe the state. These costs include co-opting the jurists to overlook transgressions, or offer an interpretation of the law, when possible, to support the rulers. Of course, jurists (even in a centralized setting) cannot make interpretations or take positions that too starkly contradict the broad understanding

of the law—else, they lose their support and risk condemnation by other clerics. However, occasionally there is room for overlooking the law or providing non-standard interpretations which may benefit the rulers. Recalling that citizens revolt when the state is $s = 0$, they observe the state $\hat{s} = s$, and the government takes action 1, the marginal benefit of reducing p to the (incongruent) government is $\beta(1, M, \gamma)/2$, both without and with institutional constraints—when both rulers are bad. The government then chooses an appropriate expenditure that equates the marginal costs and benefits of reducing p. In this extended model, co-option translates into a smaller scope of the law. One may further relax the assumption that the costs of reducing p is the same with and without institutional constraint. This, in turn, further increases the added value of the institutional constraints. Either way, a higher scope of the law reduces the added-value of institutional constraint, which is the core of our argument about Islamic and Western normative traditions.

3 Institutional Constraints on Rulers in the Islamic Tradition

As the prophet, Mohammad (d. 632) was the leader (*imam*) of the Islamic community (*umma*). The Constitution of Medina also recognizes Mohammad as the ultimate judge and arbitrator in case of disagreements among the members of the *umma* (Watt, 2003, p.130-4; Lecker, 2004). The tribal nature of early Muslim society and Mohammad's emphasis on building consensus through consultation (*mashwara/shūrā*; e.g., the Quran (3:159; 42:38)), combined with his prophetic charisma, would alleviate concerns about the concentration of coercive power. Upon Mohammad's death, Abū Bakr (d. 634) was selected as the next leader in a gathering of a group of prominent members of the Medinese Muslim community, and he adopted the title of *caliph* (*khalīfa*, meaning "successor" or "deputy"). Before his death, Abū Bakr designated 'Umar (d. 646) as his successor. 'Umar, in turn, designated a small group of notables (*shūrā*) to select the third caliph, and so 'Uthman (d. 656) was selected.[14] However, this semi-egalitarian structure changed soon after. The rebellion and

[14] Crone (2001, p.3) traces early references to *shūrā* "as a procedure for deciding who should be in charge of the government". Stasavage (2020) takes the ideas and practices of *shūrā* as evidence of "early democracy" in

killing of the third caliph, 'Uthman, led to a crisis, which evolved into the First Civil War (656-661) during the rule of the fourth caliph, 'Ali (d. 661). In turn, 'Ali was assassinated and his challenger Mu'awyah (d. 680), a kinsman of 'Uthman and the governor of Syria, became the next caliph (Madelung, 1997; Crone, 2004).

Concerns about tyranny became widespread under Mu'awyah who established hereditary succession and thereby the Umayyad dynasty. By the late 7th century, the fourth Umayyad caliph 'Abd al-Malik (d. 705) "wanted his subjects to believe that the power and the kingship... was a possession... granted by God and inalienable according to the divine will (Lambton, 1981, p.46; see also Black, 2011, p.18, and Donner, 2011, p.82-4). The title *caliph* referring to the deputy of God (*khalīfat Allāh*), as opposed to the deputy of God's messenger (*khalīfat rasūl Allāh*), appeared on coins for the first time in 'Abd al-Malik's reign (Anjum, 2012, p.47); The policies of the Umayyad caliphs and their sumptuous lifestyle were sharp departures from the behavior of Mohammad and his immediate successors. For example, the second Umayyid caliph, Yazid ibn Mu'awyah, "is often considered the epitome of injustice, impiety, and corruption" (Abou El Fadl, 2001, p.117). Various revolts broke out over "the Umayyad manner of distributing revenues... maltreatment of the Prophet's family, tyranny and the like" (Crone and Hinds, 1986, p.64).

However, we have no record of discussions about institutional constraints on rulers in that period. This puzzling absence persists during the Abbasids, who replaced the Umayyads in 750, and through various dynasties and kingdoms in the following millennium.[15]

To establish this puzzle, following Rosenthal's (1971, p.17-33; see also Lambton, 1981) classic categories, we divide political writings in Islamic civilization into three groups, depending on whether their primary foundation is Islamic law, philosophy, or advice-giving in

the Islamic community, inherited from the pre-Islamic Arabia, which was overturned when the Islamic state adopted the elaborate Sassanian bureaucratic apparatus of the conquered regions. However, in his history of the early Islamic state and its political and military structures, Donner (1981) highlights the emergence of a ruling elite from the early years. For our purposes, different notions of *shūrā* did not develop (in theory or practice) into institutional constraints on rulers once they assumed power, via *shūrā* or other mechanisms. Of course, rulers were encouraged to consult the community and religious scholars.

[15]There are hints of institutional constraints in the reported statements of a few individuals, e.g., *al-Ḥārith ibn Surayj* (d. 746), a rebel leader against the Ummayads, or *al-Aṣamm* (d. 816/7), a Mu'tazilī theologian. However, these sparks did not turn into any coherent discussions (Crone, 2004, p.277-8)—see also van Ess (*Encyclopaedia of Islam, 2nd Ed.*), Crone (2000), and Stern (1970).

the manner of Mirrors of Princes.[16] We provide brief discussions of a few well-known examples in each category to touch on the political themes that Muslim thinkers engaged with and to demonstrate the absence of discussions about institutional constraints on rulers. Such discussions are also absent in comprehensive surveys of Islamic political thought (Rosenthal, 1958; Lambton, 1981; Crone and Hinds, 1986; Crone, 2004; Black, 2011; Cook, 2014).

Obviously, the corpus of Islamic writings with direct political implications is vast. For example, the above categories do not include the writings and traditions of mystic orders that sometimes had direct political implications (Babayan, 2002; Ziad, 2021). However, mystic orders with their emphasis on the spiritual (and sometimes temporal) leaders with divine inspiration tended to be even less concerned with institutional constraints.

3.1 Juristic Writings

Jurists insisted that caliphs must follow Islamic law and resisted attempts by caliphs to modify the law. In this narrow conception of constitutionalism (Klosko, 2012, p.297-9), they were "constitutionalists" (Watt, 2003). For example, when the caliph Hārūn al-Rashīd commissioned his chief judge Abu Yusuf (d. 798) to review taxation, the resulting *Kitāb al-Kharāj* detailed Islamic law "for the rates of taxation and the expenditure of the revenue according to the source from which it derived" (Lambton, 1981, p.55). Shāfi'ī's (d. 820) emphasis on the role of *ijmāʿ* (consensus) as a source of Islamic law (Bernand, *Encyclopaedia of Islam, 2nd Ed.*) further reduced the legislative power of caliphs: "To al-Shāfi'ī the ultimate arbiter was the consensus of the entire community: the caliph counted only so far as every member of the *umma* did" (Crone and Hinds, 1986, p.93). Hallaq (2009, p.70) goes as far as arguing that "[w]hereas law – as a legislated system – was often 'state'–based in other imperial and complex civilizations, in Islam the ruling powers had, until the dawn of modernity, almost nothing to do with the production and promulgation of legal knowledge."

Of course, rulers tried to give themselves more freedom to change or reinterpret the law in ways that suited them. An example of such an attempt is the inquisition (*mihna*), which

[16] Many writings have multiple elements. For example, religious concerns and orthodoxy are intertwined with politics, justice, and stability in Nizam al-Mulk's *Sīyāsat Nāmih*, in contrast to Machiavelli's instrumentalist approach to Christianity. However, the dominant theme of each work is typically clear.

was started in 833 by the Abbasid caliph al-Ma'mūn (d. 833) and lasted until 848/9. During the inquisition, religious judges and scholars were pressured to accept the doctrine that the Quran was created (Lapidus, 1975, p.380). This theological point had critical implications. Watt (2003, p.88) argues that "if the Qur'ān was created, God could presumably have created a different Qur'ān in other circumstances. Or... God's plenipotentiary, the imam or the charismatic head of the state, acting with divine authority, could set aside... specific commands of the Qur'ān and, more generally, the provisions of the Sharī'a" (p.87-88). According to Lapidus (1975, p.380), during the inquisition, "the theological opposition is clearly linked to popular demonstrations against the policy of the regime".

Some jurists specified conditions under which a sitting ruler could be deposed, and some such depositions "may have been accompanied by a formal *fatwa* [a legal opinion issued by a jurist] authorizing it on various moral or religious grounds" (Gibb, 1982, p.161)—see Section 4.2. Critically, while jurists discussed conditions that disqualify a sitting imam, they did not "lay down any procedure by which an Imam may be deposed" (Gibb, 1982, p.161) (imam/imamate and caliph/caliphate are virtually the same in our discussions). According to Lambton (1981, p.19), "a command contrary to *sharī'a* was not to be obeyed. The jurists, however, did not specify in what way or by what tribunal it was to be decided that the leader of the community had failed to remain faithful to the *sharī'a*". There were some procedures for the election of caliphs in theory – see below. However, as Crone (2004, p.277) argues, "once elected, the caliph was free to ignore all the advice he received."

Systematized juristic formulations of government authority appeared in the High Middle Ages. An example is Mawardi's (d. 1058) *al-Aḥkām al-Sulṭāniyah* (The Ordinances of Government), which provides a theory of the caliphate. Like his Sunni and proto-Sunni predecessors, Mawardi viewed the institution of imamate as obligatory for the Muslim community. He argued that the imam should be elected by qualified electors from qualified candidates. However, even one elector could suffice; the previous imam could designate the next one, or limit the candidate pool. Mawardi's main contribution was to offer a juristic theory of imamate, partly based on the principles of necessity and expediency, that would allow for a caliph with little de facto power to remain as caliph, so that the umma is not left without an imam while the de facto power remains with various Muslim dynasties that

ruled different Islamic regions—e.g., the Buyids (945-1055). His theory was a response to a crisis in the theory of caliphate. On the one hand, the institution of imamate was obligatory and it was necessary for the fulfilment of various Islamic laws. On the other, the Abbasid caliphs were at the mercy of the Buyid rulers. Thus, Mawardi developed the notion of *imārat al-istīlā*, in which "the governor of a province, instead of being appointed and revocable by the caliphs, imposes his rule by force" (Gibb, 1982, p.162). As long as such rulers maintained their allegiance to the caliph and upheld Sharia, Mawardi argued, the caliph must validate and legitimize their government. An imam/caliph could coexist with various kings/sultans.

Importantly for our purposes, Mawardi detailed conditions that disqualify a sitting imam, but without offering any institutional procedure for deposing him. Neither did he suggest institutional procedures for holding regional kings accountable. Even if a sultan acted egregiously, all Mawardi offered was that "it becomes the duty of the Imam to call to his aid those who will restrain the usurper's hand" (Gibb, 1982, p.160). Like other jurists, he seems to have preferred rebellion and war as the instruments of accountability. Baghdadi (d. 1037), Juwayni (d. 1085) and Ghazali (d. 1111) provided variations of these themes—see Hallaq (1984) on Juwayni's and Hillenbrand (1988) on Ghazali's political views.[17]

The next major development occurred after the fall of the Abbasid caliphate to the Mongols. The problem was to provide an Islamic foundation for the government without the caliphate. Ibn Taymiyya (d. 1328), in *al-Sīyāsa al-Shar*ʿ*iyya*, argued that the legitimacy of government lies in upholding sharia, encompassing obedience to God and the prophet and establishing peace and prosperity. So there could be many imams and governments in different parts of Islamic lands, established by force or some form of election. He viewed the "oath of allegiance" between the community and the imam as a form of contract. "It was, he states, a contract and was, like all contracts, defined by its end, which was the common will to obey God and His prophet. Also, like all contracts, it presupposed two parties: on the one side there was the *imam* and on the other Ibn Taymiyya sets not only the *'ulama'* [(jurists)] but all those who by their learning, talent, wealth, or personal influence actually held authority over the community" (Lambton, 1981, p.148). Mutual obligations included the obedience of

[17]They disagreed on details. For example, while Baghdadi (d. 1037), allowed for multiple imams on distant lands with no overlapping jurisdictions, Mawardi insisted on there being a sole imam at a given time.

the community, but also their duty to offer council to the ruler, enjoin the good and forbid evil. Conversely, the ruler was obligated to consult the community, uphold justice (including just division of government funds), and appoint good officials. In his view, Lambton (1981, p.149) argues, government was to be "a harmonious association of complementary qualities which had originally been centered in one person and which were indispensable for the perfect functioning of the state. He conceives such co-operation as existing between the *'ulama'*, the depositories of the law and the *'umara'* [(rulers)], the holders of political power." Whether the government was a contract or "a harmonious association" between the ruler and the community, Ibn Taymiyya did not propose institutional constraints on rulers to hold them accountable. Khunji (d. 1521) provided a variation of these themes, e.g., recognizing as an imam someone who establishes the government by military force, without any form of election or expression of allegiance on the part of the community—effacing the distinction between imams/caliphs and kings/sultans (Lambton, 1981, Ch. XI).

By the 16th century, following the Ottoman conquests, the title of caliph was claimed by Ottoman kings. Jurists obtained significant power in the Ottoman Empire by the 17th century. Islamic law was invoked in matters of "succession, the legitimacy of a particular sultan, and the question of legitimate revolt against the government in a manner and frequency unmatched in the history of the Islamic world before the Ottomans" (Tezcan, 2010, p.237). For example, jurists were involved in the rebellions against Othman II in 1622, Ibrahim in 1648, Mehmed IV in 1678 (Tezcan, 2010), and Mustafa II in 1703 (Abou-El-Haj, 1984). In fact, Abdurrahim, the grand mufti [(chief jurist)], "gave the legal opinion that legitimized the regicide [of Ibrahim], and oversaw the execution personally" (Tezcan, 2010, p.220). Abou-El-Haj (1984, p.71-2) provides legal opinions (*fatwā*s) by Ottoman jurists, which legitimized revolt against an unjust imam/caliph/sultan/king and rendered supporting such an unjust imam against the rebels unlawful. Of these legal opinions, Aksan (2022, Ch. 2) describes, "The second legitimated the right of a Muslim community to stand up to an unjust ruler. The third condemned those who sided with an unjust ruler." The ability to constrain the sultan through rebellion, Tezcan (2010, p.238) argues, contributed to the longevity of the dynasty: "If an emperor could be 'recalled' and replaced by another one, not only was there no longer any need to challenge the dynasty but... there was also a considerable incentive

to keep the dynasty in operation to maintain its openness to political representation". Ottoman jurists, like their predecessors, seem to have preferred the "revolt channel" over more institutional means of accountability.

Of course, there were various branches of Islam. Our focus on Sunni Islam reflect its predominance in Islamic history. In some regions (e.g., Iran), variations of Shia Islam became prevalent with Shi'i rulers controlling political power. Proto-Shi'i and Shi'i jurists, with their various divisions, believed in the divine spiritual and temporal mandates of their imams (Modarressi, 1993; Dakake, 2007). Given this theology, their minority position, and messianic beliefs, Shi'i jurists did not contemplate institutional constraints on rulers. A tenet of the Twelver Shia is that the 12th imam, who has the public authority, is in occultation since the 10th century. Thus, through the 19th century, to the extent that Twelver Shi'i jurists engaged in developing political thought, their focus was on justifying governmental authority in the absence of the 12th imam. Most jurists refrained from discussing political authority, some argued for a Shi'i kingship (e.g., Majlisi (d. 1699)), and a group assigned political authority to Shi'i jurists (e.g., Naraqi (d. 1829)). However, they did not contemplate institutional constraints on rulers, whatever their identity, until the early 20th century, when a sense of crisis, that of imminent colonial domination compelled juristic justifications for constitutional constraints (Arjomand, 1984, 1988; Amanat, 2009; Ansari and Shadmehr, 2021).

3.2 Philosophical Writings

"From about the middle of the eighth century to the end of the tenth, almost *all* non-literary and non-historical secular Greek books that were available throughout the Eastern Byzantine Empire and the Near East were translated into Arabic" (Gutas, 1998, p.1). Muslim philosophers adapted the Greek tradition but aimed to make it compatible with Islamic teachings. Thus, for Farabi (d. 950), who is sometimes called the founder of Islamic philosophy, "Religion is an imitation of philosophy... In everything of which philosophy gives an account based on intellectual perception or conception, religion gives an account based on imagination" (Lerner and Mahdi, 1963, p.77). Later Muslim philosophers, most notably Ibn Sina (d. 1037) and Ibn Rushd (d. 1198), further synthesized the Greek tradition with Islamic philosophy; see Goodman (1992) on Ibn Sina, and Leaman (1988) on Ibn Rushd.

In *al-Madīnah al-Fāḍilah* (Virtuous City), Farabi developed a political theory reminiscent of Plato's *Republic*. Thus, "the founder of a virtuous city was a person endowed with an exceptional set of outstanding characteristics... Such a man, the first or ultimate chief... was imam, king, philosopher, and prophet alike. In short, he was Plato's lawgiver and the prophet of the Islamic tradition rolled together" (Crone, 2004, p.178; Rosenthal, 1958, p.128); see also *al-Sīyāsat al-Madaniyyah* (known as *Political Regimes*) (Farabi and Butterworth, 2015).

"What is missing in al-Farabi is any concept - let alone discussion - of civic institutions as central to the political life" (Gutas, 2004, p. 276, 263-4; quoted in Black, 2011). Absent an Islamicized "philosopher-king", lesser leaders should take charge, which necessitates the memorization of laws laid down by the founder. Farabi recognizes that "various qualities that went to make a first chief might also be dispersed in many people; if so, they could take the place of the first chief and rule as a team. This was how al-Farabi understood aristocracy (*riyasat al-afadil/al-akhyar*): a virtuous regime in which several philosophically trained people managed things together, perhaps as king, vizier, military leader, and advisors, though he does not say precisely how" (Crone, 2004, p.179). Critically, this allusion to conciliary government did not invoke a discussion of how dividing power could reduce its abuses.

Ibn Sina (d. 1037) and Ibn Rushd (d. 1198) generally followed Farabi. Ibn Sina "adopted a Sunni view on *succession* to Prophet-Legislator. This can be either by testamentary designation – the 'Abbasid practice – or by 'consensus of the elders'. Ibn Sina recommended designation because it avoids strife" (Black, 2011, p.75, quoting Lerner and Mahdi, 1963). However, he "roundly condemns usurpation, and actually demands the death of a tyrant (*mutaghallib*) and the punishment of those who fail to carry out such a tyrannicide if they have means to do it" (Rosenthal, 1958, p.153).

Ibn Rushd's political theory appears in his commentaries on Plato's *Republic* and Aristotle's *Ethics* and *Rhetoric* – he states that he did not have access to Aristotle's *Politics* (Averroes and Lerner, 1974, p.4). Ibn Rushd's commentary on Plato's *Republic* adapts it to his Islamic environment – with the remarkable exception of promoting women's participation in public life. Thus, in discussing the philosopher-king he writes: "Hence these names are... synonymous – i.e., 'philosopher,' 'king,' 'Lawgiver'; and so also is 'Imam,' since *imām* in Arabic means one who is followed in his action. He who is followed in these actions by

which he is a philosopher, is an Imam in the absolute sense" (Averroes and Lerner, 1974, p.72). He provides examples from his environment to demonstrate regime types and their transformation, e.g., stating that during Mu'awyah the government transformed from virtuous to timocratic (Averroes and Lerner, 1974, p.121). Like Farabi, Ibn Rushd recognizes that characteristics of a good ruler may not all be present in one person: "However, it may not happen that both these [qualifications] are found in one man, rather the one [capable of] waging Holy War being another than the legal expert. Yet of necessity both will share in the rule, as in the case with many of the Muslim kings" (Averroes and Rosenthal, 1966, p.208-9). In his commentary on Aristotle's *Rhetoric*, Ibn Rushd also "Platonizes" Aristotle's brief discussion of political regimes by elevating "rulership of the king" and "imamate" (Butterworth, 1998, p. 236-7; Averroes and Ezzaher, 2015, p. 129-131). Again, allusions to division of power or discussions on the transformation of political regimes from Greek political thought stopped short of discussions about institutional constraints to check tyranny.

In sum, Farabi, Ibn Sina, Ibn Rushd, and many other Muslim philosophers, while engaged in political theory, did not discuss institutional mechanisms to constrain rulers.

3.3 Mirrors for Princes (*Sīyāsat Nāmih*)

Among the earliest survived political writings of the Islamic period are *Rasālih fi al-Ṣaḥābih* and *Adab Kabīr* by Ibn Muqaffa (d. 759), an Iranian bureaucrat and literary figure in the Umayyad and Abbasid caliphates. Ibn Muqaffa asserted that the general public cannot obtain their welfare on their own, and they need an imam to guide them. He advocated that the ruler imposes consistency in law and argued that the imam's opinions and policies must be followed unless they explicitly contradict God's orders. Ibn Muqaffa was among the early transmitters of political writings in the Persian mirrors for princes traditions, which aimed to rationalize government to promote peace and prosperity.[18]

A classic writing in the Mirrors of Princes tradition is Nizam al-Mulk's (d. 1092) *Sīyāsat Nāmih*. Nizam al-Mulk was an Iranian vizier during the Seljuk Empire and the de facto ruler after Alp Arslan. In the tradition of Persian kingship, he asserts that just kings are

[18]Mirrors for Princes were also translated from Greek during the Umayyads when Greek was the language of bureaucracy in Damascus, adapted from the Byzantine Empire (Gutas, 1998, p.23).

chosen by God, have royal charisma, and are the shepherds of their people. When God becomes angry with the people, good kings disappear, and war replace peace and prosperity, so that wrongdoers are killed. He highlights that power is fragile, because of overt and covert contenders, but a competent and just ruler maximizes peace and prosperity. The ruler should follow God's law and respect religious scholars and the pious. He should have religious scholars advise him "once or twice a week" on God's law and Islamic traditions and stories of past just kings (Tabatabai, 2006/1385, p.97) – see also Yavari (2014).

Some works in this genre rely more on earlier philosophical writings (e.g., Tusi's *Akhlāq Nāṣirī*), while some have more religious overtone (e.g., Ghazali's *Naṣīḥat al-Mulūk*). Overall, this genre is a middle ground between theoretical works on ethics and moral philosophy and manuals for governance.[19] While offering useful advice on good governance, like juristic and philosophical writings, these works do not provide discussions of institutional constraints on rulers. Rather, they rely on invoking the rulers' intrinsic motivations (moral or selfish interests in prosperity and longevity) to achieve good governance.

4 Islamic Law, Rebellion, and Accountability

The previous section provided evidence for establishing the puzzle of missing discussions. We also suggested an explanation for this puzzle. This section provides further evidence for the key assumptions that our explanation relies on.

Two anecdotes highlight the key features of the explanation. The first caliph Abū Bakr, in his speech upon assuming leadership, stated: "I have been given the authority over you, and I am not the best of you. If I do well, help me; and if I do wrong, set me right... Obey me so long as I obey Allah and His Messenger. But if I disobey Allah and His Messenger, you owe me no obedience" (from Ibn Hisham's *Sīrah* quoted in Cook, 2014, p.320). 'Umar, the second caliph, "asks that anyone who sees any crookedness in him should tell him; a distinguished Companion of the Prophet responds that in that event 'we will straighten you out with our swords,' a sentiment to which ʿUmar responds with strong approval" (Cook, 2014, p.320).

[19] Other examples includes Davani's *Akhlāq Jalālī*, Amasi' *Kitāb Mirʾāt al-Mulūk*, Bitlisi's *Hasht Bihisht*, and Çelebi's *Akhlāq ʿalāʾī*; see Sariyannis (2019) for the adoption of this genre in the Ottoman period.

These anecdotes show (1) the presence and importance of Islamic law (reflected, e.g., in "obey Allah and His [deceased] Messenger"); (2) the deviations from the law are presumed clear and easily observable; (3) rebellion against a caliph who does not follow the law is not only allowed, but also encouraged; (4) Muslims are vigilant to hold rulers accountable through revolt if needs be. According to Cook (2014, p.320-1), this reflects "a political culture in which it is not just conceded that subjects are entitled, and perhaps obligated, to act in such ways; they are portrayed as ready to do so at the drop of a hat".

We now provide more evidence for these themes. In Section 4.1, we ask: Can we even speak of Islamic law given that there are always disagreements over details and different interpretations? To what extent can we speak of Islamic law as independent from the political, social, and economic environment? Did Islamic law have a significantly larger scope than, say, the divine law in Christian tradition? Importantly, is there evidence that Muslim thinkers believed that there was such a thing as Islamic law and that it had a wide scope? In Section 4.2, we ask: What was the normative view on revolt in Islam? Did it happen often or it was rare? Did distinguished figures in Islam participate in rebellions? What, if any, was the Islamic law of rebellion? Do we have evidence that some rebels invoked the violation of the law as a key reason for their revolt? Finally, in Section 4.3, we discuss some cultural elements, moral themes, and religious obligations that would imply a high degree of solidarity.

4.1 Islamic Law

Most Muslim scholars took as given that individuals can only infer or interpret Islamic law, so that pious and knowledgeable scholars can have legitimate disagreements over details. But the range of interpretations, while it surely would evolve in the long run, was relatively narrow and stable in the short run. Islamic law was not monolithic. However, these differences were small compared to potential differences in laws that could be. Hence the disagreements took place within a highly constrained space. Despite geographical variations (perhaps most apparently between Sunni and proto-Shi'i or Shi'i societies), the relatively narrow range of acceptable interpretations in a given region and period projected a coherent notion of Islamic law, from Andalus (as Ibn Rushd's quote below indicates) to India. For example,

"[d]escribing the late Mughals of India, the eighteenth-century English scholar Alexander Dow observed that the Sharia 'circumscribed the will of the Prince' and 'the House of Timur always observed [the law]; and the practice of ages had rendered some ancient usages and edicts so sacred in the eyes of the people, that no prudent monarch would choose to violate either by a wanton act of power' " (Hallaq, 2009, p.211). As late as the 20th century, jurists' chief concerns about the codification and unification of laws in Iran and the Ottoman Empire was that the codes conform to Islamic law. For example, Article 2 of the Supplementary Laws to the Iranian (1906) Constitution, proposed by Shi'i jurists, required that a few jurists supervise the laws passed by the parliament to ensure their consistency with Islamic law (Bayat, 1991; Afary, 1996). Later, jurists such as Modarres were involved in drafting unified civil and criminal codes and Kharaqani took the initiative to codify 831 items of Islamic law, offering the resulting booklet to the government (Jafarian, 2003/1382). All the quibbling among jurists was slim next to the wide range of potential alternatives.

Crucially, that was how Muslim scholars, jurists, and philosophers perceived their environment, even a cosmopolitan philosopher such as Ibn Rushd in Andalus. For example, comparing the nature of law in Islamic and Christian societies in his commentary on Aristotle's *Rhetoric*, Ibn Rushd wrote (Averroes and Ezzaher, 2015, p.130):

> Perhaps the laws instituted in these cities were definite, invariable, and permanent, as in the case of our Islamic law. And perhaps these cities did not have definite laws, but the matter was delegated to those who held the power, depending on what was more useful at each moment, as in the case of Byzantine laws.

Similarly, Ibn Khaldûn (2015, p.189-90) states:

> The religious laws govern all (governmental positions) and apply to each one of them in all its aspects, because the religious law governs all the actions of human beings. Jurists, therefore, are concerned with the rank of ruler... and with the conditions under which it is assumed... Furthermore, (they are concerned with the causes) that necessitate (the ruler's) removal, should (such causes) present themselves, and with other things connected with the ruler or sultan.

Ibn Sina, Mawardi, Juwayni, Ghazali, among others had similar (or more rigid) views.

This perception was rooted in reality. Rulers, even imams/caliphs, could not easily change the law. At best, they could assert their superior understanding—often with little success. Rulers attempted to define what divine law was, as evidenced by the inquisition project of the Abbasids discussed in Section 3.1. However, these attempts failed. Even jurists who preached against rebellion insisted that rulers cannot determine the law: caliphs must implement God's law, not theirs. "By locating the power to legislate outside the political system, it [(Islamic law)] denied to rulers the ability to make law to suit their fancies. It is thus a significant point about the Sharīʿa that...it is in principle the antithesis of the legislative autocracy or a traditional patrimonial state or a modern dictatorship" (Cook, 2014, p.329-30). Namik Kemal, during the Ottoman Tanzimat period, stated that "even the greatest tyrants cannot alter" Sharia for it is protected by God (Mardin, 1962, p.315, cited in Cook, 2014, p.330). This perception is a persistent feature of Islamic tradition that continues to modern times across branches of Islam. In his lectures on Islamic government, Khomeini (2006/1385, p.72-3) stated that "Islamic government is the government of laws...the law is the real ruler". Some surveys suggest that many contemporary Muslims in Turkey, Iran, and Egypt still believe that Sharia "limits the power of rulers" (Rheault and Mogahed, 2008); the law institutes rulers, not vice versa.

Even jurists could not alter the range of interpretations more than marginally at a given time. There was no hierarchical ecclesiastical structure as in Christianity. There were many scholars at a given time, and eventually a well-established juristic culture. Rulers could bribe or threaten judges to look the other way, or to misrepresent the facts, but they had extreme difficulty to rewrite the Islamic law. As Abou El Fadl (2001) shows in the context of the Islamic law of rebellion (which rulers had great interest to influence), "Once legal precedent is set, and the legal culture becomes institutionalized and developed, legal doctrines often assume a life of their own. These legal doctrines set their own base of authority and their own doctrinal imperative" (p.162). Jurists and rulers had some common interests "but this did not mean that the jurists simply became the ideologues of the state. The emerging corporate or institutional culture of the jurists demanded that order and stability be maintained, but once the precedents of the law of rebellion had come into existence, these precedents became

an imperative force by themselves" (p. 187). As Kuran (2023, p.278) argues, "the freedom to interpret Islam was bounded".

Islamic law had a wide scope. It covered "subjects such as taxation, the conduct of holy war, the suppression of rebels, the punishment of criminals, and the appointment of judges... The law left much to the discretion of rulers, but its letter was often detailed and its spirit was unmistakably protective of the believers" (Crone, 2004, p.282). Crone and Hinds (1986) argue that the Abbasid caliphs "found that the past which they were supposed to imitate consisted of narrowly defined rules, not the ancestral practice compatible with any interpretation they might wish to put on it. In practice, their hands had thus been tied... The law was the sum total of God's guidance... it dealt with every aspect of life from taxation to the proper way of wearing moustaches" (p.92-3, see also p.109-110). Legitimate public policy then was restricted, or it was so perceived, by the wide scope of Islamic law, which covered topics from taxation, inheritance, and family laws to tort and contract laws.

Hallaq (2009, p.551-5) provides a breakdown of topics in Islamic law books, covering 57 topics, including *zakāt*, various contracts, tort, and rules of procedure such as testimonies. Abu Yusuf's *Kitāb al-Kharāj*, mentioned above, was an early example of Islamic law on taxation. Modarressi (1983) provides a detailed description of the origins of *kharāj* in Islam and the jurists' opinions about its justification and rate, lands subject to it and the expenditure of the revenues. As Hallaq (2014, p.62) argues, "the benchmark of taxation was the Sharʿī-stipulated rates... In other words, taxation could be determined by fixed and objective criteria, and thus overtaxation was relatively easy to evaluate and dispute in a Sharʿī court". Based on Johansen (1988)'s study of land tax in Islamic law, Khoury (1997, p.179) argues that in the Ottoman Empire, the "sphere of action of the sultan was at all times confined within the parameters of a concept of justice which ensured the rights of the proprietor against the absolute and ultimate control by the sultan". Even in the 20th century, during the Iranian Constitutional Revolution, it was repeatedly argued: "it is obvious that our Divine Law is not limited to acts of worship but, on the contrary, embraces every major and minor political issue, down to the indemnity for a minor abrasion. Consequently, we will never be in need of man-made law" (Dabashi, 1988, p.361-2).

4.2 Rebellion and its Status in Islamic Law

The speed and magnitude of Islamic conquests and the relative prosperity of Islamic civilization for several centuries may give the impression of political stability in the Islamic polity. This is a false impression.[20] Revolt by Muslims against Muslim rulers was ubiquitous throughout much of the Islamic history. There were revolts against Abu Bakr (d. 634) (Donner, 1981, p.82-90), and against the third and forth caliphs, 'Uthmān and 'Ali (both of whom were killed), including the Battle of Siffin, Battle of Nahrawan, and the Battle of the Camel, in which different groups of Muslims fought with each other. This First Fitna (civil war) led to the establishment of the Umayyad Caliphate in Damascus. The Second Fitna (680-692) began two decades later and included the revolts of Husayn Ibn Ali, Tawwabin, Mukhtar, and Abd Allah ibn al-Zubayr. The third Fitna was another civil war in the 740s, which blended into the fourth Fitna, the Abbasid Revolution and the establishment of the Abbasid Caliphate in 750. There were many other revolts by Muslims and also by non-Muslims throughout the Islamic empire (e.g., see Wasserstrom (1995, Ch. Two) for Jewish revolts, and Crone (2012) for Zoroasterian-inspired revolts).

Rebellions continued throughout Islamic history though with some relatively stable periods, especially under the Ottomans and Safavids. Abou El Fadl (2001) mentions many examples and states that "there is hardly any period in Islamic history that was not plagued by rebellions" (p.107). Rebellions are common occurrences in Islamic history books, from the general histories of Tabari (d. 923) and Ibn Khaldun to specialized histories in the genres of *Maqātal* (e.g., Abu al-Faraj al-Isbahānī's (d. 967)) or *Ṭabaqāt*.

Rebellions can form from coalitions of different groups with different grievances. These grievances can also vary, including one or a combination of ethnic or religious conflicts (e.g., Zoroastrians, Hindus, or Jews may rebel against Muslim rulers), tribal or individual conflict over rulership or succession, and gross violations of the law by rulers. Islamic history seem to feature all these combinations. For our purposes, some statements of rebel leaders on the agenda of their revolts have been preserved, in which they highlight gross violations of Islamic law the reason for their rebellion. We provide three such examples, focusing on the

[20] Crone (2012, p.17) cleverly worded this impression as: "how long can a tiny minority be expected to hang on to power in a foreign land if it fights itself every thirty years?"

classic, Tabari's *History*. The following is from Husayn ibn Ali's speech during his rebellion (Ṭabarī, 1990, vol. 19, p.95-6; see also his letter to Basrans on p.32):

> People, the Apostle of God said: 'When anyone sees the authorities make permissible what God had forbidden, violating God's covenant, and opposing the Sunnah of the Apostle of God by acting against the servants of God [people] sinfully and with hostility, when anyone sees all these incidents and does not upbraid them by deed or by word, it is God's decree to make that person subject to [mis]fortune.' Indeed, these authorities... have neglected the [religious] punishment (*ḥudūd*) laid down by God; they have appropriated the *fayʾ* [war booty] exclusively to themselves; they have permitted what God has forbidden, and they have forbidden what He has permitted.

Ṭabarī (1989, vol. 26, p.37-8) records an episode in the Zayd's rebellion, when some accuse Zayd of rebelling to seek power. Zayd responds that Abu Bakr and Umar "behaved justly with the people and acted according to the Qurʾān and the *sunnah*." But Umayyad rulers "are tyrannical to me, to you, and to themselves. We are only summoning you to the Book of God and the *sunnah* of His prophet so that God's ordinances (*sunan*) may be revived and innovations (*bidaʿ*) may be wiped out."

During the movement of "commanding right and forbidding wrong" in Iraq, one of the leaders, Sahl ibn Salāmah would explicitly state: "I shall attack anyone who opposes the Book and the *sunnah* whoever it may be, the government authority itself or anyone else" (Ṭabarī, 1987, vol. 32, p.58). Importantly, Sahl's actions were not part of the power struggle and succession, e.g., between Maʿmūn and Ibrāhīm ibn al-Mahdi (p.61-83), or even part of the Kharijites' rebellion in the same period (p.67-8).

We highlighted the prevalence of rebellions, gross violations of Islamic law as the justification for some rebellions, and the participation of the Prophet's companions and family members, all of whom were well-versed in Islamic law. What about the attitude of later Muslim scholars and distinguished jurists towards rebellion? Evidence suggests a persistent Islamic juristic tradition, involving the participation of jurists in rebellions, their advocacy for the rights of rebels, and juristic writings on the conditions for legitimate rebellion. Given

the tendency of legal scholarship toward order and the prevalence of rebellion in Islamic civilization,[21] it is surprising how frequently prominent jurists opposed rulers, participated in revolts, emphasized the legal protection of rebels, and even at times, stated that rebellion against unjust rulers is permissible or even obligatory.

Some group such as Kharijites and Mu'tazilites believed that "the community was obliged to remove a wrongful ruler" (Crone, 2004, p.229). Ibadis, a branch of Kharijites who constitute a majority in contemporary Oman, routinely engaged in revolts, sanctioned by their jurists, against rulers, including Ummayads and Abbassids, to establish an Ibadi Imamate. They believed that a wrongful imam should be removed by force if necessary and possible (Cook, 2001, Ch. 15). In fact, tendencies to view rebellion as a form of forbidding wrong (an important ordinance of Islamic law) "characterise the early Khārijites, the Ībāḍīs, the Zaydīs... [and are] embalmed in the Imāmī heritage" (Cook, 2001, p.477-8).[22] For example, "Zaydism laid claim to, and continued, an old ʿAlid pattern: rebellion against unjust rule with the aim of establishing a legitimate imamate. References to forbidding wrong are a recurring... feature of accounts of such ʿAlid risings" (Cook, 2001, p.231). Next, we focus on the proto-Sunni/Sunni branch of Islam.

There are four major Sunni schools of jurisprudence, Hanafi, Maliki, Shafi'i, and Hanbali, whose origins are associated with four prominent jurists, Abu Hanifa (d. 767), Malik ibn Anas (d. 795), Shafi'i (d. 820), and Ahmad ibn Hanbal (d. 855), respectively. The interaction between Abu Hanifa and a certain goldsmith is illuminating about the logic of his attitude toward revolt and those of many other jurists (Cook, 2001, Ch. 1). The goldsmith discusses with Abu Hanifa if forbidding wrong is obligatory. Abu Hanifa confirms that it is. The goldsmith then proposes that he gives his allegiance to Abu Hanifa to start a rebellion. Abu Hanifa refuses, arguing that the rebellion will fail, and the goldsmith will be killed without bringing any good to others. Abu Hanifa repeats the same logic in subsequent in-

[21] Jurists are lawyers and legal scholars, not political theorists. As Abou El Fadl (2001, p.26) argues, "Jurists... will be concerned with issues of order, conflict resolution, and stability. They may demand that this order be just, or that it would comply with the divine command, but they can hardly be expected to advocate lawlessness or anarchy." And yet, "the idea of a revolt as a means to power was neither alien nor abhorrent to Muslim jurists" (p.75).

[22] Zaydis, Isma'ilis, and Imamis were eventually formed as branches of Shi'i Islam. The majority of Shi'is today (e.g., in Iran, Iraq, and Lebanon) are Imamis. The Fatimid caliphate in Egypt (prior to Mamluks) were Isma'ilis. The majority of Shi'is in Yemen are Zaydis.

teractions, even highlighting that such actions come close to one becoming an accomplice in one's own death. All the while, the jurist transmits to the goldsmith the Prophetic tradition that "The lord of the martyrs is Hamza ibn Abd al-Muttalib and a man who stands up to an unjust ruler, commanding and forbidding, and is killed by him." The goldsmith is said to have also transmitted another Prophetic tradition from Abu Hanifa: "The finest form of holy war [jihad] is speaking out in the presence of an unjust ruler, and getting killed for it." The goldsmith got killed later while forbidding wrong. Variations of this last tradition appear in various canonical hadith collections. In fact, the topics of forbidding wrong and holy war are discussed together in various law-books (Cook, 2001, Ch. 1 and p.490). Abu Hanifa (d. 767) himself was imprisoned, tortured, and died in an Abbasid prison for his defiance toward the caliph (Abou El Fadl, 2001, p.73, p.76-8). While the exact reason for his imprisonment and death is disputed, according to Schacht (*Encyclopaedia of Islam, 2nd Ed.*), "The truth is probably that he compromised himself by unguarded remarks at the time of the rising of the ʿAlids al-Nafs al-Zakiyya and his brother"—see also Cook (2001, p.8-9). There are also reports of his support for Zayd's rebellion (Abou El Fadl, 2001, p.72-3).

Similar reports exist about the founders of the other schools. Asked about whether it was legal to join a rebellion, Mālik responded that it was, arguing that "the *bayʿa* given to the caliph al-Manṣūr was obtained under duress". Later, the "governor, had Mālik beaten and flogged for his views on duress and, possibly, for supporting the rebellion" (Abou El Fadl, 2001, p.76). Even the "quietist" Ibn Hanbal was imprisoned during the Abbasid inquisition for refusing to accept the Abbasid doctrine that the Quran is created, which would allow the caliph discretion to alter the law—see Section 3.1.

Shafiʿi was among the "Several jurists [who], even before Hārūn al-Rashīd came to power, had either sympathized with or pledged their allegiance to al-Daylamī" associated with some Alid rebels (Abou El Fadl, 2001, p.80-1). Shafiʿi later was arrested for suspicion of plotting against the caliph, or according to a different report, for criticizing the governor (p.83-4). Once released, he settled in Egypt and developed a (technical) legal discourse on rebellion. The juristic tradition that he initiated was such that, centuries later, "Ibn Taymiyya [d. 1328]... even accused him and those who preceded him or followed in his footsteps of suborning rebellion and spreading *fitna*" (p.98). For example, he defined a rebel

(*bāghī*) as "one who refuses to obey the just ruler (*al-imām al-ʿādil*), and intends to rebel by fighting him." Jurists discussed extensively whether Shafi'i meant a ruler who initially came to power lawfully, or a ruler who behaves justly. He "seems to have used the expression to mean the substantively just ruler. This implies that those who rebel against an unjust ruler are not rebels at all, and in fact later jurists explicitly argue that if the ruler is unjust and the rebels are just, then the ruler is to be considered the *bāghī* and not the rebels" (Abou El Fadl, 2001, p.148-9).

The juristic tradition that developed included the law of rebellion, which specified a significant degree of protection and leniency toward rebels, regardless of whether the ruler against whom they had rebelled was just or unjust. Abou El Fadl (2001, p.238) states:

> At a minimum, the majority of Sunnī jurists agreed that the rebels are not to be held liable for life and property destroyed during the course of their rebellion, and that, as a general matter, rebels may not be executed and their properties may not be confiscated.

The openly critical attitude of distinguished jurists was not limited to the founders of the schools. Abou El Fadl (2001, p.96-8, 166) and (Cook, 2001, p.148-9, 348, 355-6, 383-5) provide many examples from Thawrī (d. 788) to Nūr al-Dīn al-Bakrī (d. 1324). Summarizing the political aspect of the religious duties of *Commanding Right and Forbidding Wrong*, Cook (2001) observes that "the biographical and anecdotal record is full of sympathetically presented examples of pious Muslims harshly rebuking rulers, governors and their henchmen, often at great risk to themselves... This activity has the sanction of the Prophetic tradition... It is occasionally suggested that it is a duty to forbid wrong in this fashion, and in any case the activity is widely regarded with favour... someone who loses his life in the process is accordingly a martyr... Favourable attitudes to forbidding wrong through rebellion are less common, but they do exist" (p.476-7). In fact, many jurists were involved in rebellions. Ibn Khaldûn (2015, p.127) writes of "revolutionaries from among the common people *and of jurists* who undertake to reform evil practices. Many religious people who follow the ways of religion come to revolt against unjust amirs. They call for a change in, and prohibition of, evil practices" (our emphasis). Abou El Fadl (2001, p.70-1, 76, 100, 205)

provides other examples and Section 3.1 mentions examples of the jurists' involvements in revolts in the Ottoman empire—see also Lapidus (1975, p.381).

There was a range of juristic arguments regarding initiating, aiding, or joining rebellion. Some jurists explicitly argued that initiating rebellion to depose an unjust ruler was lawful, as long as the likelihood of success and the expected net social benefit of rebellion are sufficiently high. "Rebels should balance the chance of success and weigh it against the potential harm that will result from the rebellion. If the potential harm to society is grave, and the chances of success are limited, then rebellion is prohibited. However, if the chances of success are reasonably good, and the harm to society is limited, then rebellion is permitted" (Abou El Fadl, 2001, p.286). Some even argued that assisting just rebels against unjust rulers was obligatory. For example, Simnānī (d. 1105) "argues that if the ruler becomes oppressive and unjust, and usurps property, then it becomes incumbent upon the jurists and Muslims to overthrow him, and he even goes as far as claiming that this has been the consistent practice of Muslims in dealing with corrupt rulers" (Abou El Fadl, 2001, p.194). Cook (2001, p.337, 346, 385-6, 390) and Abou El Fadl (2001, p.286-7, fn.161) document various other examples. Moreover, some jurists who were against rebellion or aiding rebels, also insisted that Muslims must not aid an unjust ruler to repress rebels who have a just cause. Instead, they should stay entirely out of the conflict. Clearly, this subtle form of collective action, if followed, would effectively render the ruler defenseless, ensuring the rebels' victory.

We end by highlighting the attitudes of two influential and well-known jurists: Juwayni (d. 1085) and Ghazali (d. 1111). In his classic, *Irshād*, Juwayni writes (Cook, 2001, p.346):

> If the ruler of the time (*wālī al-waqt*) acts in a manifestly unjust fashion, and does not respond to verbal admonition, then it is for 'the people of binding and loosing' (*ahl al-ḥall waʾl-ʿaqd*) to prevent him, even if this means doing battle with him. [See Abou El Fadl (2001, p.182) for a discussion.]

In his influential *Iḥyā*, Ghazali allocated a chapter to commanding and forbidding rulers. He praises those who speak out harshly against rulers for their wrongdoings, providing 17 anecdotes of the earlier Muslims who stood up to rulers and died martyrs. Moreover, he argues in detail that when lower degrees of forbidding wrong fails, ordinary Mulsims, *without*

the permission of rulers, can band together, arm themselves, and use force and violence in forbidding wrong. In his *Iqtiṣād*, Ghazali argues that an imam who is not a mujtahid should be removed, if it could be done without a fighting (Crone, 2004, p.228, fn.53). Even in *Faḍāʾih*, written primarily to defend the Abbasids, he offers a conditional support for rebellion. Abou El Fadl (2001, p.184) summarizes:

> If the ruler's commands are illegal, or if he does not rely on jurists, and someone more qualified for the position is found, people should weigh the benefits and costs of attempting to overthrow him, and replace him with someone better if they are able to do so. Ultimately, despite all the polemics in favor of the ʿAbbāsid caliph, al-Ghazālī's argument... reduces itself to a balancing act between the pros and cons of attempting to overthrow the ruler.

The Christian recension of Ghazali's *Iḥyā* in Gregory Barhebraeus's (d. 1286) *Ethicos* is telling about the radically different attitude toward rebellion in the Christian West and Islamic world. *Ethicos* closely follows Ghazali's structure of arguments, but Christianizes various aspects. In particular, in *Ethicos*, the duty of admonition and rebuke (corresponding to forbidding wrong) are reserved for ecclesiastical authority, because orders can be given only from superiors to inferiors. Moreover, even churchmen must not use force, which are reserved for the secular authority. This contrasts with Ghazali's views, even in his conservative *Iḥyā*—see Cook (2001, p.600-3). Differences are more striking once we recognize that, among Muslim thinkers, Ghazali's "political thought is dominated rather by a fear of civil war (*fitna*) and disturbances (*fasād*) leading to disorder and anarchy" (Lambton, 1981, p.109).

In sum, there is a common thread among Muslims jurists from Abu Hanifa in the 8th century to Juwayni and Ghazali in the 11th and 12th centuries to ʿUbbi (d. 1424) in the 14th and 15th centuries: Rebellion is costly and its success is uncertain, and hence should only be attempted if its social benefits are sufficiently higher than its social costs and the chances of success are sufficiently high. These features are built into our model. Importantly, jurists, even Ottoman jurists of the 17th and early 18th century (see Section 3.1), sometimes found revolt lawful, even obligatory, and prohibited aiding an unjust ruler against rebels.

4.3 Solidarity and Homogeneity

Islamic tradition stressed each Muslim's responsibility to "enjoin what is good and forbid what is wrong" (Quran 31:17). As Cook (2014, p.20-3) argues, solidarity and equality before the law were integral parts of the ideal Islamic identity. Muslims are "like a body, parts of a whole" (p.22), with no caste or aristocracy: "Indeed, Muslims are brothers" (Quran 49:10); "remember the favor of Allah upon you, when you were enemies and He brought your hearts together and you became, by His favor, brothers" (Quran 3:103). These were reflected in politics. Compared to other dynasties in Eurasia, "the early Muslim state was exceptional in that it refused to adopt the title 'king' " (Anjum, 2012, p.47, fn.39). Anjum (2012, p.51) argues: "The ethic of the Qur'an is on the whole egalitarian and activist... the unyielding monotheism of the Qur'an, coupled with its insistence on rational piety that required obeying none but God and his Prophet, encouraged questioning authority and using one's own reason instead of following tradition or other men's judgment".

These normative ideas of homogeneity (in the desire to follow God's law) and solidarity were also included in the concept of Islamic *umma*, as the universal community of believers, worshipping one God and following His law. "[B]reaking with the community... was not merely unfortunate and undesirable, but positively evil, because there was only one Islamic *umma*. To break ties with the *umma* was to break both with God and man" (Donner, 1981, p.56). The Quran urges believers to "hold firmly to the rope of Allah all together and do not become divided... And do not be like the ones who became divided and differed after the clear proofs had come to them" (Quran 3:103-10). The implication was that the members of community must struggle to improve each others spiritual and material conditions.

Taking actions to uphold the law and improve the welfare of the Muslim community (the two intertwined concepts) were religious obligations (under some conditions) and were highly commendable in the Islamic normative tradition, associated with great rewards in the afterlife. Such notions would facilitate participation in costly collective action. Section 4.2 includes several examples of how these ideas are invoked in rebellions against unjust rulers.

5 Discussion and Conclusion

In this paper we have identified a new puzzle; that Islamic scholars never developed, until the reforms of the 19th century, institutional models of executive constraints. This was so even though they were aware of the dangers of tyranny. In this they diverged from the Western tradition emanating from Sparta, Aristotle, and Plato. We argued that this was because of the cultural and social contexts in which Islamic normative tradition was embedded. Unlike in the Greco-Roman-Christian world, where legislation was mostly in secular hands, in Islam the law was determined in detail by God. We argued that this made it much easier for Muslims to determine when rulers were deviating from set policies and thus they were better able to use collective action to discipline rulers without the need for institutional safeguards. In our theory this mechanism is fortified by the homogeneity of Islamic societies, since everyone was a believer, and in the basic norms of solidarity in Islam, including the stipulation that everyone should "command right and forbid wrong" (Cook, 2001).

We do not claim that our explanation is the sole reason for the historical absence of the discussion of institutional constraints in the Islamic normative tradition. Because the puzzle is new, it is unclear what alternative explanation to consider. One possibility is to attribute this absence to Muslim thinkers' limited access to the Greco-Roman philosophical writings or history. For example, while Plato's *Republic* and *Laws* and Aristotle's *Ethics* were familiar to Muslim philosophers, it seems that they did not have access to a translation of Cicero's *De re publica*, or Aristotle's *Politics* where theories of mixed constitutions were more explicitly advocated (Melamed, 2011). This view implies that, without the help of the Greeks' discoveries, many generations of Muslim thinkers could not take what Crone calls "a short step" toward even a theoretical discussion of institutional constraints on rulers. Their "political horizon... did not reach to suggesting reforms or offering alternative institutions," as Halbertal and Holmes (2017, p.166) describe some of their earlier Jewish counterparts in antiquity. This view seems implausible. The vast territory of the Islamic Empire included people of various geographical and religious backgrounds, some of whom interacted routinely with Muslim scholars and many of whom played key roles in translating the vast corpus of Greek knowledge into Arabic (Gutas, 1998). That generations of Muslim scholars over huge

geographical and time periods did not have any knowledge of the political structure of Greek city-states, the Roman Republic, or even the Roman Empire with its Senate seems unlikely. As Gutas (1998, p.23) argues, "the historian Ḥamza al-Iṣfahānī (d. after 350/961) relates that when 'he needed information on Graeco-Roman history, he asked an old Greek, who had been captured and served as a valet, to translate for him a Greek historical work orally. . . oral translation by native speakers of whatever language within the Islamic domain did occur and that. . . it must have been widely practiced". To make sense of the puzzle, one must go beyond explanations that Muslim and Jewish thinkers did not discuss institutional constraints on rulers even in theory because they did not learn their usefulness from Aristotle or Cicero.

Another possibility is that "[p]olitics in Islam had remained the domain of the barbarian", while the Islamic ideal remained a "tribal state" inspired by pre-Islamic free Arab tribes, lacking "a form of settled government" (Crone, 1980, p.91). One could deduce such interpretations from Crone's earlier works. However, Crone's subsequent works corrected this view. "Neither the pre-Islamic Arabs nor the tribesmen who continued to inhabit the peninsula after the conquests were regarded as a model of inspiration or imitation for Muslims as far as political organization was concerned" (Crone, 2004, p.268-9). In fact, the domination of Arab tribes by the Muslim elite (e.g., Ridda Wars) reflects a relatively centralized state with an energetic executive (Donner, 1981). Moreover, the importance of imams/caliphs and the executive duties to implement the law were far from notions of tribal confederacies. Muslims thinkers routinely investigated all practical matters from contracts to taxes. They theorized the caliphate even when it had departed from its ideal—see Section 3.1. It would be odd to ignore taming contemporary governments due to historical ideals of a tribal state. Finally, many Muslim thinkers were of non-Arabic descent, living in cosmopolitan city environments or in a Persianate milieu.[23] A more plausible explanation should account for why Muslim thinkers in different social environment over vast geographical regions in an extended period of time did not engage in discussing institutional constraints on rulers.

[23]Alternatively, one may argue that notions of absolutism from Persian kingship prevented theorizing about institutional constraints. However, such notions contradicted basic tenets of Islam, where only God's authority is absolute and government is constrained by God's law. Moreover, even the "divine charisma" bestowed upon some rulers in Persian traditions is not permanent and will depart if the king behaves badly. Ferdowsi's *Book of Kings* (*Shāhnāmih*) includes good rulers who turn bad (e.g., *Jamshīd* who lost his *farrah*), bad rulers who turn worse (e.g., *Zahhāk*), and popular revolts against bad rulers (e.g., the story of *Kāvih*).

Appendix

A Proofs

Proposition 6 is obtained as a corollary of the following Proposition.

Proposition 8. *Suppose $\gamma \sim U[0,1]$. Let $Q = Pr_\gamma(\mu \leq \mu^*(\gamma))$ be the probability that institutional constraints improve the majority citizen's policy payoff. Then,*

$$Q(\mu; M, p) = \begin{cases} 1 & ; \mu \leq (1-p)(q-q^2) \\ \min\left\{1, \frac{1-\delta_0}{1-T/M}\right\} & ; (1-p)(q-q^2) \leq \mu \leq (1-p(1-\delta_0))(q-q^2) \\ \min\left\{1, \frac{1}{(1-T/M)p}\left(1-\frac{\mu}{q-q^2}\right)\right\} & ; (1-p(1-\delta_0))(q-q^2) \leq \mu \leq (q-q^2) \\ 0 & ; (q-q^2) < \mu. \end{cases}$$

Proof of Proposition 8. Using Proposition 5,

$$\begin{aligned} Q &= Pr_\gamma(\mu \leq \mu^*(\gamma)) \\ &= Pr_\gamma\left(\mu \leq (1-p)(q-q^2), \beta > 1-\delta_0\right) + Pr_\gamma\left(\mu \leq (1-p\beta)(q-q^2), \beta < 1-\delta_0\right) \end{aligned}$$

Using the fact that $\beta = \beta(1, M, \gamma)$, and substituting Proposition 1, we have: $\beta = H\left((1-\frac{T}{M})\gamma\right)$. Because $H = U[0,1]$, $\beta = (1-\frac{T}{M})\gamma$. Substituting, we have:

$$\begin{aligned} Q &= Pr_\gamma\left(\mu \leq (1-p)(q-q^2), (1-\frac{T}{M})\gamma > 1-\delta_0\right) \\ &\quad + Pr_\gamma\left(\mu \leq (1-p(1-\frac{T}{M})\gamma)(q-q^2), (1-\frac{T}{M})\gamma < 1-\delta_0\right) \\ &= Pr_\gamma\left(\mu \leq (1-p)(q-q^2), \gamma > \frac{1-\delta_0}{1-\frac{T}{M}}\right) + Pr_\gamma\left(\gamma \leq \frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right), \gamma < \frac{1-\delta_0}{1-\frac{T}{M}}\right) \end{aligned}$$

Now, consider four different cases.

1. Suppose $\mu \leq (1-p)(q-q^2)$. In this case,

$$Q = Pr_\gamma\left(\gamma > \frac{1-\delta_0}{1-\frac{T}{M}}\right) + Pr_\gamma\left(\gamma \leq \frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right), \gamma < \frac{1-\delta_0}{1-\frac{T}{M}}\right) \quad (3)$$

Rearranging $\mu < (1-p)(q-q^2)$ yields: $1 < \frac{1}{p}\left(1-\frac{\mu}{q(1-q)}\right)$. Therefore, $\frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right) > \frac{1}{1-\frac{T}{M}} > \frac{1-\delta_0}{1-\frac{T}{M}}$, and Equation (3) further simplifies to:

$$Q = Pr_\gamma\left(\gamma > \frac{1-\delta_0}{1-\frac{T}{M}}\right) + Pr_\gamma\left(\gamma < \frac{1-\delta_0}{1-\frac{T}{M}}\right) = 1$$

2. Suppose $(1-p)(q-q^2) \leq \mu \leq (1-p(1-\delta_0))(q-q^2)$. In this case,

$$Q = Pr_\gamma\left(\gamma \leq \frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right), \gamma < \frac{1-\delta_0}{1-\frac{T}{M}}\right) \quad (4)$$

Since $\mu < (1-p(1-\delta_0))(q-q^2)$, $\frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right) > \frac{1-\delta_0}{1-\frac{T}{M}}$, and Equation (4) further simplifies to:

$$Q = Pr_\gamma\left(\gamma \leq \frac{1-\delta_0}{1-\frac{T}{M}}\right) = \min\left\{1, \frac{1-\delta_0}{1-\frac{T}{M}}\right\}$$

3. Suppose $(1-p(1-\delta_0))(q-q^2) \leq \mu \leq q-q^2$. Because $\mu > (1-p(1-\delta_0))(q-q^2) > (1-p)(q-q^2)$, in this case,

$$Q = Pr_\gamma\left(\gamma \leq \frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right), \gamma < \frac{1-\delta_0}{1-\frac{T}{M}}\right) \quad (5)$$

Since $\mu > (1-p(1-\delta_0))(q-q^2)$, $\frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right) < \frac{1-\delta_0}{1-\frac{T}{M}}$, and Equation (4) further simplifies to:

$$Q = Pr_\gamma\left(\gamma \leq \frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right)\right) = \min\left\{1, \frac{1}{p\left(1-\frac{T}{M}\right)}\left(1-\frac{\mu}{q(1-q)}\right)\right\}$$

4. Finally, suppose $\mu > q - q^2$. Because $\mu > (q - q^2) > (1 - p)(q - q^2)$, in this case,

$$Q = Pr_\gamma \left(\gamma \leq \frac{1}{p\left(1 - \frac{T}{M}\right)} \left(1 - \frac{\mu}{q(1-q)} \right), \gamma < \frac{1 - \delta_0}{1 - \frac{T}{M}} \right) \tag{6}$$

Since $\mu > q - q^2$, $1 - \frac{\mu}{q(1-q)} < 0$. Then, $Q = 0$.

□

The first part of Proposition 6 is then obtained as a special case of Proposition 8 for the case $\delta_0 < T/M$, and with defining $\mu' = \frac{\mu}{q-q^2}$. The second part of Proposition 6 follows because, as $Q(\mu')$ is decreasing in μ', with $\mu'_1 < \mu'_2$:

$$|Q(\mu'_2) - Q(\mu'_1)| = Q(\mu'_1) - Q(\mu'_2)$$

Moreover, since $\mu'_1 \in (1 - p(1 - T/M), 1)$, $Q(\mu'_1) = \frac{1-\mu'_1}{(1-T/M)p}$. Also, for any $\mu'_2 > \mu'_1 > 1 - p(1 - T/M)$, $Q(\mu'_2) = \max\{\frac{1-\mu'_2}{(1-T/M)p}, 0\}$. Therefore,

$$\begin{aligned} Q(\mu'_1) - Q(\mu'_2) &= \frac{1 - \mu'_1}{(1 - T/M)p} - \max\{\frac{1 - \mu'_2}{(1 - T/M)p}, 0\} \\ &= \frac{\min\{\mu'_2, 1\} - \mu'_1}{(1 - T/M)p} \end{aligned}$$

which is strictly decreasing in p and M. □

References

Abou El Fadl, Khaled. 2001. *Rebellion and Violence in Islamic Law*, Cambridge: Cambridge University Press.

Abou-El-Haj, Rifa'at. 1984. *The 1703 Rebellion and the Structure of Ottoman Politics*, Netherlands: Nederlands Historisch-Archaeologisch Instituut te Istanbul.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2005. "The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth," *American Economic Review*, 95 (3), 546–579.

Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*, Cambridge: Cambridge University Press.

Acemoglu, Daron, James A. Robinson, and Ragnar Torvik. 2013. "Why Do Voters Dismantle Checks and Balances?" *Review of Economic Studies*, 80 (3), 845–875.

Afary, Janet. 1996. *The Iranian Constitutional Revolution, 1906-1911: Grassroot Democracy, Social Democracy, the Origins of Feminism*, New York: Columbia University Press.

Aghion, Philippe, Alberto Alesina, and Francesco Trebbi. 2004. "Endogenous Political Institutions," *The Quarterly Journal of Economics*, 119 (2), 565–611.

Aksan, Virginia. 2022. *The Ottomans 1700-1923: An Empire Besieged, 2nd Edition*, New York: Routledge.

Amanat, Abbas. 2009. *Apocalyptic Islam and Iranian Shi'ism*, New York: Tauris.

Anjum, Ovamir. 2012. *Politics, Law, and Community in Islamic Thought: The Taymiyyan Moment*, New York: Cambridge University Press.

Ansari, Hassan and Mehdi Shadmehr. 2021. "Theological Origins of the Iranian Islamic Government," Mimeo.

Arjomand, Said Amir. 1984. *The Shadow of God and the Hidden Imam: Religion, Political Order, and Societal Change in Shi'ite Iran from the Beginning to 1890*, Chicago: University of Chicago Press.

Arjomand, Said Amir. 1988. *The Turban for the Crown: The Islamic Revolution in Iran*, New York: Oxford University Press.

Ashcraft, Richard. 1986. *Revolutionary Politics and Locke's Two Treatises of Government*, Princeton: Princeton University Press.

Auriol, Emmanuelle and Jean-Philippe Platteau. 2017. "Religious Co-option in Autocracy: A Theory Inspired by History," *Journal of Development Economics*, 127, 395–412.

Auriol, Emmanuelle, Jean-Philippe Platteau, and Thierry Verdier. 2023. "The Quran and the Sword," *Journal of the European Economic Association*, 21 (5), 1772–1820.

Averroes and Lahcen Elyazghi Ezzaher. 2015. *Three Arabic Treatises on Aristotle's Rhetoric: The Commentaries of Al-Fārābi, Avicenna, and Averroes. Translated, with Introduction and Notes, by Lahcen Elyazghi Ezzaher*, Carbondale: Southern Illinois University Press.

Averroes and Ralph Lerner. 1974. *Averroes on Plato's "Republic". Translated, with an Introduction and Notes, by Ralph Lerner*, Ithaca: Cornell University Press.

Averroes and Erwin I.J. Rosenthal. 1966. *Averroes' Commentary on Plato's Republic. Edited with an Introduction, Translation and Notes by E.I.J. Rosenthal*, Cambridge: Cambridge University Press.

Babayan, Kathryn. 2002. *Mystics, Monarchs, and Messiahs: Cultural Landscapes of Early Modern Iran*, Cambridge: Harvard University Press.

Bayat, Mangol. 1991. *Iran's First Revolution : Shi'ism and the Constitutional Revolution of 1905-1909*, New York: Oxford University Press.

Bernand, M. "Idjmāʿ," In Encyclopaedia of Islam, Second Edition, edited by P. Bearman, Th. Bianquis, C.E. Bosworth, E. van Donzel, W.P. Heinrichs, P.J. Bearman (Volumes X, XI, XII), Th. Bianquis (Volumes X, XI, XII), et al. doi:http://dx.doi.org/10.1163/1573-3912_islam_COM_0350, Accessed February 16, 2022.

Besley, Timothy and Torsten Persson. 2011. "The Logic of Political Violence*," *The Quarterly Journal of Economics*, 126 (3), 1411–1445.

Bickerman, Elias J. 1962. *From Ezra to the Last of the Maccabees: Foundations of Post-Biblical Judaism*, New York: Schocken.

Bisin, Alberto, Jared Rubin, Avner Seror, and Thierry Verdier. 2023. "Culture, Institutions and the Long Divergence," *Journal of Economic Growth*, 10.1007/s10887-023-09227-7, forthcoming.

Black, Antony. 2011. *History of Islamic Political Thought: From the Prophet to the Present*, Edinburgh: Edinburgh 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 (1), 16–34.

Boleslavsky, Raphael, Mehdi Shadmehr, and Konstantin Sonin. 2021. "Media Freedom in the Shadow of a Coup," *Journal of the European Economic Association*, 19 (3), 1782–1815.

Butterworth, Charles E. 1998. "Averroës' Platonization of Aristotle's *Art of Rhetoric*," in *La Rhétorique d'Aristote: traditions et commentaires de l'Antiquité au XVII^e siècle* eds. by Dahan, Gilbert and Irène Rosier-Catach, 227–40, Paris: J. Vrin.

Chaney, Eric. 2013. "Revolt on the Nile: Economic Shocks, Religion, and Political Power," *Econometrica*, 81 (5), 2033–2053.

Chaney, Eric. 2022. "Islam and Political Structure in Historical Perspective," in *The Oxford Handbook of Politics in Muslim Societies* eds. by Cammett, Melani and Pauline Jones, New York: Oxford University Press.

Cook, Michael. 2001. *Commanding Right and Forbidding Wrong in Islamic Thought*, New York: Cambridge University Press.

Cook, Michael. 2014. *Ancient Religions, Modern Politics: The Islamic Case in Comparative Perspective*, Princeton: Princeton University Press.

Crone, Patricia. 1980. *Slaves on Horses: The Evolution of the Islamic Polity*: Cambridge University Press.

Crone, Patricia. 2000. "Ninth-Century Muslim Anarchists," *Past & Present* (167), 3–28.

Crone, Patricia. 2001. "*Shūrā* as an Elective Institution," *Quaderni di Studi Arabi*, 19, 3–39.

Crone, Patricia. 2004. *God's Rule - Government and Islam: Six Centuries of Medieval Islamic Political Thought*, New York: Columbia University Press.

Crone, Patricia. 2012. *The Nativist Prophets of Early Islamic Iran: Rural Revolt and Local Zoroastrianism*, Cambridge: Cambridge University Press.

Crone, Patricia and Martin Hinds. 1986. *God's Caliph: Religious Authority in the First Centuries of Islam*, Cambridge: Cambridge University Press.

Dabashi, Hamid. 1988. "Two Clerical Tracts on Constitutionalism," in *Authority and Political Culture in Shi'ism* ed. by Arjomand, Said Amir, Albany: State University of New York Press.

Dakake, Maria. 2007. *The Charismatic Community: Shi'ite Identity in Early Islam*, New York: State University of New York Press.

Donner, Fred. 1981. *The Early Islamic Conquests*, Princeton: Princeton University Press.

Donner, Fred M. 2011. "Qur'ânicization of Religio-Political Discourse in the Umayyad Period," *Revue des mondes musulmans et de la Méditerranée* (129), 79–92.

Farabi and Charles E. Butterworth. 2015. *Alfarabi: The Political Writings, Volume II. Translated, Annotated, and with Introduction by Charles E. Butterworth*, Ithaca: Cornell University Press.

Finer, Samuel Edward. 1999. *The History of Government from the Earliest Times - Volume II: The Intermediate Ages*, Oxford: Oxford University Press.

Gibb, Hamilton A. R. 1982. *Studies on the Civilization of Islam*, Princeton: Princeton University Press.

Goodman, Lenn Evan. 1992. *Avicenna*, New York: Routledge.

Gutas, Dimitri. 1998. *Greek Thought, Arabic Culture: The Graeco-Arabic Translation Movement in Baghdad and Early 'Abbasaid Society (2nd-4th/5th-10th c.)*, Abingdon: Taylor & Francis Group.

Gutas, Dimitri. 2004. "The Meaning of Madani in al-Farabi's "Political" Philosophy," *Mélanges de l'Université Saint-Joseph*, 57, 259–282.

Halbertal, Moshe and Stephen Holmes. 2017. *The Beginning of Politics: Power in the Biblical Book of Samuel*, Princeton: Princeton University Press.

Hallaq, Wael. 1984. "Caliphs, Jurists and the Saljūqs in the Political Thought of Juwaynī," *The Muslim World (Hartford)*, 74 (1), 26–41.

Hallaq, Wael. 2009. *Sharī'a: Theory, Practice, Transformations*, New York: Cambridge University Press.

Hallaq, Wael. 2014. *The Impossible State: Islam, Politics, and Modernity's Moral Predicament*, New York: Columbia University Press.

Halpern, Baruch. 1981. *The Constitution of the Monarchy in Israel*, The Netherlands: Brill.

Hamilton, Alexander, James Madison, and John Jay. 2008. *The Federalist Papers*, Oxford World's Classics, New York: Oxford University Press.

Hillenbrand, Carole. 1988. "Islamic Orthodoxy or Realpolitik? Al-Ghazālī's Views on Government," *Iran*, 26 (1), 81–94.

Huntington, Samuel. 1996. *The Clash of Civilizations and the Remaking of World Order*, New York: Simon & Schuster.

Ibn Khaldûn. 2015. *The Muqaddimah: An Introduction to History - Abridged Edition*, Translated and introduced by Franz Rosenthal, Abridged and Edited by N. J. Dawood, with an Introduction by Bruce B. Lawrence, Princeton: Princeton University Press.

Jafarian, Rasool. 2003/1382. *Siyyid Asadullāh Kharaqānī: Ruhānī Nugarāy Mashrūṭih va Riza Shāh*, Tehran: Markaz Asnad Enqelab Eslami.

Johansen, Baber. 1988. *The Islamic Law on Land Tax and Rent: The Peasants' Loss of Property Rights as Interpreted in the Hanafite Legal Literature of the Mamluk and Ottoman Periods*, London: Croom Helm.

Khomeini, Roohollah. 2006/1385. *Vilāyat Faqīh*, Tehran: Moasseseh Tanzim va Nashr Asar Emam Khomeini.

Khoury, Dina Rizk. 1997. *State and Provincial Society in the Ottoman Empire: Mosul, 1540-1834*, Cambridge: Cambridge University Press.

Klosko, George. 2012. *History of Political Theory: An Introduction. Volume I: Ancient and Medieval. Second Edition*, Oxford: Oxford University Press.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Kuran, Timur. 2012. *The Long Divergence: How Islamic Law Held Back the Middle East*, Princeton: Princeton University Press.

Kuran, Timur. 2016. "Legal Roots of Authoritarian Rule in the Middle East: Civic Legacies of the Islamic Waqf," *The American Journal of Comparative Law*, 64 (2), 419–454.

Kuran, Timur. 2023. *Freedoms Delayed: Political Legacies of Islamic Law in the Middle East*: Cambridge University Press.

Lambton, Ann K.S. 1981. *State and Government in Medieval Islam – An Introduction to the Study of Islamic Political Theory: the Jurists*, Oxford: Oxford University Press.

Lapidus, Ira M. 1975. "The Separation of State and Religion in the Development of Early Islamic Society," *Int. J. Middle East Stud.*, 6 (4), 363–385.

Leaman, Oliver. 1988. *Averroes and His Philosophy*, London: Routledge.

Lecker, Michael. 2004. *The "Constitution of Medina": Muḥammad's First Legal Document*, Princeton: Darwin Press.

Lerner, Ralph and Muhsin Mahdi. 1963. *Medieval Political Philosophy: A Sourcebook*, Ithaca: Free Press of Glencoe.

Lewis, Bernard. 1982. *The Muslim Discovery of Europe*, New York: Norton & Company.

Lizzeri, Alessandro and Nicola Persico. 2004. "Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform'," *The Quarterly Journal of Economics*, 119 (2), 707–765.

Macpherson, C.B. 1962. *The Political Theory of Possessive Individualism*, New York: Oxford University Press.

Madelung, Wilferd. 1997. *The Succession to Muhammad: A Study of the Early Caliphate*, Cambridge: Cambridge University Press.

Mardin, Serif. 1962. *Genesis of Young Ottoman Thought*, Princeton: Princeton University Press.

Melamed, Abraham. 2011. "Aristotle's *Politics* in Medieval and Renaissance Jewish Political Thought," in *Well Begun is Only Half Done: Tracing Aristotle's Political Ideas in Medieval Arabic, Syriac, Byzantine, and Jewish Sources* ed. by Syross, Vasileios, 145–186, Tempe: Arizona Center for Medieval and Renaissance Studies.

Modarressi, Hossein. 1983. *Kharāj in Islamic Law*, London: Anchor Press.

Modarressi, Hossein. 1993. *Crisis and Consolidation in the Formative Period of Shi'ite Islam*, Princeton: Princeton University Press.

Mokyr, Joel. 1990. *The Lever of Riches: Technological Creativity and Economic Progress*, New York: Oxford University Press.

Morris, Stephen and Mehdi Shadmehr. 2023. "Inspiring Regime Change," *Journal of the European Economic Association*, 21 (6), 2635–81.

Morris, Stephen and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review*, 88 (3), 587–597.

Morris, Stephen and Hyun Song Shin. 2003. "Global Games: Theory and Applications," in *Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress, Vol. I* eds. by Dewatripont, Mathias, Lars Peter Hansen, and Stephen J. Turnovsky, Cambridge: Cambridge University Press.

North, Douglass C. and Robert P. Thomas. 1973. *The Rise of the Western World: A New Economic History*, New York: Cambridge University Press.

North, Douglass C. and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England," *Journal of Economic History*, 49 (4), 803–832.

Ober, Josiah. 2022. *The Greeks and the Rational: The Discovery of Practical Reason*, Berkeley: University of California Press.

Persson, Torsten, Gérard Roland, and Guido Tabellini. 1997. "Separation of Powers and Political Accountability," *The Quarterly Journal of Economics*, 112 (4), 1163–1202.

Persson, Torsten and Guido Tabellini. 2004. "Constitutional Rules and Fiscal Policy Outcomes," *American Economic Review*, 94 (1), 25–45.

Platteau, Jean-Philippe. 2017. *Islam Instrumentalized: Religion and Politics in Historical Perspective*, New York: Cambridge University Press.

Rheault, Magali and Dalia Mogahed. 2008. "Many Turks, Iranians, Egyptians link Sharia and Justice," https://news.gallup.com/poll/109072/many-turks-iranians-egyptians-link-sharia-justice.aspx.

Rosenthal, Erwin Isak Jakob. 1958. *Political Thought in Medieval Islam: An Introductory Outline*, New York: Cambridge University Press.

Rosenthal, Erwin Isak Jakob. 1971. *Studia Semitica, Volume II: Islamic Themes*, Cambridge: Cambridge University Press.

Roy, Olivier. 1994. *The Failure of Political Islam*, Cambridge: Harvard University Press.

Rubin, Jared. 2017. *Rulers, Religion, and Riches: Why the West Got Rich and the Middle East Did Not*, New York: Cambridge University Press.

Saleh, Mohamed and Jean Tirole. 2021. "Taxing Identity: Theory and Evidence from Early Islam," *Econometrica*, 89 (4), 1881–1919.

Sariyannis, Marinos. 2019. *A History of Ottoman Political Thought up to the Early Nineteenth Century, with a chapter by E. Ekin Tuşalp Atiyas*, Leiden: Brill.

Schacht, J. "Abū Ḥanīfa Al-Nu'mān," In Encyclopaedia of Islam, Second Edition, edited by P. Bearman, Th. Bianquis, C.E. Bosworth, E. van Donzel, W.P. Heinrichs, P.J. Bearman (Volumes X, XI, XII), Th. Bianquis (Volumes X, XI, XII). doi:http://dx.doi.org/10.1163/1573-3912_islam_SIM_0194, Accessed November 23, 2022.

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

Stern, Menahem. 1968. "The Hasmonean Revolt and its Place in the History of Jewish Society and Religion," *Cahiers d'Histoire Mondiale. Journal of World History. Cuadernos de Historia Mundial*, 11 (1), 92–106.

Stern, Samuel Miklos. 1970. "The Constitution of the Islamic City," in *Papers on Islamic History, I: The Islamic City, A Colloquium* eds. by Hourani, Albert Habib and Samuel Miklos Stern, Oxford: Bruno Cassirer; [Philadelphia] University of Pennsylvania Press.

Ṭabarī, Abū Jaʿfar Muḥammad b. Jarīr. 1987. *The History of al-Ṭabarī, Vol. XXXII: The Reunification of The ʿAbbāsid Caliphate. Translated and Annotated by C. E. Bosworth*, Albany: State University of New York Press.

Ṭabarī, Abū Jaʿfar Muḥammad b. Jarīr. 1989. *The History of al-Ṭabarī, Vol. XXVI: The Waning of the Umayyad Caliphate. Translated and Annotated by Carole Hillenbrand*, Albany: State University of New York Press.

Ṭabarī, Abū Jaʿfar Muḥammad b. Jarīr. 1990. *The History of al-Ṭabarī, Vol. XIX: The Caliphate of Yazīd b. Muʿāwiyah. Translated and Annotated by I. K. A. Howard*, Albany: State University of New York Press.

Tabatabai, Javad. 2006/1385. *Darāmadī bar Tārīkh Andīshih Dar Iran, Vīrāstih Jadīd*, Tehran: Kavir.

Tezcan, Baki. 2010. *The Second Ottoman Empire: Political and Social Transformation in the Early Modern World*, New York: Cambridge University Press.

Tierney, Brian. 2010. *The Crisis of Church and State: 1050–1300*, Toronto: University of Toronto Press.

van Ess, Josef, "al-Aṣamm," In Encyclopaedia of Islam, Second Edition, edited by P. Bearman, Th. Bianquis, C.E. Bosworth, E. van Donzel, W.P. Heinrichs, P.J. Bearman (Volumes X, XI, XII), Th. Bianquis (Volumes X, XI, XII), et al. doi:http://dx.doi.org/10.1163/1573-3912_islam_SIM_8354, Accessed January 24, 2023.

Wasserstrom, Steven M. 1995. *Between Muslim and Jew: The Problem of Symbiosis under Early Islam*, Princeton: Princeton University Press.

Watt, William Montgomery. 2003. *Islamic Political Thought*, Edinburgh: Edinburgh University Press.

Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law," *American Political Science Review*, 91 (2), 245–263.

Yavari, Neguin. 2014. *Advice for the Sultan: Prophetic Voices and Secular Politics in Medieval Islam*, New York: Oxford University Press.

Ziad, Waleed. 2021. *Hidden Caliphate: Sufi Saints beyond the Oxus and Indus*, Cambridge: Harvard University Press.

Missing Discussions: Institutional Constraints in the Islamic Political Tradition

Online Appendix

A. Arda Gitmez* James A. Robinson† Mehdi Shadmehr‡

A Welfare Inclusive of Revolt Costs

In this section, we argue that all the main results in the main text go through if the costs of revolt are included in the payoff of majority citizens, in addition to the policy payoff.

In the equilibrium of the coordination game, a citizen i revolts if and only if his direct cost of revolt c_i is below the equilibrium threshold c^* and the revolt succeeds if and only if the $\bar{c} < \bar{c}^*$. Moreover, in the limit as $\rho \to 0$: $\lim_{\rho\to 0} c^*(\rho) = \lim_{\rho\to 0} \bar{c}^*(\rho)$ (Boleslavsky, Shadmehr and Sonin, 2021). If $\bar{c} < \bar{c}^*$, then almost all citizens (members of the Majority) revolt; if $\bar{c} > \bar{c}^*$, then almost no citizen revolts. Thus, when revolt is attempted, the expected cost of revolt is:

$$\Pr\left(\bar{c} < \bar{c}^*\right) \cdot \mathbb{E}[\bar{c} \mid \bar{c} < \bar{c}^*]$$

Given $\bar{c} \sim U[0,1]$, this is equal to:

$$\Pr\left(\bar{c} < \bar{c}^*\right) \cdot \mathbb{E}[\bar{c} \mid \bar{c} < \bar{c}^*] = \frac{(\bar{c}^*)^2}{2}$$

Under the cost threshold c^*, there is a revolt with probability $\beta \in [0,1]$. Note that $\beta = \bar{c}^*$ whenever $\bar{c}^* > 0$ and $\beta = 0$ whenever $\bar{c}^* \leq 0$. Thus, $\Pr\left(\bar{c} < \bar{c}^*\right) \cdot \mathbb{E}[\bar{c} | \bar{c} < \bar{c}^*] = \beta^2/2$. This, in turn, implies that to account for the expected costs of revolt in the citizens' expected payoffs, we can simply subtract $\beta^2/2$ whenever a revolt is attempted by a strictly positive measure of citizens. Because when the revolt succeeds, the citizen payoff increases by 2, this means that to account for the expected costs of revolt in the citizen's payoffs, we can simply substitute β with $\beta_c(\beta) = \beta - \beta^2/4$ when calculating the value of the expected payoffs: $Pr(\textit{revolt attempted}) \cdot (2\beta - \beta^2/2) = 2Pr(\textit{revolt attempted}) \cdot (\beta - \beta^2/4)$. Because $\beta_c(0) = 0$,

*Department of Economics, Bilkent University. E-mail: arda.gitmez@bilkent.edu.tr

†University of Chicago, Harris School of Public Policy and Department of Political Science, 1307 East 60th street, Chicago, IL 60637. E-mail: jamesrobinson@uchicago.edu

‡Department of Public Policy, University of North Carolina at Chapel Hill. E-mail: mshadmehr@unc.edu

$\beta_c(1) = 1/2$, and $\beta_c(\beta)$ is strictly increasing in β, all our main results go through if we add the direct costs of revolt into the citizens' payoff, and then compare them under different institutional arrangements. For completeness, we derive these results below.

Recall that, in the equilibrium without institutional constraints, there is a revolt only if $\hat{s} = 0$ and the ruler takes action 1. In particular,

- When $\beta(1, M, \gamma) > 1 - \delta_0$, the threat of revolt disciplines the minority-congruent ruler, and the minority-congruent ruler takes action $a = 0$ when $\hat{s} = 0$. Consequently, there are no revolts, and there are no costs of revolt. In this case, payoffs inclusive of revolt costs are equal to policy payoffs. The majority citizens' expected payoff inclusive of revolt costs is:

$$1 - q(1 - p)$$

- When $\beta(1, M, \gamma) < 1 - \delta_0$, the minority-congruent ruler takes action $a = 1$ when $\hat{s} = 0$. When that happens, a revolt is attempted by a strictly positive measure of citizens. Therefore, a revolt is attempted by a strictly positive measure of citizens with probability

$$\Pr(t = b) \cdot \Pr(s = 0) \cdot \Pr(\hat{s} = 0 \mid s = 0) = q \cdot \frac{1}{2} \cdot p = \frac{qp}{2},$$

  in which case the expected costs of revolt is $\beta^2/2$.

  Thus, to calculate the expected payoff of majority citizens inclusive of revolt costs, one needs to subtract $(qp/2)(\beta^2/2) = qp\beta^2/4$ from the policy payoff. Recalling that $\beta_c(\beta) = \beta - \beta^2/4$, the majority citizens' expected payoff inclusive of revolt costs is:

$$1 - q\left(1 - p\beta\right) - \frac{qp\beta^2}{4} = 1 - q\left(1 - p\beta_c\right).$$

Proposition 2 in the main text is therefore modified as follows.

Proposition 1. *In equilibrium,*

$$\sigma(\hat{s}, 1) = \sigma(\hat{s} = \emptyset, 0) = 1 \quad \textit{and} \quad \sigma(\hat{s} = s, 0) = \begin{cases} 0 & ; \ \beta(1, M, \gamma) > 1 - \delta_0 \\ 1 & ; \ \beta(1, M, \gamma) < 1 - \delta_0 \end{cases}$$

There is a revolt only if $\hat{s} = 0$ and the ruler takes action 1. This revolt succeeds with probability $\beta(1, M, \gamma)$. Moreover, the expected payoff for a majority citizen, inclusive of policy payoffs and revolt costs, is

$$\begin{cases} 1 - q(1 - p) & ; \ \beta(1, M, \gamma) > 1 - \delta_0 \\ 1 - q(1 - p\beta_c(1, M, \gamma)) & ; \ \beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Following the same steps, Proposition 3 in the main text is modified as follows.

Proposition 2. *Recall that A is the aggregate government action, and $Pr_{(t_1,t_2)}(A)$ is the probability of A conditional on rulers' types (t_1, t_2). In equilibrium,*

$$Pr_{(t_1,t_2)}(A = s) = 1, \quad \text{if } (t_1, t_2) \neq (b, b).$$

Otherwise,

$$Pr_{(b,b)}(A = 1|\hat{s}, s = 1) = Pr_{(b,b)}(A = 1|\hat{s} = \emptyset, s = 0) = 1$$

and

$$Pr_{(b,b)}(A = 1|\hat{s} = s, s = 0) = \begin{cases} 1 & ;\beta(1, M, \gamma) < 1 - \delta_0 \\ 0 & ;otherwise. \end{cases}$$

There is a revolt only if $\hat{s} = 0$ and both rulers take action 1. This revolt succeeds with probability $\beta(1, M, \gamma)$. Moreover, the expected payoff for a majority citizen, inclusive of policy payoffs and revolt costs, is

$$\begin{cases} 1 - q^2(1 - p) - \mu & ;\beta(1, M, \gamma) > 1 - \delta_0 \\ 1 - q^2(1 - p\beta_c(1, M, \gamma)) - \mu & ;\beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Corollary 1 in the main text is modified as follows.

Corollary 1. *The value of institutional constraints is:*

$$\begin{cases} (1 - p)(q - q^2) - \mu & ;\beta(1, M, \gamma) > 1 - \delta_0 \\ (1 - p\beta_c(1, M, \gamma))(q - q^2) - \mu & ;\beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Proposition 4 in the main text is modified as follows.

Proposition 3. *There is threshold $p^*(M, \gamma, q, \mu)$ such that a majority citizen's expected payoff, inclusive of policy payoffs and revolt costs, is higher without institutional constraints if and only if the scope of the divine law $p > p^*$, where*

$$p^*(M, \gamma, q, \mu) = \begin{cases} 1 - \frac{\mu}{q(1-q)} & ;\beta(1, M, \gamma) > 1 - \delta_0 \\ \\ \frac{1}{\beta_c(1,M,\gamma)}\left(1 - \frac{\mu}{q(1-q)}\right) & ;\beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Moreover,

1. *If $p^*(M, \gamma, q, \mu) > 0$, then $p^*(M, \gamma, q, \mu)$ is decreasing in M and γ; strictly so if and only if $\beta(1, M, \gamma) < 1 - \delta_0$.*
2. *$p^*(M, \gamma, q, \mu = 0) \geq 1$. For $\mu > 0$, $p^*(M, \gamma, q, \mu)$ has an inverted U-shape in q, with*

$$lim_{q\to 0^+} p^*(M, \gamma, q, \mu) = lim_{q\to 1^-} p^*(M, \gamma, q, \mu) = -\infty.$$

Proposition 5 in the main text is modified as follows.

Proposition 4. *There is a cost threshold such that the majority citizen's expected payoff, inclusive of policy payoffs and revolt costs, is higher without institutional constraints if and only if $\mu > \mu^*$, where*

$$\mu^*(\beta, p, q) = \begin{cases} (1-p)(q-q^2) & ;\beta > 1-\delta_0 \\ \\ (1-p(\beta-\beta^2/4))(q-q^2) & ;\beta < 1-\delta_0, \end{cases}$$

where $\beta = \beta(1, M, \gamma)$. Moreover,

1. *μ^* is strictly decreasing in p, and weakly decreasing in $\beta(1, M, \gamma)$ (and hence in M and γ); strictly so when $\beta < 1-\delta_0$.*
2. *Suppose $\delta_0 < T/M$, so that there is sufficient conflict of interest that the threat of revolt does not deter the minority-congruent ruler ($\beta < 1-\delta_0$). Then,*

$$\frac{\partial^2 \mu^*(\beta, p, q)}{\partial p \partial \beta} = -(q-q^2)\left(1-\frac{\beta}{2}\right) < 0.$$

Finally, Proposition 6 in the main text is modified as follows.

Proposition 5. *Suppose $\gamma \sim U[0,1]$. Let $Q = Pr_\gamma(\mu \leq \mu^*(\gamma))$ be the probability that institutional constraints improve the majority citizen's expected payoff, inclusive of policy payoffs and revolt costs. Suppose $\delta_0 < T/M$, so that there is sufficient conflict of interest that the threat of revolt does not deter the minority-congruent ($\beta < 1-\delta_0$). Let $\mu' = \mu/(q-q^2)$. Then,*

1. *Q is decreasing in μ', p and in M; strictly so when $\mu' \in (1-p\left((1-T/M) - \frac{1}{4}(1-T/M)^2\right), 1)$.*
2. *$|Q(\mu_2') - Q(\mu_1')|$ is strictly decreasing in p and M for all μ_1', μ_2', with $\mu_1', \mu_2' \in (1-p\left((1-T/M) - \frac{1}{4}(1-T/M)^2\right), 1)$.*

Proof. Using Proposition 4,

$$\begin{aligned} Q &= Pr_\gamma(\mu \leq \mu^*(\gamma) \mid \beta < 1-\delta_0) \\ &= Pr_\gamma\left(\mu \leq (1-p(\beta - \frac{\beta^2}{4}))(q-q^2)\right) \end{aligned}$$

Using the fact that $\beta = \beta(1, M, \gamma)$, and substituting Proposition 1 in the main text, we have: $\beta = H\left((1-\frac{T}{M})\gamma\right)$. Because $H = U[0,1]$, $\beta = (1-\frac{T}{M})\gamma$. Substituting, we have:

$$\begin{aligned} Q &= Pr_\gamma\left(\mu \leq (1-p(\gamma(1-\frac{T}{M}) - \frac{\gamma^2(1-\frac{T}{M})^2}{4}))(q-q^2)\right) \\ &= Pr_\gamma\left(\mu' \leq 1-p(\gamma(1-\frac{T}{M}) - \frac{\gamma^2(1-\frac{T}{M})^2}{4})\right) \\ &= Pr_\gamma\left(\gamma(1-\frac{T}{M}) - \frac{\gamma^2}{4}(1-\frac{T}{M})^2 \leq \frac{1-\mu'}{p}\right) \end{aligned}$$

For any $\gamma \in [0,1]$ and $M \in (T,1]$, let:

$$\zeta(\gamma, M) \equiv \gamma(1 - \frac{T}{M}) - \frac{\gamma^2}{4}(1 - \frac{T}{M})^2$$

Then,

$$Q = Pr_\gamma \left(\zeta(\gamma, M) \leq \frac{1-\mu'}{p} \right) \tag{1}$$

We continue with a few observations that will play a crucial role in the following arguments.

- $\zeta(\gamma, M)$ is strictly increasing in γ and M, because:

$$\begin{aligned} \frac{\partial \zeta}{\partial \gamma} &= 1 - \frac{T}{M} - \frac{\gamma}{2}(1 - \frac{T}{M})^2 && > 0 \\ \frac{\partial \zeta}{\partial M} &= \gamma \frac{T}{M^2} - \frac{\gamma^2}{2}(1 - \frac{T}{M})\frac{T}{M^2} && > 0 \end{aligned}$$

- $\zeta(\gamma, M)$ is strictly concave in γ, because:

$$\frac{\partial^2 \zeta}{\partial \gamma^2} = -\frac{1}{2}(1 - \frac{T}{M})^2 \qquad < 0$$

- $\zeta(\gamma, M)$ is supermodular in γ and M, because:

$$\frac{\partial^2 \zeta}{\partial \gamma \partial M} = \frac{T}{M^2} - \gamma(1 - \frac{T}{M})\frac{T}{M^2} \qquad > 0$$

  As a result, $\zeta(\gamma, M)$ satisfies strict increasing differences in (γ, M). That is, for any $\gamma_1 < \gamma_2$ and $M_1 < M_2$,

$$\zeta(\gamma_2, M_1) - \zeta(\gamma_1, M_1) < \zeta(\gamma_2, M_2) - \zeta(\gamma_1, M_2).$$

Because $\zeta(\gamma, M)$ is strictly increasing in γ, and since $\gamma \sim U[0,1]$, Equation (1) implies:

$$Q = \begin{cases} 0 & ; \ \frac{1-\mu'}{p} < \zeta(0, M) \\ \gamma^* \text{ s.t. } \zeta(\gamma^*, M) = \frac{1-\mu'}{p} & ; \ \zeta(0, M) \leq \frac{1-\mu'}{p} \leq \zeta(1, M) \\ 1 & ; \ \frac{1-\mu'}{p} > \zeta(1, M) \end{cases}$$

Substituting $\zeta(0, M) = 0$ and $\zeta(1, M) = (1 - \frac{T}{M}) - \frac{1}{4}(1 - \frac{T}{M})^2$,

$$Q = \begin{cases} 0 & ; \ \frac{1-\mu'}{p} < 0 \\ \gamma^* \text{ s.t. } \zeta(\gamma^*, M) = \frac{1-\mu'}{p} & ; \ 0 \leq \frac{1-\mu'}{p} \leq (1 - \frac{T}{M}) - \frac{1}{4}(1 - \frac{T}{M})^2 \\ 1 & ; \ \frac{1-\mu'}{p} > (1 - \frac{T}{M}) - \frac{1}{4}(1 - \frac{T}{M})^2 \end{cases}$$

Rearranging,

$$Q = \begin{cases} 1 & ;\ \mu' < 1 - p\left((1 - \frac{T}{M}) - \frac{1}{4}(1 - \frac{T}{M})^2\right) \\ \gamma^* \text{ s.t. } \zeta(\gamma^*, M) = \frac{1-\mu'}{p} & ;\ \mu' \in [1 - p\left((1 - \frac{T}{M}) - \frac{1}{4}(1 - \frac{T}{M})^2\right), 1] \\ 0 & ;\ \mu' > 1 \end{cases}$$

The fact that Q is decreasing in μ' and p, strictly so when $\mu' \in (1-p\left((1 - \frac{T}{M}) - \frac{1}{4}(1 - \frac{T}{M})^2\right), 1)$, follows from $\zeta(\gamma, M)$ being strictly increasing in γ. Moreover, the fact that Q is decreasing in M, strictly so when $\mu' \in (1 - p\left((1 - \frac{T}{M}) - \frac{1}{4}(1 - \frac{T}{M})^2\right), 1)$, follows from $\zeta(\gamma, M)$ being strictly increasing in γ and strictly increasing in M.

Next, we show that $|Q(\mu_2') - Q(\mu_1')|$ is strictly decreasing in p for all μ_1', μ_2', with $\mu_1', \mu_2' \in (1-p\left((1 - T/M) - \frac{1}{4}(1 - T/M)^2\right), 1)$. Without loss of generality, take $\mu_1' < \mu_2'$ and $p_1 < p_2$ such that $\mu_1', \mu_2' \in (1 - p_1\left((1 - T/M) - \frac{1}{4}(1 - T/M)^2\right), 1)$.

Let $Q(\mu_1' \mid p_1)$ and $Q(\mu_2' \mid p_1)$ denote the relevant probabilities under p_1. Note that $Q(\mu_1' \mid p_1) = \gamma_{11}$ and $Q(\mu_2' \mid p_1) = \gamma_{21}$, where:

$$\zeta(\gamma_{11}, M) = \frac{1 - \mu_1'}{p_1} \qquad \zeta(\gamma_{21}, M) = \frac{1 - \mu_2'}{p_1}$$

Similarly, $Q(\mu_1' \mid p_2) = \gamma_{12}$ and $Q(\mu_2' \mid p_2) = \gamma_{22}$, where:

$$\zeta(\gamma_{12}, M) = \frac{1 - \mu_1'}{p_2} \qquad \zeta(\gamma_{22}, M) = \frac{1 - \mu_2'}{p_2}$$

Now,

$$\zeta(\gamma_{11}, M) - \zeta(\gamma_{21}, M) = \frac{\mu_2' - \mu_1'}{p_1} > \frac{\mu_2' - \mu_1'}{p_2} = \zeta(\gamma_{12}, M) - \zeta(\gamma_{22}, M)$$

Therefore,

$$\zeta(\gamma_{11}, M) - \zeta(\gamma_{21}, M) > \zeta(\gamma_{12}, M) - \zeta(\gamma_{22}, M) \tag{2}$$

Because Q is strictly decreasing in μ' in the range considered, $\gamma_{11} > \gamma_{21}$ and $\gamma_{12} > \gamma_{22}$. Because Q is strictly decreasing in p in the range considered, $\gamma_{11} > \gamma_{12}$ and $\gamma_{21} > \gamma_{22}$. Finally, recall that $\zeta(\gamma, M)$ is strictly concave in γ. For Equation (2) to hold, therefore, one must have: $\gamma_{11} - \gamma_{21} > \gamma_{12} - \gamma_{22}$. Therefore,

$$\begin{aligned} |Q(\mu_2' \mid p_1) - Q(\mu_1' \mid p_1)| &= |\gamma_{21} - \gamma_{11}| \\ &= \gamma_{11} - \gamma_{21} \\ &> \gamma_{12} - \gamma_{22} \\ &= |\gamma_{22} - \gamma_{12}| \\ &= |Q(\mu_2' \mid p_2) - Q(\mu_1' \mid p_2)| \end{aligned}$$

and the result follows.

Finally, we show that $|Q(\mu_2') - Q(\mu_1')|$ is strictly decreasing in M for all μ_1', μ_2', with $\mu_1', \mu_2' \in (1 - p\left((1 - T/M) - \frac{1}{4}(1 - T/M)^2\right), 1)$. Without loss of generality, take $\mu_1' < \mu_2'$ and $M_1 < M_2$ such that $\mu_1', \mu_2' \in (1 - p\left((1 - T/M_1) - \frac{1}{4}(1 - T/M_1)^2\right), 1)$.

Let $Q(\mu_1' \mid M_1)$ and $Q(\mu_2' \mid M_1)$ denote the relevant probabilities under M_1. Note that $Q(\mu_1' \mid M_1) = \gamma_{11}$ and $Q(\mu_2' \mid M_1) = \gamma_{21}$, where:

$$\zeta(\gamma_{11}, M_1) = \frac{1 - \mu_1'}{p} \qquad \zeta(\gamma_{21}, M_1) = \frac{1 - \mu_2'}{p}$$

Similarly, $Q(\mu_1' \mid M_2) = \gamma_{12}$ and $Q(\mu_2' \mid M_2) = \gamma_{22}$, where:

$$\zeta(\gamma_{12}, M_2) = \frac{1 - \mu_1'}{p} \qquad \zeta(\gamma_{22}, M_2) = \frac{1 - \mu_2'}{p}$$

Now,

$$\zeta(\gamma_{11}, M_1) - \zeta(\gamma_{21}, M_1) = \frac{\mu_2' - \mu_1'}{p} = \zeta(\gamma_{12}, M_2) - \zeta(\gamma_{22}, M_2) \tag{3}$$

Because Q is strictly decreasing in μ' in the range considered, $\gamma_{11} > \gamma_{21}$. Because $\zeta(\gamma, M)$ satisfies strictly increasing differences in (γ, M), $\zeta(\gamma_{11}, M_1) - \zeta(\gamma_{21}, M_1) < \zeta(\gamma_{11}, M_2) - \zeta(\gamma_{21}, M_2)$. This, along with Equation (3), implies:

$$\zeta(\gamma_{11}, M_2) - \zeta(\gamma_{21}, M_2) > \zeta(\gamma_{12}, M_2) - \zeta(\gamma_{22}, M_2) \tag{4}$$

Because Q is strictly decreasing in μ' in the range considered, $\gamma_{12} > \gamma_{22}$. Because Q is strictly decreasing in M in the range considered, $\gamma_{11} > \gamma_{12}$ and $\gamma_{21} > \gamma_{22}$. Finally, recall that $\zeta(\gamma, M)$ is strictly concave in γ. For Equation (4) to hold, therefore, one must have: $\gamma_{11} - \gamma_{21} > \gamma_{12} - \gamma_{22}$. Therefore,

$$\begin{aligned} |Q(\mu_2' \mid M_1) - Q(\mu_1' \mid M_1)| &= |\gamma_{21} - \gamma_{11}| \\ &= \gamma_{11} - \gamma_{21} \\ &> \gamma_{12} - \gamma_{22} \\ &= |\gamma_{22} - \gamma_{12}| \\ &= |Q(\mu_2' \mid M_2) - Q(\mu_1' \mid M_2)| \end{aligned}$$

and the result follows. □

B An Alternative Model of Institutional Constraints

In this section, we present an alternative model with institutional constraints and provide a characterization. Throughout this section, we maintain our assumption that $0 = \delta_1 < \delta_0 < 1$ in the main text. The difference is that we consider $y(a_1, a_2) = \max\{a_1, a_2\}$. That is, in the setup considered here, if one of the rulers choose the minority-congruent policy $a_i = 1$, the aggregate policy is $A = 1$. A majority-congruent ruler, therefore, does not have the blocking power by himself. However, since citizens observe (a_1, a_2), they can still receive information from the majority-congruent ruler's proposed policy and base their revolt decisions on this information. In this sense, the institutional arrangement has a learning benefit for the citizens.

B.1 Formal Definition of Equilibrium

The majority-congruent ruler $j \in \{1, 2\}$ (i.e., ruler j of type $t_j = g$) always chooses $a_j = s$ by assumption.

The strategy of the minority-congruent ruler 1 (i.e., ruler 1 of type $t_1 = b$) in state s, when public signal is $\hat{s}$ and ruler 2's type is $t_2 \in \{b, g\}$ is:

$$\sigma_1(\widehat{s}, s, t_2) \equiv \Pr(a_1 = 1 | s, \hat{s}, t_2) \quad \in [0, 1]$$

The strategy of minority-congruent ruler 2 (i.e., ruler 2 of type $t_2 = b$) in state s, given the public signal is $\hat{s}$ and ruler 1's action a_1 is:[1]

$$\sigma_2(\widehat{s}, s, a_1) \equiv \Pr(a_2 = 1 | s, \hat{s}, a_1) \quad \in [0, 1]$$

The posterior beliefs of citizens that the aggregate policy is incongruent, given information $(\hat{s}, a_1, a_2)$, is denoted by:

$$q(\hat{s}, a_1, a_2) \equiv \Pr\left(\max\{a_1, a_2\} \neq s | \hat{s}, a_1, a_2\right) \quad \in [0, 1]$$

The strategy of a citizen i when with posterior beliefs q' and the cost of revolt is c_i is denoted by:

$$\varphi(q', c_i) \equiv \Pr(r_i = 1 | q', c_i) \quad \in [0, 1]$$

The Perfect Bayesian Nash Equilibrium of the game is a quadruple $(\sigma_1^*, \sigma_2^*, \varphi^*, q^*)$ such that the following are satisfied.

1. $\varphi^*(q', c_i)$ maximizes the payoff of the citizens in majority for any $q' = q^*(\hat{s}, a_1, a_2)$.
2. $q^*(\hat{s}, a_1, a_2)$ is given by Bayes' Rule.
3. Given φ^* and σ_2^*, σ_1^* maximizes the payoff of the minority-congruent ruler 1. Similarly, given φ^* and σ_1^*, σ_2^* maximizes the payoff of the minority-congruent ruler 2.

[1] As discussed in the main text, ruler 2's strategy may also condition on t_1. However, because t_1 is not payoff-relevant for ruler 2, the dependence can be dropped.

We consider the symmetric cutoff strategy equilibrium with cutoffs greater than one as $\rho \to 0$. Once again, there are multiple equilibria in this model. As an equilibrium selection device, we impose the following assumption on the minority-congruent ruler.

Assumption 1. *When a minority-congruent ruler j is indifferent between the two actions, he chooses $a_j = 1$ with probability 1.*

Assumption 1 is a mild restriction on the minority-congruent ruler's behavior: it applies only when the ruler is indifferent between the two actions. It can be microfounded by assuming that the minority-congruent ruler j obtains some infinitesimal material payoff from taking action $a_j = 1$.

B.2 Equilibrium Characterization

B.2.1 Citizens' Actions

As we will show later, the members of minority never take part in a revolution in equilibrium. Therefore, the only citizens who potentially participate in a revolution are majority citizens, whose size is M. As discussed in Proposition 1 in the main text, in a symmetric cutoff strategy equilibrium as $\rho \to 0$, a successful revolution occurs with probability:

$$\beta(q', M, \gamma) = H\left((1 - \frac{T}{M}) \cdot \gamma \cdot (2q' - 1)\right)$$

B.2.2 Beliefs Following Proposed Policy

When $\hat{s} \in \{0, 1\}$, $q^*(\hat{s}, a_1, a_2) = |\hat{s} - \max\{a_1, a_2\}| \in \{0, 1\}$. When $\hat{s} = \emptyset$, the posterior beliefs are given by:

$$\begin{aligned}
q^*(\emptyset, 0, 0) &\equiv \Pr(\max\{a_1, a_2\} \neq s | a_1 = a_2 = 0, \hat{s} = \emptyset) \\
&= \Pr(s = 1 | a_1 = a_2 = 0, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 1, a_1 = a_2 = 0, \hat{s} = \emptyset)}{\Pr(s = 1, a_1 = a_2 = 0, \hat{s} = \emptyset) + \Pr(s = 0, a_1 = a_2 = 0, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2} q^2 (1 - \sigma_1^*(\emptyset, 1, b))(1 - \sigma_2^*(\emptyset, 1, 0))}{\frac{1}{2} q^2 (1 - \sigma_1^*(\emptyset, 1, b))(1 - \sigma_2^*(\emptyset, 1, 0)) + \frac{1}{2}\left(q^2 (1 - \sigma_1^*(\emptyset, 0, b))(1 - \sigma_2^*(\emptyset, 0, 0)) + q(1 - q)(1 - \sigma_1^*(\emptyset, 0, g)) + (1 - q) q (1 - \sigma_2^*(\emptyset, 0, 0)) + (1 - q)^2\right)} \\
&= \frac{q^2 (1 - \sigma_1^*(\emptyset, 1, b))(1 - \sigma_2^*(\emptyset, 1, 0))}{q^2 (1 - \sigma_1^*(\emptyset, 1, b))(1 - \sigma_2^*(\emptyset, 1, 0)) + \left(q^2 (1 - \sigma_1^*(\emptyset, 0, b))(1 - \sigma_2^*(\emptyset, 0, 0)) + q(1 - q)(1 - \sigma_1^*(\emptyset, 0, g)) + (1 - q) q (1 - \sigma_2^*(\emptyset, 0, 0)) + (1 - q)^2\right)}
\end{aligned}$$

$$\begin{aligned}
q^*(\emptyset, 0, 1) &\equiv \Pr(\max\{a_1, a_2\} \neq s | a_1 = 0, a_2 = 1, \hat{s} = \emptyset) \\
&= \Pr(s = 0 | a_1 = 0, a_2 = 1, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 0, a_1 = 0, a_2 = 1, \hat{s} = \emptyset)}{\Pr(s = 0, a_1 = 0, a_2 = 1, \hat{s} = \emptyset) + \Pr(s = 1, a_1 = 0, a_2 = 1, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2} q^2 (1 - \sigma_1^*(\emptyset, 0, b)) \sigma_2^*(\emptyset, 0, 0) + (1 - q) q (1 - \sigma_2^*(\emptyset, 0, 0))}{\frac{1}{2} q^2 (1 - \sigma_1^*(\emptyset, 0, b)) \sigma_2^*(\emptyset, 0, 0) + (1 - q) q (1 - \sigma_2^*(\emptyset, 0, 0)) + \frac{1}{2}\left(q^2 (1 - \sigma_1^*(\emptyset, 1, b)) \sigma_2^*(\emptyset, 1, 0) + q(1 - q)(1 - \sigma_1^*(\emptyset, 1, g))\right)}
\end{aligned}$$

$$\begin{aligned}
q^*(\emptyset, 1, 0) &\equiv \Pr(\max\{a_1, a_2\} \neq s | a_1 = 1, a_2 = 0, \hat{s} = \emptyset) \\
&= \Pr(s = 0 | a_1 = 1, a_2 = 0, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 0, a_1 = 1, a_2 = 0, \hat{s} = \emptyset)}{\Pr(s = 0, a_1 = 1, a_2 = 0, \hat{s} = \emptyset) + \Pr(s = 1, a_1 = 1, a_2 = 0, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2} q^2 \sigma_1^*(\emptyset, 0, b)(1 - \sigma_2^*(\emptyset, 0, 1)) + q(1 - q) \sigma_1^*(\emptyset, 0, g)}{\frac{1}{2} q^2 \sigma_1^*(\emptyset, 0, b)(1 - \sigma_2^*(\emptyset, 0, 1)) + q(1 - q) \sigma_1^*(\emptyset, 0, g) + \frac{1}{2}\left(q^2 \sigma_1^*(\emptyset, 1, b)(1 - \sigma_2^*(\emptyset, 1, 1)) + (1 - q) q (1 - \sigma_2^*(\emptyset, 1, 1))\right)}
\end{aligned}$$

$$
\begin{aligned}
q^*(\emptyset,1,1) &\equiv \Pr(\max\{a_1,a_2\} \neq s | a_1 = a_2 = 1, \hat{s} = \emptyset) \\
&= \Pr(s = 0 | a_1 = a_2 = 1, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 0, a_1 = a_2 = 1, \hat{s} = \emptyset)}{\Pr(s = 0, a_1 = a_2 = 1, \hat{s} = \emptyset) + \Pr(s = 1, a_1 = a_2 = 1, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2} q^2 \sigma_1^*(\emptyset,0,b)\sigma_2^*(\emptyset,0,1)}{\frac{1}{2} q^2 \sigma_1^*(\emptyset,0,b)\sigma_2^*(\emptyset,0,1) + \frac{1}{2}\left(q^2 \sigma_1^*(\emptyset,1,b)\sigma_2^*(\emptyset,1,1) + q(1-q)\sigma_1^*(\emptyset,1,g) + (1-q)q\sigma_2^*(\emptyset,1,1) + (1-q)^2\right)}
\end{aligned}
$$

B.2.3 Rulers' Actions

When the Issue is Preordained We begin by pinning down the strategies of minority-congruent ruler 2 at every history.

1. Consider the case $\hat{s} = s = 0$ and $a_1 = 0$. In this case, $\max\{a_1, a_2\} = a_2$ and $q^*(0,0,a_2) = a_2$ for any $a_2 \in \{0,1\}$.

   The majority members never revolt against $a_2 = 0$, and since $M > 1/2$, there is never a revolt against $a_2 = 0$. In contrast, the minority members never revolt against $a_2 = 1$, and therefore the probability of a successful revolt against $a_2 = 1$ is $\beta(1, M, \gamma)$. Thus, ruler 2's policy when $(\hat{s}, s, a_1) = (0,0,0)$ is:

   $$\sigma_2^*(0,0,0) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(1, M, \gamma)) + (1 - \sigma) \cdot \delta_0$$

   Therefore, ruler 2's PBE strategy is:

   $$\sigma_2^*(0,0,0) = \begin{cases} 0 & ; \delta_0 > 1 - \beta(1, M, \gamma) \\ 1 & ; \delta_0 < 1 - \beta(1, M, \gamma) \end{cases}$$

2. Consider the case $\hat{s} = s = 0$ and $a_1 = 1$. In this case, $\max\{a_1, a_2\} = 1$ regardless of a_2, and $q^*(0,1,a_2) = 1$ for any $a_2 \in \{0,1\}$. Ruler 2 is indifferent between the two actions, and by Assumption 1, $\sigma_2^*(0,0,1) = 1$.

3. Consider the case $\hat{s} = s = 1$ and $a_1 = 0$. In this case, $\max\{a_1, a_2\} = a_2$ and $q^*(1,0,a_2) = 1 - a_2$ for any $a_2 \in \{0,1\}$.

   Because $\delta_1 = 0$, ruler 2 receives a payoff of 0 if he chooses $a_2 = 0$. If he chooses $a_2 = 1$, the citizens will not revolt, and ruler 2 will receive a payoff of 1. Therefore, $\sigma_2^*(1,1,0) = 1$.

4. Consider the case $\hat{s} = s = 1$ and $a_1 = 1$. In this case, $\max\{a_1, a_2\} = 1$ regardless of a_2, and $q^*(1,1,a_2) = 0$ for any $a_2 \in \{0,1\}$. Ruler 2 is indifferent between the two actions, and by Assumption 1, $\sigma_2^*(1,1,1) = 1$.

Next, we pin down the strategy of minority-congruent ruler 1 in every history.

1. Consider the case $\hat{s} = s = 0$ and $t_2 = g$. In this case, $a_2 = 0$, and $\max\{a_1, a_2\} = a_1 \in \{0,1\}$. Moreover, $q^*(0,a_1,a_2) = a_1$ for any $a_1 \in \{0,1\}$.

   The majority members never revolt against $a_1 = 0$, and since $M > 1/2$, there is never a revolt against $a_1 = 0$. The minority members never revolt against $a_1 = 1$, and

therefore the probability of a successful revolt against $a_1 = 1$ is $\beta(1, M, \gamma)$. Thus, ruler 1's policy when $(\hat{s}, s, t_2) = (0, 0, g)$ is:

$$\sigma_1^*(0, 0, g) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(1, M, \gamma)) + (1 - \sigma) \cdot \delta_0$$

Therefore, ruler 1's PBE strategy is:

$$\sigma_1^*(0, 0, g) = \begin{cases} 0 & ; \delta_0 > 1 - \beta(1, M, \gamma) \\ 1 & ; \delta_0 < 1 - \beta(1, M, \gamma) \end{cases}$$

2. Consider the case $\hat{s} = s = 0$ and $t_2 = b$. In this case,

$$a_2 = \begin{cases} 0 & ; \delta_0 > 1 - \beta(1, M, \gamma) \\ 1 & ; \delta_0 < 1 - \beta(1, M, \gamma) \end{cases}$$

   - If $\delta_0 > 1 - \beta(1, M, \gamma)$, ruler 1's optimal strategy is:

$$\sigma_1^*(0, 0, b) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(1, M, \gamma)) + (1 - \sigma) \cdot \delta_0$$

     which is maximized when $\sigma_1^*(0, 0, b) = 0$.

   - If $\delta_0 < 1 - \beta(1, M, \gamma)$, $\max\{a_1, a_2\} = 1$ for any $a_1 \in \{0, 1\}$ in any PBE. Ruler 1 is indifferent between the two actions. By Assumption 1, $\sigma_1^*(0, 0, b) = 1$.

3. Consider the case $\hat{s} = s = 1$ and $t_2 = g$. In this case, $a_2 = 1$, and $\max\{a_1, a_2\} = 1$ for any $a_1 \in \{0, 1\}$ in any PBE. Ruler 1 is indifferent between the two actions. By Assumption 1, $\sigma_1^*(1, 1, g) = 1$.

4. Consider the case $\hat{s} = s = 1$ and $t_2 = b$. Since $\sigma_2^*(1, 1, 0) = \sigma_2^*(1, 1, 1) = 1$, $a_2 = 1$ with probability one. Then, $\max\{a_1, a_2\} = 1$ for any $a_1 \in \{0, 1\}$ in any PBE. Ruler 1 is indifferent between the two actions, and by Assumption 1, $\sigma_1^*(1, 1, b) = 1$.

Note that $\sigma_1^*(1, 1) = \sigma_2^*(1, 1, 1) = 1$ in any PBE. That is, when $\hat{s} = s = 1$, the aggregate policy is $A = 1$ with probability one.

If $\delta_0 < 1 - \beta(1, M, \gamma)$, $\sigma_1^*(0, 0, b) = \sigma_1^*(0, 0, g) = \sigma_2^*(0, 0, 0) = 1$. That is, when $\hat{s} = s = 0$, the aggregate policy taken by two rulers, when at least one of them is minority-congruent, is $A = 1$ with probability one. This is accompanied by a revolt with probability $\beta(1, M, \gamma)$.

If $\delta_0 > 1 - \beta(1, M, \gamma)$, $\sigma_1^*(0, 0, b) = \sigma_1^*(0, 0, g) = \sigma_2^*(0, 0, 0) = 0$. That is, when $\hat{s} = s = 0$, the aggregate policy is $A = 0$ with probability one.

When the Issue is Non-Preordained We begin this analysis with two observations, which will considerably simplify the following arguments.

Remark 1. *In any PBE, $\sigma_2^*(\emptyset, 1, 0) = 1$. This is because when $s = 1$ and $a_1 = 0$, choosing $a_2 = 0$ yields a payoff of 0 to ruler 2 (recall that $\delta_1 = 0$). On the other hand, choosing $a_2 = 1$ yields a strictly positive payoff because the probability of revolt is strictly less than one.*

Remark 2. *In any PBE, $q^*(\emptyset, 0, 0) = 0$. This follows from Remark 1 and the Equation defining $q^*(\emptyset, 0, 0)$ in Section B.2.2. In words, the citizens know that when $s = 1$, ruler 2 follows up $a_1 = 0$ with $a_2 = 1$. Therefore, whenever $a_1 = 0$ is followed up with $a_2 = 0$, the citizens deduce that the state is $s = 0$.*

By Remark 2, the majority citizens do not revolt upon observing $(\hat{s}, a_1, a_2) = (\emptyset, 0, 0)$. Since $M > 1/2$, the minority members do not revolt either, and there are no revolts. In any other $(\hat{s}, a_1, a_2) = (\emptyset, a_1, a_2)$, the aggregate action is $A = 1$. The minority citizens never attempt revolt against this action, and thus the only citizens possibly attempting revolt are the majority citizens. The probability of revolt is given by $\beta(q', M, \gamma)$.

Given these observations, the equilibrium strategy of minority-congruent ruler 2 the remaining histories is characterized by the following equations.

$$\sigma_2^*(\emptyset, 0, 0) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset, 0, 1), M, \gamma)) + (1 - \sigma) \cdot \delta_0 \tag{5}$$

$$\sigma_2^*(\emptyset, 0, 1) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset, 1, 1), M, \gamma)) + (1 - \sigma) \cdot (1 - \beta(q^*(\emptyset, 1, 0), M, \gamma)) \tag{6}$$

$$\sigma_2^*(\emptyset, 1, 1) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset, 1, 1), M, \gamma)) + (1 - \sigma) \cdot (1 - \beta(q^*(\emptyset, 1, 0), M, \gamma)) \tag{7}$$

We continue with two observations.

- In any PBE, $1 - \beta(q^*(\emptyset, 1, 1), M) \geq 1 - \beta(q^*(\emptyset, 1, 0), M)$. To see this, suppose not: suppose $1 - \beta(q^*(\emptyset, 1, 1), M) < 1 - \beta(q^*(\emptyset, 1, 0), M)$. Then, by (6), $\sigma_2^*(\emptyset, 0, 1) = 0$. Then, by the equation defining $q^*(\emptyset, 1, 1)$ in Section B.2.2, $q^*(\emptyset, 1, 1) = 0$. But then, $\beta(q^*(\emptyset, 1, 1), M) = 0$, a contradiction.
- The observation above, along with Assumption 1, implies that $\sigma_2^*(\emptyset, 0, 1) = \sigma_2^*(\emptyset, 1, 1) = 1$ in any PBE.

The only part of ruler 2's PBE strategy we have not pinned down so far is $\sigma_2^*(\emptyset, 0, 0)$.

We now proceed with ruler 1. For the equilibrium strategy of minority-congruent ruler 1, consider four possible histories.

1. Consider the case when $\hat{s} = \emptyset$, $s = 0$ and $t_2 = g$. Ruler 2 chooses $a_2 = 0$ with probability one, and the aggregate action is $A = \max\{a_1, a_2\} = a_1$.

   If ruler 1 chooses $a_1 = 1$, there is a revolt with probability $\beta(q^*(\emptyset, 1, 0), M, \gamma)$. If ruler 1 chooses $a_1 = 0$, there is a revolt with probability $\beta(q^*(\emptyset, 0, 0), M, \gamma) = 0$. Therefore, minority-congruent ruler 1's optimal strategy when $(\hat{s}, s, t_2) = (\emptyset, 0, g)$ is:

   $$\sigma_1^*(\emptyset, 0, g) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset, 1, 0), M, \gamma)) + (1 - \sigma) \cdot \delta_0 \tag{8}$$

2. Consider the case when $\hat{s} = \emptyset$, $s = 0$ and $t_2 = b$.

   If ruler 1 chooses $a_1 = 1$, ruler 2 will follow with $a_2 = 1$ with probability one, because we established that $\sigma_2^*(\emptyset, 0, 1) = 1$. The aggregate action will be $A = 1$ and there will be a revolt with probability $\beta(q^*(\emptyset, 1, 1), M, \gamma)$.

If ruler 1 chooses $a_1 = 0$, ruler 2 will follow with a_2 with probability $\sigma_2^*(\emptyset, 0, 0)$. The aggregate action will be a_2, and ruler 1's payoff will be:

$$\sigma_2^*(\emptyset, 0, 0) \cdot (1 - \beta(q^*(\emptyset, 0, 1), M, \gamma)) + (1 - \sigma_2^*(\emptyset, 0, 0)) \cdot \delta_0$$

which, by (5), equals: $\max\{1 - \beta(q^*(\emptyset, 0, 1), M, \gamma), \delta_0\}$.

Therefore, minority-congruent ruler 1's optimal strategy when $(\hat{s}, s, t_2) = (\emptyset, 0, b)$ is:

$$\sigma_1^*(\emptyset, 0, b) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset, 1, 1), M, \gamma)) + (1 - \sigma) \cdot \max\{1 - \beta(q^*(\emptyset, 0, 1), M, \gamma), \delta_0\} \tag{9}$$

3. Consider the case when $\hat{s} = \emptyset$, $s = 1$ and $t_2 = g$. Ruler 2 chooses $a_2 = 1$ with probability one, and the aggregate action will be $A = 1$ with probability one.

   If ruler 1 chooses $a_1 = 1$, there will be a revolt with probability $\beta(q^*(\emptyset, 1, 1), M, \gamma)$. If ruler 1 chooses $a_1 = 0$, there will be a revolt with probability $\beta(q^*(\emptyset, 0, 1), M, \gamma)$. Therefore, minority-congruent ruler 1's optimal strategy when $(\hat{s}, s, t_2) = (\emptyset, 1, g)$ is:

$$\sigma_1^*(\emptyset, 1, g) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset, 1, 1), M, \gamma)) + (1 - \sigma) \cdot (1 - \beta(q^*(\emptyset, 0, 1), M, \gamma)) \tag{10}$$

4. Consider the case when $\hat{s} = \emptyset$, $s = 1$ and $t_2 = b$. We have already established that $\sigma_2^*(\emptyset, 1, 0) = \sigma_2^*(\emptyset, 1, 1) = 1$ in any PBE. Thus, ruler 2 chooses $a_2 = 1$ with probability one, and the aggregate action will be $A = 1$ with probability one.

   If ruler 1 takes $a_1 = 1$, there will be a revolt with probability $\beta(q^*(\emptyset, 1, 1), M, \gamma)$. If ruler 1 chooses $a_1 = 0$, there will be a revolt with probability $\beta(q^*(\emptyset, 0, 1), M, \gamma)$. Therefore, minority-congruent ruler 1's optimal strategy when $(\hat{s}, s, t_2) = (\emptyset, 1, b)$ is:

$$\sigma_1^*(\emptyset, 1, b) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset, 1, 1), M, \gamma)) + (1 - \sigma) \cdot (1 - \beta(q^*(\emptyset, 0, 1), M, \gamma)) \tag{11}$$

Once again, we continue with two observations.

- In any PBE, $1 - \beta(q^*(\emptyset, 1, 1), M, \gamma) \geq 1 - \beta(q^*(\emptyset, 0, 1), M, \gamma)$. To see this, suppose not: suppose $1 - \beta(q^*(\emptyset, 1, 1), M, \gamma) < 1 - \beta(q^*(\emptyset, 0, 1), M, \gamma)$. Then, by (9), $\sigma_1^*(\emptyset, 0, b) = 0$. Then, by the equation defining $q^*(\emptyset, 1, 1)$ in Section B.2.2, $q^*(\emptyset, 1, 1) = 0$. But then, $\beta(q^*(\emptyset, 1, 1), M, \gamma) = 0$, a contradiction.
- The observation above, along with Assumption 1, implies that $\sigma_1^*(\emptyset, 1, g) = \sigma_1^*(\emptyset, 1, b) = 1$ in any PBE.

So far we have argued that $\sigma_2^*(\emptyset, 0, 1) = \sigma_2^*(\emptyset, 1, 1) = \sigma_1^*(\emptyset, 1, g) = \sigma_1^*(\emptyset, 1, b) = 1$. Substituting these into the equation defining $q^*(\emptyset, 1, 1)$ in Section B.2.2,

$$q^*(\emptyset,1,1) = \frac{q^2\sigma_1^*(\emptyset,0,b)\sigma_2^*(\emptyset,0,1)}{q^2\sigma_1^*(\emptyset,0,b)\sigma_2^*(\emptyset,0,1) + (q^2\sigma_1^*(\emptyset,1,b)\sigma_2^*(\emptyset,1,1) + q(1-q)\sigma_1^*(\emptyset,1,g) + (1-q)q\sigma_2^*(\emptyset,1,1) + (1-q)^2)}$$
$$= \frac{q^2\sigma_1^*(\emptyset,0,b)}{q^2\sigma_1^*(\emptyset,0,b) + q^2 + q(1-q) + (1-q)q + (1-q)^2} \leq \frac{1}{2}$$

Therefore, $\beta(q^*(\emptyset,1,1),M,\gamma) = 0$. This, along with Assumption 1 and $\delta_0 < 1$, implies that $\sigma_1^*(\emptyset,0,b) = 1$ in any PBE.

The only part of ruler 1's PBE strategy we have not pinned down so far is $\sigma_1^*(\emptyset,0,g)$. The rest of the analysis considers two separate cases.

- Suppose $\delta_0 < 1 - \beta(1,M,\gamma)$. By (5), $\sigma_2^*(\emptyset,0,0) = 1$. By (8), $\sigma_1^*(\emptyset,0,g) = 1$. This completes the characterization of equilibrium strategies.

  Note that under these strategies, $q^*(\emptyset,1,0) = q^*(\emptyset,0,1) = 1$. Therefore, whenever $(\hat{s},s) = (\emptyset,0)$ and $t_1 \neq t_2$, there is a mismatch in the actions, and there is a successful revolution with probability $\beta(1,M,\gamma)$.

- Suppose $\delta_0 > 1 - \beta(1,M,\gamma)$.

  Our first claim is that $\sigma_2^*(\emptyset,0,0) = 0$. To see why, suppose not: $\sigma_2^*(\emptyset,0,0) > 0$. Given the strategies pinned down so far and the equation defining $q^*(\emptyset,0,1)$ in Section B.2.2, $q^*(\emptyset,0,1) = 1$. But then, by (5), $\sigma_2^*(\emptyset,0,0) = 0$, a contradiction. On the other hand, when $\sigma_2^*(\emptyset,0,0) = 0$, the history $(\emptyset,0,1)$ is never reached on equilibrium path. The Bayes' rule does not apply to $q^*(\emptyset,0,1)$. Then, any choice of $q^*(\emptyset,0,1)$ high enough so that $1 - \beta(q^*(\emptyset,0,1),M,\gamma) \leq \delta_0$ is consistent with $\sigma_2^*(\emptyset,0,0) = 0$ as an equilibrium strategy.

  Next, we similarly claim that $\sigma_1^*(\emptyset,0,g) = 0$. Suppose not: $\sigma_1^*(\emptyset,0,g) > 0$. Given the strategies pinned down so far and the equation defining $q^*(\emptyset,1,0)$ in Section B.2.2, $q^*(\emptyset,1,0) = 1$. But then, by (8), $\sigma_1^*(\emptyset,0,g) = 0$, a contradiction. On the other hand, when $\sigma_1^*(\emptyset,0,g) = 0$, the history $(\emptyset,1,0)$ is never reached on equilibrium path. The Bayes' rule does not apply to $q^*(\emptyset,1,0)$. Then, any choice of $q^*(\emptyset,1,0)$ high enough so that $1 - \beta(q^*(\emptyset,1,0),M,\gamma) \leq \delta_0$ is consistent with $\sigma_1^*(\emptyset,0,g) = 0$ as an equilibrium strategy.

  We conclude that $\sigma_2^*(\emptyset,0,0) = \sigma_1^*(\emptyset,0,g) = 0$ in any PBE. This completes the characterization of equilibrium strategies.

  Note that under these strategies, whenever $(\hat{s},s) = (\emptyset,0)$ and $t_1 \neq t_2$, the aggregate action is $A = 0$ and there are no revolts.

Our findings imply the following result.

Proposition 6. *Recall that A is the aggregate government action, and $Pr_{(t_1,t_2)}(A)$ be the equilibrium probability of A conditional on rulers' types (t_1,t_2).*

- *When $\beta(1,M,\gamma) > 1 - \delta_0$, the equilibrium outcomes are identical to those of the model*

in the main text. That is, in equilibrium,

$$Pr_{(t_1,t_2)}(A = s) = 1, \quad \text{if } (t_1, t_2) \neq (b, b).$$

Otherwise,

$$Pr_{(b,b)}(A = s|\hat{s} = s) = Pr_{(b,b)}(A = 1|\hat{s} = \emptyset) = 1$$

There are no revolts in equilibrium.

- *When* $\beta(1, M, \gamma) < 1 - \delta_0$,

$$Pr_{(t_1,t_2)}(A = 1) = 1, \quad \text{if } (t_1, t_2) \neq (g, g).$$

There is a revolt when $\hat{s} = 0$ *and at least one ruler takes action* 1*, and when* $\hat{s} = \emptyset$ *and the rulers' actions do not match each other. These revolts succeed with probability* $\beta(1, M, \gamma)$*.*

The expected policy payoff for a majority citizen is

$$\begin{cases} 1 - q^2(1-p) - \mu & ;\beta(1, M, \gamma) > 1 - \delta_0 \\ (1-q)^2 + (2q(1-q) + pq^2)\,\beta(1, M, \gamma) - \mu & ;\beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Corollary 1 of the main text is then modified as follows:

Corollary 2. *The value of institutional constraints is:*

$$\begin{cases} (1-p)(q-q^2) - \mu & ;\beta(1, M, \gamma) > 1 - \delta_0 \\ ((2-p)\beta(1, M, \gamma) - 1)\,(q-q^2) - \mu & ;\beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Proposition 4 of the main text is modified as follows.

Proposition 7. *There is threshold* $p^*(M, \gamma, q, \mu)$ *such that a majority citizen's policy payoff is higher without institutional constraints if and only if the scope of the divine law* $p > p^*$*, where*

$$p^*(M, \gamma, q, \mu) = \begin{cases} 1 - \frac{\mu}{q(1-q)} & ;\beta(1, M, \gamma) > 1 - \delta_0 \\ 2 - \frac{1}{\beta(1,M,\gamma)}\left(1 + \frac{\mu}{q(1-q)}\right) & ;\beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

Moreover,

1. $p^*(M, \gamma, q, \mu)$ *is increasing in* M *and* γ*; strictly so if and only if* $\beta(1, M, \gamma) < 1 - \delta_0$*.*
2. *For* $\mu > 0$*,* $p^*(M, \gamma, q, \mu)$ *has an inverted U-shape in* q*, with*

$$lim_{q\to 0^+} p^*(M, \gamma, q, \mu) = lim_{q\to 1^-} p^*(M, \gamma, q, \mu) = -\infty.$$

As in the model in the main text, a higher scope of the law makes it less likely for a society to adopt institutional constraints. The difference is regarding the comparative statics with respect to M and γ. In this model, a more homogeneous society and a society with higher solidarity is *more* likely to adopt institutional constraints.

Why are the comparative statics going in the opposite direction? In the model with $y(a_1, a_2) = \min\{a_1, a_2\}$, the main advantage of institutional constraints is that a good ruler can *block* the bad ruler: he can just impose $a_j = 0$ on the aggregate action. Therefore, when institutional constraints are imposed, society needs to resort to revolt *less* than it would without institutional constraints. However, in the model with $y(a_1, a_2) = \max\{a_1, a_2\}$, the bad ruler *cannot* block the good ruler: even when the good ruler takes $a_j = 0$, the aggregate action is dictated by the other ruler's choice. In this model, the main advantage of institutional constraints is that the good ruler can *inform* the citizens by taking a different action than the bad ruler. The citizens can learn the state better with institutional constraints, yet, it still needs to revolt against an incongruent policy. In this model, therefore, when institutional constraints are imposed, the society resorts to revolt *more* than it would without institutional constraints. Because higher M and higher γ facilitate revolt, they favor the adoption of institutional constraints.

Proposition 5 of the main text is modified as follows.

Proposition 8. *There is a cost threshold such that the majority citizen's policy payoff is higher without institutional constraints if and only if $\mu > \mu^*$, where*

$$\mu^*(\beta, p, q) = \begin{cases} (1-p)(q-q^2) & ;\beta > 1-\delta_0 \\ \\ ((2-p)\beta - 1)\,(q-q^2) & ;\beta < 1-\delta_0, \end{cases}$$

where $\beta = \beta(1, M, \gamma)$. Moreover,

1. *μ^* is strictly decreasing in p. μ^* is weakly increasing in $\beta(1, M, \gamma)$ (and hence in M and γ), strictly so when $\beta < 1 - \delta_0$.*
2. *Suppose $\delta_0 < T/M$, so that there is sufficient conflict of interest that the threat of revolt does not deter the minority-congruent ruler ($\beta < 1 - \delta_0$). Then,*

$$\frac{\partial^2 \mu^*(\beta, p, q)}{\partial p \partial \beta} = -(q - q^2) < 0.$$

As in the model in the main text, higher scope of the law p improves the majority's ability to control the ruler, thereby reducing the marginal value of institutional constraints, and hence the cost threshold below which they are adopted. Recall that societal homogeneity M or solidarity γ improve the majority's ability to revolt. Contrary to the model in the main text, in this model, institutional constraints provide information about incongruent policies, leading the majority towards revolting more. Therefore, societal homogeneity and solidarity increase the marginal value of institutional constraints, and hence they increase the cost threshold below which they are adopted. Indeed, if M and γ are low enough so that

$\beta < 1 - \delta_0$ and $(2-p)\beta < 1$, it follows that $\mu^*(\beta, p, q) < 0$, and institutional constraints are never adopted. That is, in societies where homogeneity and solidarity are extremely low, it is never worth adopting institutional constraints. Intuitively, in this model, institutional constraints provide information to citizens and citizens use this information to revolt against incongruent policies. When the threat of revolt does not discipline the ruler and it is not likely to overturn incongruent policies, such information has no value, and it is not worth bringing in a second ruler for the sole purpose of providing information.

Note, however, that even though the comparative statics with respect to M and γ change, the second part of Proposition 5 remains intact. Recall that μ^* is decreasing in p, and it decreases faster when β is higher. Therefore, this model maintains the idea that homogeneity M and solidarity γ complements the scope of the law p. Intuitively, higher scope of the law is useful insofar as it is accompanied by a revolt. On the other hand, μ^* is increasing in β, and it increases slower when p is higher. Therefore, in this model, the scope of the law p substitutes homogeneity M and solidarity γ. Intuitively, the information provided through institutional constraints is more useful when revolt capabilities are higher. Yet, a higher scope of the law renders this information (and therefore the revolt capability) less useful by providing information regardless of institutions.

Regarding the inertia of institutional constraints, Proposition 6 of the main text is modified as follows. As in Proposition 6 in the main text, we focus on the case where institutional constraints may be adopted or not. This means restricting attention to the $(2-p)\beta > 1$ case; otherwise, institutional constraints are never adopted.

Proposition 9. *Suppose $\gamma \sim U[0,1]$. Let $Q = Pr_\gamma(\mu \leq \mu^*(\gamma))$ be the probability that institutional constraints improve the majority citizen's policy payoff. Suppose $\delta_0 < T/M$, so that there is sufficient conflict of interest that the threat of revolt does not deter the minority-congruent ($\beta < 1 - \delta_0$). Moreover, suppose $(2-p)(1-\frac{T}{M}) > 1$, so that the institutional constraints are sometimes adopted ($(2-p)\beta > 1$ for high enough γ). Then,*

$$Q(\mu'; M, p) = \begin{cases} 1 - \frac{1+\mu'}{(2-p)(1-T/M)} & ;\mu' \leq (2-p)(1-T/M) - 1 \\ 0 & ;\mu' > (2-p)(1-T/M) - 1, \end{cases}$$

where $\mu' = \mu/(q-q^2)$. Moreover,

1. *Q is decreasing in p and increasing in M; strictly so when $\mu' \leq (2-p)(1-T/M) - 1$.*
2. *$|Q(\mu'_2) - Q(\mu'_1)|$ is strictly increasing in p and strictly decreasing in M for all $\mu'_2 > \mu'_1$, with $\mu'_2 \leq (2-p)(1-T/M) - 1$.*

Proof. Using Proposition 8,

$$\begin{aligned} Q &= Pr_\gamma(\mu \leq \mu^*(\gamma) \mid \beta < 1 - \delta_0) \\ &= Pr_\gamma\left(\mu \leq \left((2-p)\beta - 1\right)(q - q^2)\right) \end{aligned}$$

Using the fact that $\beta = \beta(1, M, \gamma)$, and substituting Proposition 1 in the main text, we have:

$\beta = H\left((1-\frac{T}{M})\gamma\right)$. Because $H = U[0,1]$, $\beta = (1-\frac{T}{M})\gamma$. Substituting, we have:

$$\begin{aligned} Q &= Pr_\gamma\left(\mu \leq \left((2-p)(1-\frac{T}{M})\gamma - 1\right)(q-q^2)\right) \\ &= Pr_\gamma\left(\mu' \leq \left((2-p)(1-\frac{T}{M})\gamma - 1\right)\right) \\ &= Pr_\gamma\left((2-p)(1-\frac{T}{M})\gamma \geq 1+\mu'\right) \\ &= Pr_\gamma\left(\gamma \geq \frac{1+\mu'}{(2-p)(1-\frac{T}{M})}\right) \end{aligned}$$

Recall that $\gamma \sim U[0,1]$. Under the restriction $(2-p)(1-\frac{T}{M}) > 1$, $\frac{1}{(2-p)(1-\frac{T}{M})} < 1$, which means Q is strictly positive for $\mu' = 0$. Moreover, as long as $\mu' \leq (2-p)(1-T/M) - 1$, $\frac{1+\mu'}{(2-p)(1-\frac{T}{M})} \leq 1$, which means Q is positive. Indeed, when $\mu' \leq (2-p)(1-T/M) - 1$,

$$Q = Pr_\gamma\left(\gamma \geq \frac{1+\mu'}{(2-p)(1-\frac{T}{M})}\right) = 1 - \frac{1+\mu'}{(2-p)(1-\frac{T}{M})}$$

On the other hand, when $\mu' > (2-p)(1-T/M) - 1$, $\frac{1+\mu'}{(2-p)(1-\frac{T}{M})} > 1$, which means $Q = 0$.

The first part of Proposition 9 is evident from these formulas. Regarding the second part, as $Q(\mu')$ is decreasing in μ', with $\mu'_1 < \mu'_2$:

$$|Q(\mu'_2) - Q(\mu'_1)| = Q(\mu'_1) - Q(\mu'_2)$$

Moreover, since $\mu'_1 < \mu'_2 \leq (2-p)(1-T/M) - 1$, $Q(\mu'_1) = 1 - \frac{1+\mu'_1}{(2-p)(1-\frac{T}{M})}$ and $Q(\mu'_2) = 1 - \frac{1+\mu'_2}{(2-p)(1-\frac{T}{M})}$. Therefore,

$$\begin{aligned} Q(\mu'_1) - Q(\mu'_2) &= \left(1 - \frac{1+\mu'_1}{(2-p)(1-\frac{T}{M})}\right) - \left(1 - \frac{1+\mu'_2}{(2-p)(1-\frac{T}{M})}\right) \\ &= \frac{\mu'_2 - \mu'_1}{(2-p)(1-\frac{T}{M})} \end{aligned}$$

which is strictly increasing in p and strictly decreasing in M. □

Proposition 9 provides new insights into the effect of changes in the costs of institutions. For a given μ', societies with sufficiently high solidarity levels adopt institutional constraints. Consider a reduction in the costs of institutional constraints from μ'_2 to μ'_1, e.g., due to peacetime. Then, societies with even lower levels of γ tend to adopt institutional constraints. As part 2 of the Proposition shows, this change tends to be larger when p is larger. This is because the scope of the law substitutes solidarity in this model: when the scope of the law p is larger, the capacity of revolt obtained through γ needs to change a lot for a society to adopt institutional constraints. Therefore, the cutoff of solidarity above which institutional

constraints are adopted varies strongly with μ'. Consequently, societies with high scope of law are more responsive to a decrease in μ'.

We conclude our discussion by presenting the analogue of Proposition in the main text. Given that institutional constraints make revolt more likely by providing information, the following result is not surprising.

Proposition 10. *Suppose that $p^* \in (0,1)$ and that $\delta_0 < T/M$, so that there is sufficient conflict of interest and the threat of revolt does not deter the minority-congruent ruler ($\beta < 1 - \delta_0$). Focusing on the scope of the law p as the only source of variation, the equilibrium probabilities of revolt attempts and successful revolt are both higher in societies with institutional constraints. Formally,*

$$\mathbb{E}[\frac{pq}{2} \mid p > p^*] < \mathbb{E}[q(1-q)+\frac{pq^2}{2} \mid p < p^*] \quad \textit{and} \quad \mathbb{E}[\frac{pq\beta}{2} \mid p > p^*] < \mathbb{E}[\left(2q(1-q) + \frac{pq^2}{2}\right)\beta \mid p < p^*],$$

for a given q and $\beta = \beta(1, M, \gamma)$.

C An Extended Model of Institutional Constraints

In this section, we present an extended version of the model with institutional constraints in Section 2.1 of the main text and provide a characterization of the equilibrium. The extended model is different from our main model in two ways. First, we do not require that the rulers observe each others' types. Second, we allow for $\delta_1 > 0$, but we still maintain the assumption that $\delta_1 < \delta_0 < 1$. That is, throughout this section, we will maintain the following assumption.

Assumption 2. $\delta_1 < \delta_0 < 1$, *i.e., the minority-congruent ruler always prefers to propose* $a = 1$, *and his incentives to propose* $a = 0$ *are stronger in state* $s = 0$.

Note that under Assumption 2, the PBE with one ruler discussed in the main text (Proposition 2) applies verbatim. This is because $\delta_1 < \delta_0$ ensures $\sigma(\emptyset, 1) \geq \sigma(\emptyset, 0)$ in any PBE. Then, Bayesian updating implies that following $\hat{s} = \emptyset$ and $a = 0$, the belief that the ruler's action does not match the state satisfies $q'(a) \leq \frac{1}{2}$. As a result, there are no revolts following $\hat{s} = \emptyset$. Throughout the reminder of this section, we analyze the game with two rulers.

Timing The timing of the game is as follows.

1. The nature determines the realizations of rulers' types, the state of the world s, signal $\hat{s}$, the common value of costs $\bar{c}$, and idiosyncratic elements of costs ϵ_i's.
2. Each ruler observes his own type, the state s, and $\hat{s}$. Each citizen i observes $\hat{s}$ and her private cost c_i.
3. Ruler 1 proposes action a_1, which ruler 2 and the citizens observe.
4. Ruler 2 proposes action a_2, which the citizens observe.
5. The aggregate policy is $A = \min\{a_1, a_2\}$. Citizens simultaneously decide whether or not to revolt against the aggregate policy A.
6. Success of revolution r is determined, payoffs are received, and the game ends.

We consider the Perfect Bayesian Nash Equilibrium of this game. The existence of two rulers who do not observe each others' types can generate multiple equilibria, in which case we use *forward induction* criterion of Govindan and Wilson (2009) to select an outcome. This criterion implies the Intuitive Criterion of Cho and Kreps (1987) for simple signaling games with one sender. The formal definition of forward induction criterion is provided below in Section C.2.

C.1 Formal Definition of Equilibrium

The majority-congruent ruler $j \in \{1, 2\}$ (i.e., ruler j of type $t_j = g$) always chooses $a_j = s$ by assumption. The strategy of the minority-congruent ruler 1 (i.e., ruler 1 of type $t_1 = b$) in state s when public signal is $\hat{s}$ is:

$$\sigma_1(\widehat{s}, s) \equiv \Pr(a_1 = 1 | s, \hat{s}) \quad \in [0, 1]$$

The strategy of minority-congruent ruler 2 (i.e., ruler 2 of type $t_2 = b$) in state s, given public signal $\hat{s}$ and ruler 1's action a_1 is:

$$\sigma_2(\widehat{s}, s, a_1) \equiv \Pr(a_2 = 1 | s, \hat{s}, a_1) \quad \in [0,1]$$

The posterior beliefs of citizens that the aggregate policy is incongruent, given information $(\hat{s}, a_1, a_2)$, is denoted by:

$$q(\hat{s}, a_1, a_2) \equiv \Pr\left(\min\{a_1, a_2\} \neq s | \hat{s}, a_1, a_2\right) \quad \in [0,1]$$

Let $r_i \in \{0,1\}$ denote the revolting decision of citizen i, with $r_i = 1$ corresponding to revolting. The strategy of a majority citizen i when posterior beliefs are q' and the cost of revolt is c_i is denoted by:

$$\varphi(q', c_i) \equiv \Pr(r_i = 1 | q', c_i) \quad \in [0,1]$$

As we will see later, in this version of the model, the minority citizens sometimes participate in revolt against $A = 0$ when they believe a sufficient number of majority citizens participate as well. The strategy of a minority citizen i when the aggregate action is $A = 0$, the posterior beliefs are q', and the cost of revolt is c_i, is denoted by:

$$\phi(q', c_i) \equiv \Pr(r_i = 1 | q', c_i) \quad \in [0,1]$$

The Perfect Bayesian Nash Equilibrium of the game is a tuple $(\sigma_1^*, \sigma_2^*, \varphi^*, \phi^*, q^*)$ such that the following are satisfied.

1. $\varphi^*(q', c_i)$ maximizes the payoff of the citizens in majority for any $q' = q^*(\hat{s}, a_1, a_2)$.
2. $\phi^*(q', c_i)$ maximizes the payoff of the citizens in minority for any $q' = q^*(\hat{s}, a_1, a_2)$ when $A = 0$.
3. $q^*(\hat{s}, a_1, a_2)$ is given by Bayes' Rule.
4. Given φ^*, ϕ^* and σ_2^*, σ_1^* maximizes the payoff of the minority-congruent ruler 1. Similarly, given φ^*, ϕ^* and σ_1^*, σ_2^* maximizes the payoff of the minority-congruent ruler 2.

C.2 Forward Induction

The following definitions are adapted from Govindan and Wilson (2009).

A **terminal node** of the game with two rulers is:

$$(s, \hat{s}, a_1, a_2, r) \qquad \in \{(0,0),(1,1),(0,\emptyset),(1,\emptyset)\} \times \{0,1\} \times \{0,1\} \times \{0,1\}$$

As we will demonstrate later, any subgame following $\hat{s} = s$ has a unique PBE. While refining the equilibrium, we will focus on the subgame following $\hat{s} = \emptyset$.

We begin by a formal definition of an outcome.

Definition 1. *The* **outcome** *of a Perfect Bayesian Nash Equilibrium is the induced probability distribution over the terminal nodes.*

Definition 2. *Consider an outcome. A pure strategy of a player is* **relevant** *for that outcome if:*

1. *There is a Perfect Bayesian Nash Equilibrium with that outcome, and,*
2. *The pure strategy is optimal under the beliefs in the said equilibrium.*

In words, given an outcome, the relevant strategies are reasonable deviations that a player may consider.

Definition 3. *Consider an outcome. An information set is* **relevant** *for that outcome if it is reached with strictly positive probability by some relevant strategy for that outcome.*

In words, relevant information sets are those that can be reached via reasonable deviations by some players.

Definition 4. *An outcome satisfies* **forward induction** *if it results from a Perfect Bayesian Nash Equilibrium in which at every information set that is relevant for that outcome the support of the beliefs are confined to profiles of Nature's strategies and other players' strategies that are relevant for that outcome.*

In words, an outcome satisfies forward induction if, in any relevant information set, the players believe that information set is reached via a reasonable deviation.

C.3 Equilibrium Characterization

C.3.1 Citizens' Actions

Suppose $A = 1$. In this case, the minority citizens never participate in the revolt, and the measure of citizens who may contemplate a revolt is M. As discussed in Proposition 1 of the main text, in a symmetric cutoff strategy equilibrium as $\rho \to 0$, a successful revolution occurs with probability:

$$\beta(q', M, \gamma) = H\left((1 - \frac{T}{M}) \cdot \gamma \cdot (2q' - 1)\right)$$

In contrast, when $A = 0$, the minority citizens always prefer to participate in the revolt, if they believe a sufficient number of majority citizens also revolt. In this case, the measure of citizens who may contemplate a revolt is 1. In any equilibrium, let the probability of a successful revolt be given by:

$$\bar{\beta}(q') \qquad \in [0, 1]$$

For our purposes, a closed-form equation characterizing $\bar{\beta}(q')$ is unnecessary. This is because we will show that in PBE that survives forward induction, there are no revolts against

$A = 0$.[2] However, we will maintain the assumption that $\bar{\beta}(q')$ is continuous in q', and:

$$\bar{\beta}(q') = 0 \qquad \text{for any } q' \leq \frac{1}{2} \tag{12}$$

Equation (12) holds because, in any equilibrium, majority citizens do not participate in a revolt against $A = 0$ when $q' \leq \frac{1}{2}$. Foreseeing this, minority members do not participate either, and hence there are no revolts.

C.3.2 Beliefs Following Proposed Policy

When the issue is predordained ($\hat{s} \in \{0, 1\}$), $q^*(\hat{s}, a_1, a_2) = |\hat{s} - \min\{a_1, a_2\}| \in \{0, 1\}$.

When the issue is non-preordained ($\hat{s} = \emptyset$), the posterior beliefs are given by:

$$\begin{aligned}
q^*(\emptyset,0,0) &\equiv \Pr(\min\{a_1,a_2\} \neq s | a_1 = a_2 = 0, \hat{s} = \emptyset) \\
&= \Pr(s = 1 | a_1 = a_2 = 0, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 1, a_1 = a_2 = 0, \hat{s} = \emptyset)}{\Pr(s = 1, a_1 = a_2 = 0, \hat{s} = \emptyset) + \Pr(s = 0, a_1 = a_2 = 0, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2}q^2(1-\sigma_1^*(\emptyset,1))(1-\sigma_2^*(\emptyset,1,0))}{\frac{1}{2}q^2(1-\sigma_1^*(\emptyset,1))(1-\sigma_2^*(\emptyset,1,0)) + \frac{1}{2}\left(q^2(1-\sigma_1^*(\emptyset,0))(1-\sigma_2^*(\emptyset,0,0)) + q(1-q)(1-\sigma_1^*(\emptyset,0)) + (1-q)q(1-\sigma_2^*(\emptyset,0,0)) + (1-q)^2\right)} \\
&= \frac{q^2(1-\sigma_1^*(\emptyset,1))(1-\sigma_2^*(\emptyset,1,0))}{q^2(1-\sigma_1^*(\emptyset,1))(1-\sigma_2^*(\emptyset,1,0)) + \left(q^2(1-\sigma_1^*(\emptyset,0))(1-\sigma_2^*(\emptyset,0,0)) + q(1-q)(1-\sigma_1^*(\emptyset,0)) + (1-q)q(1-\sigma_2^*(\emptyset,0,0)) + (1-q)^2\right)}
\end{aligned}$$

$$\begin{aligned}
q^*(\emptyset,0,1) &\equiv \Pr(\min\{a_1,a_2\} \neq s | a_1 = 0, a_2 = 1, \hat{s} = \emptyset) \\
&= \Pr(s = 1 | a_1 = 0, a_2 = 1, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 1, a_1 = 0, a_2 = 1, \hat{s} = \emptyset)}{\Pr(s = 1, a_1 = 0, a_2 = 1, \hat{s} = \emptyset) + \Pr(s = 0, a_1 = 0, a_2 = 1, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2}q^2(1-\sigma_1^*(\emptyset,1))\sigma_2^*(\emptyset,1,0) + q(1-q)(1-\sigma_1^*(\emptyset,1))}{\frac{1}{2}q^2(1-\sigma_1^*(\emptyset,1))\sigma_2^*(\emptyset,1,0) + q(1-q)(1-\sigma_1^*(\emptyset,1)) + \frac{1}{2}\left(q^2(1-\sigma_1^*(\emptyset,0))\sigma_2^*(\emptyset,0,0) + (1-q)q\sigma_2^*(\emptyset,0,0)\right)}
\end{aligned}$$

$$\begin{aligned}
q^*(\emptyset,1,0) &\equiv \Pr(\min\{a_1,a_2\} \neq s | a_1 = 1, a_2 = 0, \hat{s} = \emptyset) \\
&= \Pr(s = 1 | a_1 = 1, a_2 = 0, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 1, a_1 = 1, a_2 = 0, \hat{s} = \emptyset)}{\Pr(s = 1, a_1 = 1, a_2 = 0, \hat{s} = \emptyset) + \Pr(s = 0, a_1 = 1, a_2 = 0, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2}q^2\sigma_1^*(\emptyset,1)(1-\sigma_2^*(\emptyset,1,1)) + (1-q)q(1-\sigma_2^*(\emptyset,1,1))}{\frac{1}{2}q^2\sigma_1^*(\emptyset,1)(1-\sigma_2^*(\emptyset,1,1)) + q(1-q)(1-\sigma_2^*(\emptyset,1,1)) + \frac{1}{2}\left(q^2\sigma_1^*(\emptyset,0)(1-\sigma_2^*(\emptyset,0,1)) + q(1-q)\sigma_1^*(\emptyset,0)\right)}
\end{aligned}$$

[2] It is, however, possible to characterize the value of $\bar{\beta}(q')$. In any symmetric cutoff strategy equilibrium as $\rho \to 0$, the minority citizens use a cutoff c^m such that $r_i = 1$ if and only if $c_i \leq c^m$. Similarly, majority citizens use a cutoff c^M such that $r_i = 1$ if and only if $c_i \leq c^M$. The revolution is successful as long as $\bar{c} \leq \bar{c}^*$ for some c^*. The three cutoff values, c^m, c^M and $\bar{c}^*$ satisfy:

$$\begin{aligned}
\gamma \cdot \Pr\left(\bar{c} \leq \bar{c}^* | c_i = c^m\right) &= c^m \\
\gamma \cdot (2q' - 1) \cdot \Pr\left(\bar{c} \leq \bar{c}^* | c_i = c^M\right) &= c^M \\
(1 - M) \cdot \Pr(c_i \leq c^m | \bar{c} = \bar{c}^*) + M \cdot \Pr(c_i \leq c^M | \bar{c} = \bar{c}^*) &= T
\end{aligned}$$

The probability of a successful revolt is $\bar{\beta}(q') = H(\bar{c}^*)$.

$$
\begin{aligned}
q^*(\emptyset,1,1) &\equiv \Pr(\min\{a_1,a_2\} \neq s | a_1 = a_2 = 1, \hat{s} = \emptyset) \\
&= \Pr(s = 0 | a_1 = a_2 = 1, \hat{s} = \emptyset) \\
&= \frac{\Pr(s = 0, a_1 = a_2 = 1, \hat{s} = \emptyset)}{\Pr(s = 0, a_1 = a_2 = 1, \hat{s} = \emptyset) + \Pr(s = 1, a_1 = a_2 = 1, \hat{s} = \emptyset)} \\
&= \frac{\frac{1}{2} q^2 \sigma_1^*(\emptyset,0) \sigma_2^*(\emptyset,0,1)}{\frac{1}{2} q^2 \sigma_1^*(\emptyset,0) \sigma_2^*(\emptyset,0,1) + \frac{1}{2}\left(q^2 \sigma_1^*(\emptyset,1) \sigma_2^*(\emptyset,1,1) + q(1-q)\sigma_1^*(\emptyset,1) + (1-q)q\sigma_2^*(\emptyset,1,1) + (1-q)^2\right)}
\end{aligned}
$$

C.3.3 Rulers' Actions

When the Issue is Preordained We proceed in the fashion of backward induction, first pinning down the strategies of minority-congruent ruler 2 at every history.

1. Consider the case $\hat{s} = s = a_1 = 0$. In this case, $\min\{a_1, a_2\} = 0$ regardless of a_2, and $q^*(0,0,a_2) = 0$ for any $a_2 \in \{0,1\}$. We conclude that any $\sigma_2^*(0,0,0) \in [0,1]$ can be a part of a PBE.

   Note that because $q^*(0,0,a_2) = 0$, the majority citizens never participate in revolt, and consequently, there are no revolts. Therefore, the payoff of minority-congruent ruler 2 is δ_0 in any PBE.

2. Now, consider the case $\hat{s} = s = 0$ and $a_1 = 1$. In this case, $\min\{a_1, a_2\} = a_2$ and $q^*(0,1,a_2) = a_2$ for any $a_2 \in \{0,1\}$. When $a_2 = 0$, majority citizens do not participate in the revolt and there are no revolts. When $a_2 = 1$, only majority citizens participate in the revolt, which is successful with probability $\beta(1, M, \gamma)$. Thus, ruler 2's optimal strategy when $(\hat{s}, s, a_1) = (0,0,1)$ is:

$$
\sigma_2^*(0,0,1) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (1 - \beta(1, M, \gamma)) + (1 - \sigma) \cdot \delta_0
$$

   Therefore, ruler 2's PBE strategy is:

$$
\sigma_2^*(0,0,1) = \begin{cases} 0 & ; \delta_0 > 1 - \beta(1, M, \gamma) \\ 1 & ; \delta_0 < 1 - \beta(1, M, \gamma) \end{cases}
$$

3. Now, consider the case $\hat{s} = s = 1$ and $a_1 = 0$. In this case, $\min\{a_1, a_2\} = 0$ regardless of a_2, and $q^*(1,0,a_2) = 1$ for any $a_2 \in \{0,1\}$. We conclude that any $\sigma_2^*(1,1,0) \in [0,1]$ can be a part of a PBE.

   Note that because $q^*(1,0,a_2) = 1$, majority citizens participate in a revolt against $A = 0$. Foreseeing this, minority citizens also participate. Therefore, all citizens contemplate participating in a revolt, which is successful with probability $\bar{\beta}(1)$. The payoff of minority-congruent ruler 2 is $\delta_0 \cdot \left(1 - \bar{\beta}(1)\right)$ in any PBE.

4. Finally, consider the case $\hat{s} = s = a_1 = 1$. In this case, $\min\{a_1, a_2\} = a_2$ and $q^*(1,1,a_2) = 1 - a_2$ for any $a_2 \in \{0,1\}$. When $a_2 = 0$, all citizens contemplate participating in a revolt, which is successful with probability $\bar{\beta}(1)$. When $a_2 = 1$, none of the citizens revolt. Thus, ruler 2's optimal strategy when $(\hat{s}, s, a_1) = (1,1,1)$ is:

$$
\sigma_2^*(1,1,1) \in \arg\max_{\sigma \in [0,1]} \sigma + (1 - \sigma) \cdot \delta_0 \cdot \left(1 - \bar{\beta}(1)\right)
$$

   Given Assumption 2, we conclude that $\sigma_2^*(1,1,1) = 1$.

Next, we pin down the strategy of minority-congruent ruler 1 in every history.

1. Consider the case $\hat{s} = s = 0$. If ruler 1 chooses a_1, the probability that ruler 2 chooses $a_2 = 0$ is:

$$(1 - q) + q \cdot (1 - \sigma_2^*(0, 0, a_1))$$

and the probability that ruler 2 chooses $a_2 = 1$ is:

$$q \cdot \sigma_2^*(0, 0, a_1)$$

Therefore, ruler 1's optimal strategy when $\hat{s} = s = 0$ is:

$$\sigma_1^*(0,0) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot (((1-q) + q \cdot (1 - \sigma_2^*(0,0,1))) \cdot \delta_0 + q \cdot \sigma_2^*(0,0,1) \cdot (1 - \beta(1, M, \gamma))) + (1 - \sigma) \cdot \delta_0$$

- If $1 - \beta(1, M, \gamma) > \delta_0$, $\sigma_2^*(0, 0, 1) = 1$ and thus ruler 1's optimal strategy is:

$$\sigma_1^*(0,0) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot ((1-q) \cdot \delta_0 + q \cdot (1 - \beta(1, M, \gamma))) + (1 - \sigma) \cdot \delta_0$$

  which is maximized when $\sigma_1^*(0, 0) = 1$.

- If $1 - \beta(1, M, \gamma) < \delta_0$, $\sigma_2^*(0, 0, 1) = 0$ and $\min\{a_1, a_2\} = 0$ for any $a_1 \in \{0, 1\}$ in any PBE. We conclude that any $\sigma_1^*(0, 0) \in [0, 1]$ can be a part of a PBE.

  Note that majority citizens never participate in a revolt, and there are no revolts. Therefore, the payoff of minority-congruent ruler 1 is δ_0 in any PBE.

2. Now, consider the case $\hat{s} = s = 1$. If ruler 1 chooses a_1, the probability that ruler 2 chooses $a_2 = 0$ is:

$$q \cdot (1 - \sigma_2^*(1, 1, a_1))$$

and the probability that ruler 2 chooses $a_2 = 1$ is:

$$(1 - q) + q \cdot \sigma_2^*(1, 1, a_1)$$

Thus, ruler 1's optimal strategy when $\hat{s} = s = 1$ is:

$$\sigma_1^*(1,1) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot \big(q \cdot (1 - \sigma_2^*(1,1,1)) \cdot (1 - \bar{\beta}(1)) \cdot \delta_1 + (1-q) + q \cdot \sigma_2^*(1,1,1)\big) + (1 - \sigma) \cdot (1 - \bar{\beta}(1)) \cdot \delta_1$$

But recall that $\sigma_2^*(1, 1, 1) = 1$. Thus, ruler 1's choice simplifies to:

$$\sigma_1^*(1,1) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot 1 + (1 - \sigma) \cdot \delta_1 \cdot (1 - \bar{\beta}(1))$$

Given Assumption 2, we conclude that $\sigma_1^*(1, 1) = 1$.

Note that $\sigma_1^*(1,1) \cdot \sigma_2^*(1,1,1) = 1$ in any PBE. That is, when $\hat{s} = s = 1$, the aggregate policy is $A = 1$ with probability one and there are no revolts.

If $1 - \beta(1, M, \gamma) > \delta_0$, $\sigma_1^*(0,0) \cdot \sigma_2^*(0,0,1) = 1$. That is, when $\hat{s} = s = 0$, the aggregate policy taken by two minority-congruent rulers is $A = 1$ with probability one. This is followed with a revolt with probability $\beta(1, M, \gamma)$.

If $1 - \beta(1, M, \gamma) < \delta_0$, $\sigma_1^*(0,0) \cdot \sigma_2^*(0,0,1) = 0$. That is, when $\hat{s} = s = 0$, the aggregate policy is $A = 0$ with probability one and there are no revolts.

When the Issue is Non-Preordained The equilibrium strategy of minority-congruent ruler 2 in any history is characterized by the following equations.

$$\sigma_2^*(\emptyset,0,0) \in \arg\max_{\sigma\in[0,1]} \sigma \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset,0,1))) + (1-\sigma) \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset,0,0))) \tag{13}$$

$$\sigma_2^*(\emptyset,0,1) \in \arg\max_{\sigma\in[0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset,1,1), M, \gamma)) + (1-\sigma) \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset,1,0))) \tag{14}$$

$$\sigma_2^*(\emptyset,1,0) \in \arg\max_{\sigma\in[0,1]} \sigma \cdot \delta_1 \cdot (1 - \bar{\beta}(q^*(\emptyset,0,1))) + (1-\sigma) \cdot \delta_1 \cdot (1 - \bar{\beta}(q^*(\emptyset,0,0))) \tag{15}$$

$$\sigma_2^*(\emptyset,1,1) \in \arg\max_{\sigma\in[0,1]} \sigma \cdot (1 - \beta(q^*(\emptyset,1,1), M, \gamma)) + (1-\sigma) \cdot \delta_1 \cdot (1 - \bar{\beta}(q^*(\emptyset,1,0))) \tag{16}$$

For the equilibrium strategy of minority-congruent ruler 1, consider two possible histories.

1. Consider the case when $\hat{s} = \emptyset$ and $s = 0$. If ruler 1 chooses a_1, the probability that ruler 2 chooses $a_2 = 0$ is:

   $$(1-q) + q \cdot (1 - \sigma_2^*(\emptyset, 0, a_1))$$

   and the probability that ruler 2 chooses $a_2 = 1$ is:

   $$q \cdot \sigma_2^*(\emptyset, 0, a_1)$$

   Therefore, minority-congruent ruler 1's policy when $(\hat{s}, s) = (\emptyset, 0)$ is:

   $$\begin{aligned}\sigma_1^*(\emptyset,0) \in \arg\max_{\sigma\in[0,1]} & \sigma \cdot ((1-q) + q \cdot (1 - \sigma_2^*(\emptyset,0,1))) \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset,1,0))) \\ & + \sigma \cdot q \cdot \sigma_2^*(\emptyset,0,1) \cdot (1 - \beta(q^*(\emptyset,1,1), M, \gamma)) \\ & + (1-\sigma) \cdot ((1-q) + q \cdot (1 - \sigma_2^*(\emptyset,0,0))) \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset,0,0))) \\ & + (1-\sigma) \cdot q \cdot \sigma_2^*(\emptyset,0,0) \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset,0,1)))\end{aligned} \tag{17}$$

2. Consider the case when $\hat{s} = \emptyset$ and $s = 1$. If ruler 1 chooses a_1, the probability that ruler 2 chooses $a_2 = 0$ is:

   $$q \cdot (1 - \sigma_2^*(\emptyset, 1, a_1))$$

   and the probability that ruler 2 chooses $a_2 = 1$ is:

   $$(1-q) + q \cdot \sigma_2^*(\emptyset, 1, a_1)$$

Therefore, minority-congruent ruler 1's optimal strategy when $(\hat{s}, s) = (\emptyset, 1)$ is:

$$\begin{aligned}\sigma_1^*(\emptyset,1) \in \arg\max_{\sigma\in[0,1]} &\sigma \cdot q \cdot (1-\sigma_2^*(\emptyset,1,1)) \cdot \delta_1 \cdot (1-\bar{\beta}(q^*(\emptyset,1,0))) \\ &+ \sigma \cdot ((1-q) + q \cdot \sigma_2^*(\emptyset,1,1)) \cdot (1-\beta(q^*(\emptyset,1,1),M,\gamma)) \\ &+ (1-\sigma) \cdot q \cdot (1-\sigma_2^*(\emptyset,1,0)) \cdot \delta_1 \cdot (1-\bar{\beta}(q^*(\emptyset,0,0))) \\ &+ (1-\sigma) \cdot ((1-q) + q \cdot \sigma_2^*(\emptyset,1,0)) \cdot \delta_1 \cdot (1-\bar{\beta}(q^*(\emptyset,0,1)))\end{aligned}$$

The analysis proceeds in a number of claims.

Claim 1. *In any PBE of the game with two rulers,* $\sigma_2^*(\emptyset,1,1) = 1$.

Proof. Suppose, towards a contradiction, that $\sigma_2^*(\emptyset,1,1) < 1$. By Equation (16), this implies:

$$1-\beta(q^*(\emptyset,1,1),M,\gamma) \leq \delta_1 \cdot (1-\bar{\beta}(q^*(\emptyset,1,0)))$$

By Assumption 2, then,

$$1-\beta(q^*(\emptyset,1,1),M,\gamma) < \delta_0 \cdot (1-\bar{\beta}(q^*(\emptyset,1,0)))$$

which, by Equation (14), implies: $\sigma_2^*(\emptyset,0,1) = 0$.

By equations in Section C.3.2, this implies: $q^*(\emptyset,1,1) = 0$. But then, $\beta(q^*(\emptyset,1,1),M,\gamma) = 0$. By Assumption 2, then:

$$1-\beta(q^*(\emptyset,1,1),M,\gamma) > \delta_1 \cdot (1-\bar{\beta}(q^*(\emptyset,1,0)))$$

and therefore $\sigma_2^*(\emptyset,1,1) = 1$, a contradiction. □

Given Claim 1, the beliefs following $(\hat{s}, a_1, a_2) = (\emptyset, 1, 1)$ in any PBE is:

$$q^*(\emptyset,1,1) = \frac{q^2\sigma_1^*(\emptyset,0)\sigma_2^*(\emptyset,0,1)}{q^2\sigma_1^*(\emptyset,0)\sigma_2^*(\emptyset,0,1) + q^2\sigma_1^*(\emptyset,1) + q(1-q)\sigma_1^*(\emptyset,1) + (1-q)q + (1-q)^2} \tag{18}$$

and minority-congruent ruler 1's optimal strategy when $(\hat{s}, s) = (\emptyset, 1)$ is:

$$\begin{aligned}\sigma_1^*(\emptyset,1) \in \arg\max_{\sigma\in[0,1]} &\sigma \cdot (1-\beta(q^*(\emptyset,1,1),M,\gamma)) \\ &+ (1-\sigma) \cdot q \cdot (1-\sigma_2^*(\emptyset,1,0)) \cdot \delta_1 \cdot (1-\bar{\beta}(q^*(\emptyset,0,0))) \\ &+ (1-\sigma) \cdot ((1-q) + q \cdot \sigma_2^*(\emptyset,1,0)) \cdot \delta_1 \cdot (1-\bar{\beta}(q^*(\emptyset,0,1)))\end{aligned} \tag{19}$$

Claim 2. *In any PBE of the game with two rulers,* $\sigma_1^*(\emptyset,1) = 1$.

Proof. Suppose, towards a contradiction, that $\sigma_1^*(\emptyset,1) < 1$. Then, by Equation (19),

$$\begin{aligned}1-\beta(q^*(\emptyset,1,1),M,\gamma) \leq &\delta_1 \cdot q \cdot (1-\sigma_2^*(\emptyset,1,0)) \cdot (1-\bar{\beta}(q^*(\emptyset,0,0))) \\ &+ \delta_1 \cdot ((1-q) + q \cdot \sigma_2^*(\emptyset,1,0)) \cdot (1-\bar{\beta}(q^*(\emptyset,0,1)))\end{aligned} \tag{20}$$

Since the right hand-side of this inequality at most δ_1, and since $\delta_1 < 1$ by Assumption 2, we must have: $\beta(q^*(\emptyset, 1, 1), M, \gamma) > 0$. This means $q^*(\emptyset, 1, 1) > \frac{1}{2}$. By Equation (18), a necessary condition for this is:

$$\sigma_1^*(\emptyset, 0) \cdot \sigma_2^*(\emptyset, 0, 1) > \sigma_1^*(\emptyset, 1) \tag{21}$$

In particular, this requires $\sigma_1^*(\emptyset, 0) > 0$ and $\sigma_2^*(\emptyset, 0, 1) > 0$. We will investigate the implications of these observations separately.

- By Equation (17), $\sigma_1^*(\emptyset, 0) > 0$ implies:

$$\begin{aligned}&\delta_0 \cdot ((1-q) + q \cdot (1 - \sigma_2^*(\emptyset, 0, 1))) \cdot (1 - \bar{\beta}(q^*(\emptyset, 1, 0))) + q \cdot \sigma_2^*(\emptyset, 0, 1) \cdot (1 - \beta(q^*(\emptyset, 1, 1), M, \gamma)) \\ &\geq \delta_0 \cdot ((1-q) + q \cdot (1 - \sigma_2^*(\emptyset, 0, 0))) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 0))) + \delta_0 \cdot q \cdot \sigma_2^*(\emptyset, 0, 0) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1)))\end{aligned} \tag{22}$$

- By Equation (14), $\sigma_2^*(\emptyset, 0, 1) > 0$ implies:

$$1 - \beta(q^*(\emptyset, 1, 1), M, \gamma) \geq \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset, 1, 0))) \tag{23}$$

By (23), the left-hand side of Equation (22) is bounded above by $1 - \beta(q^*(\emptyset, 1, 1), M, \gamma)$. By (20), this is further bounded above by $\delta_1 \cdot q \cdot (1 - \sigma_2^*(\emptyset, 1, 0)) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 0))) + \delta_1 \cdot ((1-q) + q \cdot \sigma_2^*(\emptyset, 1, 0)) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1)))$. Therefore, the following inequality must hold:

$$\begin{aligned}&\delta_1 \cdot q \cdot (1 - \sigma_2^*(\emptyset, 1, 0)) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 0))) + \delta_1 \cdot ((1-q) + q \cdot \sigma_2^*(\emptyset, 1, 0)) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1))) \\ &\geq \delta_0 \cdot ((1-q) + q \cdot (1 - \sigma_2^*(\emptyset, 0, 0)) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 0))) + \delta_0 \cdot q \cdot \sigma_2^*(\emptyset, 0, 0) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1)))\end{aligned} \tag{24}$$

Recall, by Assumption 2, that $\delta_0 > \delta_1$. Therefore, inequality (24) cannot hold when $1 - \beta(q^*(\emptyset, 0, 0), 1, \gamma) = 1 - \beta(q^*(\emptyset, 0, 1), 1, \gamma)$. We conclude that $1 - \beta(q^*(\emptyset, 0, 0), 1, \gamma) \neq 1 - \beta(q^*(\emptyset, 0, 1), 1, \gamma)$. There are two mutually exhaustive possibilities.

- Suppose $1 - \bar{\beta}(q^*(\emptyset, 0, 0)) > 1 - \bar{\beta}(q^*(\emptyset, 0, 1))$. Then, by Equation (13), $\sigma_2^*(\emptyset, 0, 0) = 0$. Moreover, by Equation (15), $\sigma_2^*(\emptyset, 1, 0) = 0$. Substituting these into (24):

$$\begin{aligned}&\delta_1 \cdot q \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 0))) + \delta_1 \cdot (1-q) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1))) \\ &\geq \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 0)))\end{aligned}$$

  But recall that, by Assumption 2, $\delta_0 > \delta_1$. For the above inequality to hold, then, one must have $1 - \bar{\beta}(q^*(\emptyset, 0, 1)) > 1 - \bar{\beta}(q^*(\emptyset, 0, 0))$. This is a contradiction to the case we consider.

- Suppose $1 - \bar{\beta}(q^*(\emptyset, 0, 0)) < 1 - \bar{\beta}(q^*(\emptyset, 0, 1))$. Then, by Equation (15), $\sigma_2^*(\emptyset, 1, 0) = 1$. By equations in Appendix C.3.2, this implies $q^*(\emptyset, 0, 0) = 0$. But then, by Equation (12), $\bar{\beta}(q^*(\emptyset, 0, 0)) = 0$ and $1 - \bar{\beta}(q^*(\emptyset, 0, 0)) \geq 1 - \bar{\beta}(q^*(\emptyset, 0, 1))$, a contradiction to the case we consider.

In any case, we obtain a contradiction, and the result follows. □

Using Claim 2 to substitute $\sigma_1^*(\emptyset, 1) = 1$ into Equation (18) gives that, in any PBE:

$$\begin{aligned} q^*(\emptyset, 1, 1) &= \frac{q^2\sigma_1^*(\emptyset, 0)\sigma_2^*(\emptyset, 0, 1)}{q^2\sigma_1^*(\emptyset, 0)\sigma_2^*(\emptyset, 0, 1) + q^2 + q(1-q) + (1-q)q + (1-q)^2} \\ &= \frac{q^2\sigma_1^*(\emptyset, 0)\sigma_2^*(\emptyset, 0, 1)}{q^2\sigma_1^*(\emptyset, 0)\sigma_2^*(\emptyset, 0, 1) + 1} \leq \frac{1}{2} \end{aligned}$$

Then, $\beta(q^*(\emptyset, 1, 1), M, \gamma) = 0$. Because $\delta_0 < 1$ by Assumption 2, Equation (14) implies that $\sigma_2^*(\emptyset, 0, 1) = 1$ in any PBE.

Given these observations, Equation (17) simplifies to:

$$\begin{aligned} \sigma_1^*(\emptyset, 0) \in \arg\max_{\sigma \in [0,1]} & \sigma \cdot \delta_0 \cdot (1-q) \cdot (1 - \bar{\beta}(q^*(\emptyset, 1, 0))) \\ & + \sigma \cdot q \\ & + (1-\sigma) \cdot \delta_0 \cdot ((1-q) + q \cdot (1 - \sigma_2^*(\emptyset, 0, 0))) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 0))) \\ & + (1-\sigma) \cdot \delta_0 \cdot q \cdot \sigma_2^*(\emptyset, 0, 0) \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1))) \end{aligned} \tag{25}$$

Meanwhile, using Claim 2 to substitute $\sigma_1^*(\emptyset, 1) = 1$ into the equation defining $q^*(\emptyset, 0, 0)$ in Section C.3.2 gives that, in any PBE:

$$q^*(\emptyset, 0, 0) = 0$$

Then, by Equation (12), $\bar{\beta}(q^*(\emptyset, 0, 0)) = 0$ in any PBE. Equation (13) simplifies to:

$$\sigma_2^*(\emptyset, 0, 0) \in \arg\max_{\sigma \in [0,1]} \sigma \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1))) + (1-\sigma) \cdot \delta_0$$

This implies:

$$\sigma_2^*(\emptyset, 0, 0) \cdot \delta_0 \cdot (1 - \bar{\beta}(q^*(\emptyset, 0, 1))) + (1 - \sigma_2^*(\emptyset, 0, 0)) \cdot \delta_0 = \delta_0$$

Substituting this into (25), it further simplifies to:

$$\begin{aligned} \sigma_1^*(\emptyset, 0) \in \arg\max_{\sigma \in [0,1]} & \sigma \cdot \delta_0 \cdot (1-q) \cdot (1 - \bar{\beta}(q^*(\emptyset, 1, 0))) \\ & + \sigma \cdot q \\ & + (1-\sigma) \cdot \delta_0 \end{aligned} \tag{26}$$

Where, subtituting our findings so far into the equation defining $q^*(\emptyset, 1, 0)$ gives that, in any PBE:

$$\begin{aligned} q^*(\emptyset, 1, 0) &= \frac{q\sigma_1^*(\emptyset, 1)(1 - \sigma_2^*(\emptyset, 1, 1)) + (1-q)(1 - \sigma_2^*(\emptyset, 1, 1))}{q\sigma_1^*(\emptyset, 1)(1 - \sigma_2^*(\emptyset, 1, 1)) + (1-q)(1 - \sigma_2^*(\emptyset, 1, 1)) + q\sigma_1^*(\emptyset, 0)(1 - \sigma_2^*(\emptyset, 0, 1)) + (1-q)\sigma_1^*(\emptyset, 0)} \\ &= \frac{q \cdot 1 \cdot 0 + (1-q) \cdot 0}{q \cdot 1 \cdot 0 + (1-q) \cdot 0 + q\sigma_1^*(\emptyset, 0)(1 - \sigma_2^*(\emptyset, 0, 1)) + (1-q)\sigma_1^*(\emptyset, 0)} \end{aligned}$$

Note that whenever $\sigma_1^*(\emptyset, 0) > 0$, Bayes' rule applies and $q^*(\emptyset, 1, 0) = 0$. In this case, by Equation (12) and (26), $\sigma_1^*(\emptyset, 0) = 1$. We conclude that there is always a PBE where $\sigma_1^*(\emptyset, 0) = 1$. This completes the description of one PBE.

Remark 3. *There is always a PBE of the game with two rulers where:*

$$\sigma_1^*(\emptyset, 0) = \sigma_1^*(\emptyset, 1) = 1$$
$$\sigma_2^*(\emptyset, 0, 1) = \sigma_2^*(\emptyset, 1, 1) = 1$$

In this PBE, when the issue is non-preordained,

- *If $s = 1$, the aggregate policy is $a = 1$.*
- *If $s = 0$, the aggregate policy is $a = 1$ if and only if both rulers are minority-congruent.*

In any case, there are no revolts.

If $\delta_0 \cdot (1-q) \cdot (1-\bar{\beta}(q')) + q > \delta_0$ for all $q' \in [\frac{1}{2}, 1]$, any PBE of the game with two rulers is a PBE that is described in Remark 3. For the rest of the analysis, suppose $\delta_0 \cdot (1-q) \cdot (1-\bar{\beta}(q')) + q \leq \delta_0$ for some $q' \in [\frac{1}{2}, 1]$. In this case, there is another PBE where $\sigma_1^*(\emptyset, 0) = 0$. Now, Bayes' Rule does not apply to $(\hat{s}, a_1, a_2) = (\emptyset, 1, 0)$, so $q^*(\emptyset, 1, 0)$ can be chosen arbitrarily. Choosing it so that $\delta_0 \cdot (1-q) \cdot (1 - \bar{\beta}(q^*(\emptyset, 1, 0))) + q \leq \delta_0$ ensures that $\sigma_1^*(\emptyset, 0) = 0$ is optimal.

Remark 4. *Suppose $\delta_0 \cdot (1-q) \cdot (1 - \bar{\beta}(q')) + q \leq \delta_0$ for some $q' \in [\frac{1}{2}, 1]$. There is a PBE of the game with two rulers where:*

$$\sigma_1^*(\emptyset, 0) = 0$$
$$\sigma_1^*(\emptyset, 1) = 1$$
$$\sigma_2^*(\emptyset, 1, 1) = 1$$

In this PBE, when the issue is non-preordained,

- *If $s = 1$, the aggregate policy is $a = 1$.*
- *If $s = 0$, the aggregate policy is $a = 0$.*

In any case, there are no revolts on the equilibrium path.

Although the PBE described in Remark 4 is a theoretical possibility, it is a very fragile equilibrium. This is because it relies on the belief $q^*(\emptyset, 1, 0)$ being above $\frac{1}{2}$, even though the scenario where $(\hat{s}, a_1, a_2) = (\emptyset, 1, 0)$ occurs with zero probability. In particular, for this equilibrium to be sustained, the citizens must believe, with high probability, that $s = 1$ following $(\hat{s}, a_1, a_2) = (\emptyset, 1, 0)$. Then, the minority-congruent type of ruler 1 is worried about having $a_2 = 0$ by the majority-congruent type of ruler 2.[3] The reason for this worry is **not** the aggregate action changing. Rather, it is the worry of **revolt**: when citizens encounter $(\hat{s}, a_1, a_2) = (\emptyset, 1, 0)$, they incorrectly infer that state is $s = 1$ with high probability and revolt against aggregate action $A = 0$.

Given the minority-congruent ruler's preference towards $A = 1$ (by Assumption 2), this is a counterintuitive equilibrium. If anything, $(\hat{s}, a_1, a_2) = (\emptyset, 1, 0)$ should make citizens infer that "The state must be $s = 0$, but ruler 1 is minority-congruent and could not resist the temptation of $a_1 = 1$. He is corrected by a majority-congruent ruler 2. But since $A = 0$,

[3]Note that this reasoning falls apart when ruler 1 can observe ruler 2's type, which is the reason why the setup described in the main text does not suffer from equilibrium multiplicity.

I will not revolt against it." The counterintuitivity of PBE described in Remark 4 can be formalized by showing that it fails forward induction. The next result shows this.

Claim 3. *Any PBE described in Remark 4 fails forward induction.*

Proof. Consider a PBE described in Remark 4. In the subgame following $\hat{s} = \emptyset$, the outcome of this PBE is:

$$(s, \hat{s}, a_1, a_2, r) = \begin{cases} (0, \emptyset, 0, 1, 0), & w.p.\ \frac{1}{2}\sigma_2^*(\emptyset, 0, 0), \\ (0, \emptyset, 0, 0, 0), & w.p.\ \frac{1}{2}(1 - \sigma_2^*(\emptyset, 0, 0)), \\ (1, \emptyset, 1, 1, 0), & w.p.\ \frac{1}{2} \end{cases}$$

Our first observation is that the pure strategy of ruler 1 defined as

$$\sigma_1(\emptyset, 0) = \sigma_1(\emptyset, 1) = 1 \tag{27}$$

is a relevant strategy for this outcome. To see this, among the PBE's described in Remark 4, take the one with $q^*(\emptyset, 1, 0)$ such that:

$$\delta_0 \cdot (1 - q) \cdot (1 - \bar{\beta}(q^*(\emptyset, 1, 0))) + q = \delta_0$$

This is the belief that leaves ruler 1 just indifferent between the two actions when $(s, \hat{s}) = (0, \emptyset)$, and such a belief exists due to continuity of $\bar{\beta}(q')$ in q'.

The strategy in (27) is optimal under these beliefs, and therefore it is a relevant strategy. Intuitively, under this PBE, ruler 1 may consider deviating to $a_1 = 1$ when $s = 0$.

Our next observation is that any strategy that includes

$$\begin{aligned} &\sigma_2(\emptyset, 0, 1) = 0, \qquad or \\ &\sigma_2(\emptyset, 1, 1) = 0 \end{aligned}$$

is irrelevant for this outcome. This is because, as discussed above, $q^*(\emptyset, 1, 1) < \frac{1}{2}$ in any PBE. Therefore, $\beta(q^*(\emptyset, 1, 1), M, \gamma) = 0$ in any PBE. By equations (14) and (16), and by Assumption 2, then, $\sigma_2^*(\emptyset, 0, 1) = 1$ and $\sigma_2^*(\emptyset, 1, 1) = 1$ are strict best responses in any PBE. Intuitively, because $a = 1$ is the minority-congruent ruler's favorite outcome, any minority-congruent ruler 2 will not consider deviating to $a_2 = 0$ following $a_1 = 1$.

The discussion above shows that information set $(\hat{s}, a_1, a_2) = (\emptyset, 1, 0)$ is relevant for the outcome under PBE in Remark 4. Moreover, for the outcome to satisfy forward induction, any beliefs in this information set must contain $\sigma_1(\emptyset, 0) = 1$ and rule out $\sigma_2(\emptyset, 0, 1) = 0$ as well as $\sigma_2(\emptyset, 1, 1) = 0$. Under this restriction, the only scenario consistent with $(\hat{s}, a_1, a_2) = (\emptyset, 1, 0)$ occurs when $s = 0$. Therefore, $q^*(\emptyset, 1, 0) = 0$. Under these beliefs, $\bar{\beta}(q^*(\emptyset, 1, 0)) = 0$, and $\sigma_1^*(\emptyset, 0) = 0$ ceases to be optimal. We conclude that any PBE of the type described in Remark 4 fails forward induction. □

In contrast, the outcome of the PBE described in Remark 3 survives forward induction. This is because:

- The strategies that include:

$$\sigma_1^*(\emptyset, 0) = \sigma_1^*(\emptyset, 1) = 1$$
$$\sigma_2^*(\emptyset, 0, 1) = \sigma_2^*(\emptyset, 1, 1) = 1$$

  are relevant for this outcome. After all, they are part of the PBE strategies, and thus they are always optimal.

- The information set $(\hat{s}, a_1, a_2) = (\emptyset, 1, 1)$ is always relevant, because they are reached by the strategies above with strictly positive probability.

- In the information set $(\hat{s}, a_1, a_2) = (\emptyset, 1, 1)$, with the relevant strategies specified above, an equilibrium belief such that $q^*(\emptyset, 1, 1) < \frac{1}{2}$ can always be constructed. Then, the relevant strategies mentioned above remain optimal.

Our findings so far imply the following result.

Proposition 11. *In the extended model of institutional constraints, there is a unique outcome that satisfies forward induction of the Perfect Bayesian Nash Equilibrium of the game. In this outcome,*

$$Pr_{(t_1,t_2)}(A = s) = 1, \quad \textit{if } (t_1, t_2) \neq (b, b).$$

Otherwise,

$$Pr_{(b,b)}(A = 1|\hat{s}, s = 1) = Pr_{(b,b)}(A = 1|\hat{s} = \emptyset, s = 0) = 1$$

and

$$Pr_{(b,b)}(A = 1|\hat{s} = s, s = 0) = \begin{cases} 1 & ; \beta(1, M, \gamma) < 1 - \delta_0 \\ 0 & ; otherwise. \end{cases}$$

There is a revolt only if $\hat{s} = 0$ and both rulers take action 1. This revolt succeeds with probability $\beta(1, M, \gamma)$. Moreover, the expected policy payoff for a majority citizen is

$$\begin{cases} 1 - q^2(1 - p) - \mu & ; \beta(1, M, \gamma) > 1 - \delta_0 \\ 1 - q^2(1 - p\beta(1, M, \gamma)) - \mu & ; \beta(1, M, \gamma) < 1 - \delta_0. \end{cases}$$

D Institutional Constraints on Rulers in Jewish, Greco-Roman, and Christian Traditions

D.1 Ancient Jewish Tradition

Institutional constraints on rulers are also absent in ancient Jewish traditions, covering the ancient Israelites to the end of the Hasmonean Kingdom in 37 BCE. After the period of tribal confederacy, kingship was established by the people (1 Samuel 8) as a Hobbesian remedy for a state of nature in which "everyone did what was right in his own eyes" (Judges 21:25). It is clear that Deuteronomic editors were aware of the downsides of centralized power. The arguments against monarchy in 1 Sam 8 are striking (1 Sam 8: 11-8):[4] "he [(the king)] will take your sons and place them for himself in his chariots and among his horsemen and they will run before his chariots. He will appoint for himself commanders of thousands and of fifties, and some to do his plowing and to reap his harvest and to make his weapons of war and equipment for his chariots. He will also take your daughters for perfumers and cooks and bakers. He will take the best of your fields and your vineyards and your olive groves and give them to his servants. He will take a tenth of your seed and of your vineyards and give to his officers and to his servants. He will also take your male servants and your female servants and your best young men and your donkeys and use them for his work. He will take a tenth of your flocks, and you yourselves will become his servants. Then you will cry out in that day because of your king whom you have chosen for yourselves, but the Lord will not answer you in that day." Subsequent history, according to the Bible, confirmed these prophecies. Halbertal and Holmes (2017, p.67) go as far as arguing that the books of Samuel are early political science, which draw attention to the problem of constraining rulers: "If the sovereign ruler amasses sufficient power to safeguard his people from outside threat, he will also be in a position to redirect that power to torment and abuse his people with sovereign impunity".

However, no institutional remedy is offered from antiquity throughout the Middle Ages, "Instead, the author [of 1-2 Samuel] turned a penetrating gaze onto the punishing costs of sovereign power as such" (Halbertal and Holmes, 2017, p.167). In his study of pre-modern Jewish political thought, Walzer (2012, p.71) argues that "the Bible does not provide... any effective constitutional or political check on the power of kings". "The body negotiating the elevation of the monarch has the opportunity to impose conditions, to extract promises, and to level ultimata. Whether the king after his accession actually paid attention to them is, of course, another matter, about which our sources are too inadequate to permit speculation" (Halpern, 1981, p.222). There was a separation of duties between the king, priests, and prophets. But that was not a substitute for institutional constraints, as the recorded actions of rulers from Saul to the Hasmoneans attest (1-2 Samuel, 1-2 Kings, Josephus's *Antiquities of the Jews*, books XIII-XVII). Kings appointed priests and judges and they promoted, banished or killed prophets to advance their interests.

The king was supposed to follow the divine law. In fact, according to Halpern (1981, p.xx), "Israel's was the first monarchy known to have deposited and preserved a written consti-

[4] All Scripture quotations are taken from the *New American Standard Bible* version 1995.

tution, a document imposing strictures on the exercise of royal authority (Deuteronomy 17-18)". For example, Deuteronomy 17-18 specifies that the king must be an Israelite, must not amass wealth, take many wives, or consider himself better than others; and he must write the laws and read them every day. "Throughout its history, then, Israel's elective autocracy was kingship under the law" (Halpern, 1981, p.249). However, once in power, there were no external constraints on kings except rebellion. The mode of holding a king accountable was mostly internal to the king (God, his conscience, and the prophets' advice and warnings). "A policy focus on political reason, debate in the assembly, popular decision-making – what we might think of as the Greek alternative – was never considered" (Walzer, 2012, p.211).

Why is it, then, that no institutional remedy was provided even in theory? As we discussed, Halbertal and Holmes (2017, p.167) argue that the problem was that the "The political horizons of the author of the Samuel". Similarly, Halpern (1981, p.239) senses "a charming naivite, an idealistic reliance on tribal conservatism, in Samuel's assumption that the 'prophet' could constrain a new and vigorous executive". Walzer (2012, p.204) argues that Jewish thinkers whose works have survived simply put the blame on human imperfections: "Worldly rulers, the power that be, whatever their social or political character, are more likely to disobey than to obey, but disobedience is a function of human recalcitrance and stiffneckedness, not of institutional imperfection".

These arguments ultimately place the problem in the inability of thinkers to even contemplate institutional solutions for a problem that they keenly identified. Thus, according to this literature, for centuries, these thinkers' "political horizon" did not reach that of the Greco-Roman traditions. We find this explanation unsatisfactory. Even more so if we recognize the interactions and cultural exchanges since Alexander's conquests of the late 4th century BCE. Indeed, the 1 Maccabees records a working knowledge of the institutions of the Roman Republic: "Yet with all this, they [Romans] never any of them put on a diadem or wore purple as a mark of magnificence. And they built themselves a senate house, and every day three hundred and twenty men deliberated, constantly planning for the people, that they might conduct themselves properly, and they intrusted the government to one man every year..." (1 Maccabees 8: 14-16).

Moreover, Melamed (2011, p.163) argues that even when Aristole's *Politics* became available to Jewish scholars through Christian-Latin tradition, "Jewish writers continued to translate, expound, and reproduce Plato's *Republic*, the *Ethics*, and commentaries on these works – and not by chance. Their conceptual framework remained Platonic, given the inertia of tradition and their theological commitment". From the 14th to early 17th century (before Spinoza), when, on rare occasions, they directly used *Politics*, it was "mainly to criticize the Platonic model of social organization... rather than the construction of a new political theory" (p.169). An exception is Rabbi Isaac Abravanel's analysis in the context of his commentary on 1 Samuel 8. Possibly reading *Politics* through misrepresentations of Medieval Christian scholars (p.174), he "mistakenly looked upon Aristotle as a partisan of absolute kingship" (p.173, also p. 174). However, he "insists... that this position is wrong. He maintains that monarchy is not a necessity and sees it as doomed to degenerate into tyranny, preferring a mixed regime like that of the Venetian Republic" (p. 173). In sum, in a period when we know that *Politics* was available to Jewish scholars, it was never used to develop a

discussion of institutional constraints on rulers. When such discussions appeared, the author thought Aristotle was in favor of monarchy.

We argue that the comprehensive scope of the law in the Jewish tradition helps make sense of the absence of discussions about institutional constraints on rulers. While the law did not specify institutional constraints on rulers, its scope was extensive, covering various topics including inheritance, marriage, contracts, foreign policy, and various other aspects of criminal and civil law. As Walzer (2012, p.206) argues, "both the legal and prophetic texts have a great deal to say about what political leaders, whoever they are, ought to do. Policy is not free. Leaving royalist ideology [God's anointed king] aside, and speaking still in Greek mode, we can say that God as he was conceived in ancient Israel, did not decree a politics, but he certainty did decree an ethics [policy]". Walzer (2012) derives one consequence of this observation: "Obedience to God's law doesn't require deliberation or arguments or votes; it only requires a moral choice" (p.211). Our focus is on the consequences of these features for political thought. Walzer's (and others') observations point to the theoretical homogeneity of the population's preferences regarding public policy: preferences for God's law. Moreover, when divine law is more extensive, a ruler's wrongdoing is more observable. This, combined with higher societal homogeneity, facilitates disciplining rulers through rebellion. Indeed, Deuteronomic history records various such popular rebellions, e.g., against David and Rehoboam, Solomon's successor who refused to reduce taxes.

D.2 The Western Tradition

Constraining the executive is a common thread in the tradition that starts from Greco-Roman political thought. The existence of these constraints clearly antedates the written justifications we have for them. The Spartan Constitution of Lycurgus, possibly dating to the 7th century BC, divided powers in several important ways. Plutarch (1914) records how the period before Lycurgus had been one with "excessive absolutism" (p.209) and with kings "hated for trying to force their way with the multitude" (p.209). Aristotle comments on Lycurgus' attitudes towards the Spartan kings that "he shows a great distrust of their virtue" (Aristotle, 1996, p.53). Lycurgus therefore created a council of elders which countered the fact that the "ruling power was still in a feverish condition" (Plato, 2016, p.123) and "by having an equal vote with them in the matters of highest importance, brought safety and due moderation into the councils of state" (Plutarch, 1914, p.219-221). This was critical because "the civil polity was veering and unsteady, inclining at one time to follow the kings towards tyranny, and at another to follow the multitude towards democracy" (Plutarch, 1914, p.221). About 130 years later the *ephors* were added to the system of government and, as Plato puts it, "curbed it" (Plato, 2016, p.123). They were specifically tasked with monitoring the kings. The constitutional experiments of Athens as documented by Aristotle (1996) involve similar attempts to balance powers. By the time of the famous reforms of Solon in 594 BC, Athenian kings had already disappeared with the main executive body being nine archons who served for one year. There was an assembly of all adult male citizens and two councils the Boule and the Areopagus, where the latter had "the duty of watching over the laws" (Aristotle, 1996, p.216). Plutarch notes that this was designed "thinking that the city with its two councils, riding as it were at double anchor, would be less tossed by the surges, and would keep its

populace in great quiet" (Plutarch, 1914, p.455). Solon tinkered with the organization and membership of the different councils and explicitly justified what he was doing as balancing power between different groups, particularly the rich and the poor. Further reforms which democratized and reorganized the institutions were implemented by Cleisthenes.

Plato and Aristotle subsequently theorized the success and failings of Greek constitutions.[5] Though Plato's *Republic* advanced a utopian solution, proposing mechanisms for abolishing political conflict, in his *Laws* he developed more practical institutions if utopia proved not to be possible. As von Fritz (1954, p.v) puts it "Plato is concerned with the danger inherent in absolute political power, and that he is of the opinion that there must be a check to all political power, and that this must be done by distributing power over several government agencies which counterbalance one another." Aristotle outlined a famous ranking of constitutions which started with the three ideal forms of government, followed by their perversions. The ideal forms ran in order from best to worst: kingship, aristocracy, polity. Their perversions were tyranny, oligarchy, democracy. Critically, while kingship might be best in theory, it relied on having someone of unlikely "excellence" and quickly deteriorated into tyranny, which was the worst form of government, even worse than democracy, the perversion of polity. Indeed, Aristotle follows his discussion of the likely character of kings with an exposition of the institution of ostracism (Aristotle, 1996, p.81-82).[6] Instead, Aristotle preferred a blend of aristocracy and polity – mixed government. In contrast to Plato's *Republic*, which focuses on the selection and training of rulers, institutional mechanisms to constrain rulers appear in Aristotle's *Politics* (Aristotle, 1996). These institutional constraints include term limits (Book 5, Ch. 8, Paragraphs 6-7, 12-13), audits (6,4,5-7), prevention of excessive power disparity (5,8,11; 3,16,16), control by setting interest against interest (5,8,14), and collective decision-making/multiple rulers (3,15,8). Ryan (2012, p.98-99) sums up the lessons from Aristotle's analysis in the following terms: "The problem in designing a constitution is to distribute power so as to give every incentive to those who have it to use it for the common good.... What is needed is what later came to be called checks and balances".

These Greek beginnings had a profound influence over subsequent constitutional thought, particularly of the Roman Empire. Polybius, who was himself Greek, conducted a famous analysis of the success of Rome attributing it to the mixed constitution initially supposedly devised by Romulus. In it power was distributed between "the consuls... the Senate... and the common people" (Polybius, 2010, p.380). Polybius attributed the idea of such a system to the Spartans who "bundled together all the merits and distinctive characteristics of the best systems of government in order to prevent any of them going beyond the point where it would degenerate into its congenital vice" (Polybius, 2010, p.378-379). He is very clear, referring to the basic systems of government that were mixed, that Lycurgus "wanted the potency of each system to be counteracted by the others" (Polybius, 2010, p.379) so that "nowhere would any of them tip the scales or outweigh the others". Any one of them on their own has the same sorts of problems that Aristotle identified so that in the past, for example, "kingship gave way to tyranny" (Polybius, 2010, p.376). He is definitive that "we should take the best system of government to be the one that combines all three of these

[5]Previous writers discussed some aspects of them, though less comprehensively; see Sinclair (2012).

[6]See Teegarden (2013) for an analysis of ancient Greek legislation aimed at blocking the rise of tyrants.

constitutions" (Polybius, 2010, p.372). The view that the secret of the Romans' success was due to the type of mixed government that emerged was also asserted by Cicero. In his political life, contesting with Caesar and Pompey, Cicero was well aware of the danger of tyranny. In *The Republic* he discusses at length the dangers, pointing out that "although Cyrus of Persia was an exceptionally just and wise monarch" it was highly dangerous to have a government "managed by one man's nod and wish" since this led to the rule of the "cruelly capricious Phalaris. His is the image into which, by a smooth and easy process, the rule of one man degenerates" (Cicero, 1998, p.20-21). Cicero was also clear that the main advantage of a mixed government was "although those three original forms easily degenerate into their corrupt versions... such things rarely happen in a political structure which represents a combination and judicious mixture" (Cicero, 1998, p.32).

The rise of Christianity and the collapse of the western Roman empire created some significant challenges to the Greco-Roman tradition. This is most obvious is the work of St. Augustine, who wrote right after Alaric's sack of Rome in 410. For Augustine, the type of state Cicero had imagined here on earth was an impossibility and everything was focused on the afterlife. This led to a downgrading in the importance of political institutions. As he put it:

> As far as this mortal life is concerned, which is spent and finished in a few days, what difference does it make under what rule a man lives who is soon to die, provided only that those who rule him do not compel him to do what is impious and wicked. – Augustine (1998, p.217)

The standard interpretation of this is that God created the king and that unless one's religious beliefs were threatened, one had to accept his authority. In this, he built upon earlier churchmen, particularly St. Paul who argued that (Colossians 1:16):

> For by him were all things created that are in heaven, and that are in earth, visible and invisible, whether they be thrones, or dominions, or principalities, or powers: all things were created by him and for him.

Furthermore, "the powers that be are ordained by God... whosoever therefore resisteth the power, resisteth the ordinance of God" (Romans, 13:1-5). Augustine put it in the following way: "all these things he bestows upon good and evil men alike. And among these things is imperial sway also, of whatever scope, which He dispenses according to His plan for the government of the ages" (Augustine, 1998, p.235). Augustine, therefore, did not take a view on things like the mixed constitution, and tyrannicide, which was explicitly advocated by Cicero, was definitely out. The powers that be were created by God. In addition, the only reason that states existed was because of sin, and "the discipline that even bad rulers imposed provided a partial remedy for sin in that it restrained men from indulging to the full criminal proclivities of fallen nature" (Tierney, 2008, p.39).

Though Ryan (2012, p.199) uses the statements of St. Paul and St. Augustine to argue that "The conventional view down to the sixteenth century was that if a ruler required his subjects to repudiate Christ, they did not have to comply; short of that they had to obey", it is also clear that the rise of Christianity and Christian approaches to politics left the old concerns about tyranny alive. These concerns took different forms and institutional guises

and parted ways until coming together in the late Middle Ages (see Acemoglu and Robinson, 2019 for a discussion of these channels).

First, and most directly, though works such as Aristotle's *Politics* were lost until the middle of the 13th century, and Polybius and Cicero re-discovered only later, clear manifestations of Greco-Roman political institutions persisted. This is most evident in the Italian city-states. Before Aristotle was translated into Latin, Venice already had its elaborate mixed constitution with its "monarchic doge, aristocratic Senate, and democratic Great Council" (Blythe, 1992, p.278). At the same time a score of northern polities, including Arezzo, Milan and Pisa, had created republican institutions, consuls, and were governed by an annually elected executive, known as the podestà, who was always an outsider and who was subjected to an elaborate system of accountability (Waley and Dean, 2010). Just as in classical Greece, the emergence of these institutions preceded their written justifications. Ryan (2012, p.281) argues that by "the eleventh century they reinvented many features of the early Roman republic, in particular the appointment of magistrates to very short periods of office as a defense against tyranny.... These city states were in many respects genuine revivals of the city-state of antiquity". These institutions were heavily theorized later, notably by Florentine writers such as Guicciardini and Machiavelli (particularly Machiavelli, 1903).

The second stream stemmed from the political institutions of the Germanic tribes that conquered the western Roman empire. They maintained key elements of their highly participatory politics based around assemblies; see King (1988) and Wickham (2017). These were famously described by the Roman historian Tacitus in his book *Germania*: "The leading men take counsel over minor issues, the major ones involve them all... The assembly is also the place to bring charges and initiate trials in capital cases.... Likewise in these assemblies are chosen the leaders who administer justice" (Tacitus, 1999, p.81-82). Almost 800 years later similar political institutions during the Carolingian polity were described by Hincmar of Rheims: "At that time the custom was followed that no more than two general assemblies were to be held each year.... All the important men, both clerics and laymen attended this general assembly. The important men came to participate in the deliberations, and those of lower station were present in order to hear the decisions and occasionally also to deliberate concerning them, and to confirm them not out of coercion but by their own understanding and agreement" (Hincmar, 1980, p.222). In Britain this assembly was called the *witan*. It is not a coincidence that King John signed the Magna Carta in 1215 on a site at Runnymede where the Anglo-Saxon *witans* used to meet (Pantos and Semple, 2004). This shows a direct continuity between pre-Norman institutions and the regime begun by William the Conqueror in 1066. Interestingly, the Magna Carta also specified a complex institutional design to monitor whether or not John implemented the policies. Maddicott (2012) develops in detail the argument that the roots of England's parliament are in its pre-Norman Germanic representative institutions and this view was common already in the 16th century, e.g., Fortescue (1997). In 1583 the Elizabethan courtier Sir Thomas Smith could write "The most high and absolute power or the realme of Englande, is in Parliament" (Smith, 1982, p.78). Part of the mechanism through which these institutions perpetuated themselves and ended up in theories of the state was via feudalism, since this was a set of institutions based on contract. In line with this, Figgis (1956, p.9) notes: "it is in the feudal system that the

contractual theory of government took its rise".[7] Echos of these Germanic institutions arise all over western Europe. Charters similar to the Magna Carta were granted to Catalonia in 1205; Hungary in 1222; and Germany in 1220. Parliaments, estates and similar institutions sprouted up (Bisson, 1973; Myers, 1975), all prior to the rediscovery of Aristotle or Polybius.

The third stream flowed through the organization of the Catholic Church fused with elements of Roman Law. The church was viewed as a voluntary community and the pope was elected by the bishops. Roman law contained the idea of a corporation, which was an entity with a legal existence separate from that of its particular members, and the will of the corporation could be determined by a majority of its members. The members delegated power to an official who acted on behalf of the community. "In the normal doctrine of Roman private corporation law, the agent's powers were not only derivative, but revocable and subject to modification" (Tierney, 2008, p.26). In 1140 Gratian produced an influential collection of church law which led to a great deal of debate on the organization of the church. This debate entertained the fact that a pope could misbehave (Tierney, 2008, p.16). Then "Around 1200 [religious scholars] began to discern that the legal concept of a corporation could define the structure... of the universal church itself and of a general council representing the church" (Tierney, 2008, p.20). As early as 1214 Pope Innocent III convoked a general council of not just bishops but representatives of many churches and religious chapters. The implications of this Roman law model for secular authority were profound. "In this theory the ruler held a position analogous to that of any elected official of a Roman law corporation" (Tierney, 2008, p.26) and Tierney argues that it led to notions of government by consent and "a complex doctrine of mixed or limited monarchy" (Tierney, 2008, p.27). These arguments became particularly powerful within the church at the time of the Great Schism when rival popes emerged and a series of councils met to settle the dispute, most notably in Constance in 1415. These councils claimed supreme authority within the church and ended up deposing three popes. This "conciliar movement", for a constitutionally governed church, had repercussions for the organization of secular authority; see Black (1988).

In short, though Augustine's view was influential in the 840 years between the sack of Rome and the rediscovery of Aristotle, the old views about the potential abuse of power by kings, and the need to take institutional precautions against it, persisted. Supporting this, Ryan (2012, p.219) suggests in the context of the reaffirmation of John of Salisbury's vindication of tyrannicide in the mid-12th century, that "similar ideas must have been in circulation from the end of antiquity without leaving any written evidence of their existence". In the context of feudal institutions, Ryan (2012, p.195) also notes: "The Polybian view of mixed government aligns easily with the medieval idea that a king should rule with the advice of an aristocratic council and seek consent for taxation".

These different streams start to come together in Thomas Aquinas' 13th century attempt to synthesize Catholic teaching with classical philosophical ideas. He was perhaps the first writer to absorb the newly rediscovered works of Aristotle and, reflecting this, he notes that

[7]There is an extensive and controversial literature about the origins of representative institutions in Medieval Europe. Particularly disputed is the connection to Germanic tribal institutions. For our purposes, the main point is the prevalence of these institutions which clearly balanced and checked monarchical power; see Bisson (1973) for key essays and an overview of the literature.

"the rule of one, which is the best, is preferred, but that it can turn into tyranny, which is the worst" (Aquinas, 2002, p.17). When it came to political institutions the solution to this was that "all should have some share in the government; for an arrangement of this kind secures the peace of people, and all men love and defend it, as is stated in *Politics* II" (Aquinas, 2002, p.53-54). As in Cicero, there is no compunction against removing tyrants. In addition, political institutions should be structured to avoid tyranny: "governance of the kingdom should be so arranged that the opportunity to tyrannize be removed and the king's power should be so tempered that he cannot easily become a tyrant" (Blythe, 1992, p.48-49). Blythe (1992, p.49) concludes that Aquinas's discussion implies that "the king's power be limited or controlled by other governmental institutions so that it cannot exceed what is proper". Aquinas found direct inspiration for mixed government in the Bible in particular arguing that this was how the state was organized at the time of Moses:

> Moses and his successors governed the people in such a way that each of them was ruler over all. But they chose seventy two elders according to their virtue... and this was aristocracy. But this arrangement was also democratic in that they were chosen from all the people. – Aquinas (2002, p.54)

Tierney's summary of the logic is that "The mixed regime was best, he wrote, because each element checked, 'tempered', the other two" (Tierney, 2008, p.90).

Aquinas was followed by a series of writers who elaborated on his ideas and extended them in various ways sketching out theories of consent and constitutional rule. Marsilius of Padua (d. 1342) and William of Ockham (d. 1347) further advanced justifications for popular sovereignty. Marsilius extensively quotes Aristotle and discusses his taxonomy of different forms of government and makes it clear that a key advantage of popular sovereignty is that it avoids tyranny. He notes that government "savours of tyranny... the more it departs from these conditions, viz. the consent of those subjects and a law established to the common advantage" (Marsilius of Padua, 2005, p.47). Moreover, "giving the power of legislation to one alone creates a space for tyranny" (Marsilius of Padua, 2005, p.78). Marsilius also discusses other institutional mechanisms to reduce the potential for tyranny, for example, elected monarchs are to be preferred to hereditary ones (p.105). Ockham advocated for a mixed constitution with a king and council where "the element of balance is present in that the council exists in part to check the excesses of the king" (Blythe, 1992, p.183). One of his arguments in favor of such a constitution, as opposed to a simple monarchy, was that "one can be more easily corrupted than many" Blythe (1992, p.182). Finally, John of Paris advanced ideas about both mixed government and notions based on the corporation. His position was that "government is a stewardship... exercised for the common good of individual and corporate owners. Should it not carry out its mandate, it is removable on the authority of the people" (Coleman, 2000, p.133).

Nevertheless, sixteenth century Europe was ruled by powerful kings, even if most had to deal with parliaments. The century saw an ideological struggle between those who wished to make kings subject to popular sovereignty and those who wished to make kings more absolutist. Advocates of popular sovereignty coalesced around what is known as "resistance theory" – whether, contrary to the Augustine tradition, people had the legitimate right to resist and dethrone a king (see Kingdon, 1991 and Skinner, 1978 for authoritative discussions). Early

versions of this emanated from the struggle of Luther and Calvin against papal control. Interestingly, the advocates of absolutism explicitly set themselves against the notion of a mixed constitution, instead emphasizing that many classical writers, such as Aristotle, Aquinas, and Cicero (e.g. Cicero, 1998, p.25), thought kingship the best type of government. Theoretically, as Bodin (1992, p.92) put it "to combine monarchy with democracy and aristocracy is impossible and contradictory.... For if sovereignty is indivisible, as we have shown, how can it be shared by a prince, the nobles, and the people at the same time?" To sustain this argument he went on to argue that previous writers, like Polybius or Cicero, had in fact misinterpreted the nature of the Spartan and Roman constitutions stating "We shall conclude, then, that there is not now, and never was, a state compounded of aristocracy and democracy, much less of the three forms of state" (Bodin, 1992, p.103). It was not just that sovereignty was indivisible, dividing powers led to anarchy as Sir Robert Filmer put it in a famous tract of 1648, *The Anarchy of a Limited or Mixed Monarchy* (Filmer, 1991).

Resistance theory began to take on a more institutionalized form at the start of the seventeenth century (Llord, 1991; Sommerville, 1999). Franklin (1991, p.304) notes, for example, that though notions of mixed government and executive constraints were well understood, other concepts like the separation of powers were only nascent in the sixteenth century. The first constitution to feature explicit separation of powers was the English Instrument of Government written after the parliamentary victory in the civil wars; see Vile (1967). This provided the basis for Locke's analysis in his *Second Treatise on Government*. Locke provides a clear rationale for the existence of the state but warns against tyranny since "monarchs are but men" and he asks whether "men are so foolish, that they take care to avoid what mischiefs may be done them by pole-cats and foxes; but are content, nay think it safety, to be devoured by lions?" (Locke, 2003, p.140). Locke then argues that the design of institutions is key to constraining potential lions. Power has to be devolved to a legislature containing "collective bodies of men, call them senate, parliament, or what you please" (Locke, 2003, p.141) and because of potential conflicts of interest, "the legislative and executive power come often to be separated" (Locke, 2003, p.164).

This tradition, by way of Montesquieu's *Spirit of the Laws*, subsequently had a major impact on the thinking and institutional design of the US and French constitutions. Though the Federalist Papers mention only Montesquieu explicitly, other writings confirm the importance of Locke; see, for example, Mace (1979) and Wills (1981). Of particular interest are the writings of John Adams. In his 1778 book *A Defence of the Constitutions of Government of the United States of America* he traces the genealogy of the key ideas of the constitution, particularly executive constraints, checks and balances, and the separation of powers. Included in the sources are Plato and Solon, with Polybius and Machiavelli's *Discourses of Livy* receiving particular attention.

An important factor underpinning this intellectual history is the fact that in the Greco-Roman and Christian traditions, humans legislated much of their own laws – the collection of Roman law in the 6th century under Justinian was one manifestation. Church law, Canon law, never had the same status as the Sharia. Indeed Pennington (2008, p.386) notes "Christian communities lived without a comprehensive body of written law for more than five centuries. Consequently, in the early church, 'canon law' as a system of norms that

governed the church or even a large number of Christian communities did not exist." Instead, in Europe, local traditions and Roman law were powerful and "no single authoritative compilation of Church law came into existence before the twelfth century" (Herzog, 2018, p.49). When finally Canon law was systematized by Gratian, his compilation, the *Decretum* (Decree), had to compete with other sources of law. Pirie (2021, p.163) notes how there were "interminable debates about its relation to the 'civil law' ". Moreover, while the Decree was emerging "rulers and judges were inspired by the example of Justinian to create new codes for their people" (Pirie, 2021, p.163) all a very far cry from the Islamic world. At the same time there was also a clear sense of legislation and the legitimacy of legislation. Thus, Marsilius of Pauda wrote in *Defensor Pacis*, "the judgment, command, and execution of any arraignment of the prince for his demerit or transgression should take place through the legislator, or through a person or persons established for this purpose by the authority of the legislator" (Klosko, 2012, p.312-3; see also Coleman, 2000, Ch.4).

However, the concern with setting the best law or with the concentration of both legislative and executive power in the prince, king, caliph, *ḥākim*, *ulu al-amr*, or whoever was in charge was less concerning in Islamic (and Jewish) traditions, in which it was assumed that much of the law was divine and set by God. The comprehensive scope of the law and the perceived homogeneity of the society in Islamic and Jewish traditions (all were supposed to follow the divine law), in turn, would make a ruler's deviations more observable and coordination on revolt against such deviant rulers more expedient.

References

Acemoglu, Daron and James A. Robinson. 2019. *The Narrow Corridor*. New York: Penguin.

Aquinas, Thomas. 2002. *Aquinas: Political Writings*. Cambridge Texts in the History of Political Thought New York: Cambridge University Press.

Aristotle. 1996. *Aristotle: The Politics and the Constitution of Athens*. Cambridge Texts in the History of Political Thought New York: Cambridge University Press.

Augustine. 1998. *Augustine: The City of God against the Pagans*. Cambridge Texts in the History of Political Thought Cambridge: Cambridge University Press.

Bisson, Thomas N. 1973. *Medieval Representative Institutions, Their Origins and Nature*. New York: The Dryden Press.

Black, Antony. 1988. The Conciliar Movement. In *The Cambridge History of Political Thought: 350-1450*, ed. J.H. Burns. New York: Cambridge University Press.

Blythe, James M. 1992. *Ideal Government and the Mixed Constitution in the Middle Ages*. Princeton: Princeton University Press.

Bodin, Jean. 1992. *On Sovereignty*. New York: Cambridge University Press.

Boleslavsky, Raphael, Mehdi Shadmehr and Konstantin Sonin. 2021. "Media Freedom in the Shadow of a Coup." *Journal of the European Economic Association* 19(3):1782–1815.

Cho, In-Koo and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *The Quarterly Journal of Economics* 102(2):179–221.

Cicero. 1998. *The Republic and The Laws*. Oxford World's Classics London: Oxford University Press.

Coleman, Janet. 2000. *A History of Political Thought: From the Middle Ages to the Renaissance*. Malden: Blackwell.

Figgis, John Neville. 1956. *Studies of Political Thought from Gerson to Grotius, 1414-1625*. New York: Cambridge University Press.

Filmer, Robert. 1991. *Filmer: 'Patriarcha' and Other Writings*. Cambridge Texts in the History of Political Thought New York: Cambridge University Press.

Fortescue, John. 1997. *Sir John Fortescue: On the Laws and Governance of England*. Cambridge Texts in the History of Political Thought New York: Cambridge University Press.

Franklin, Julian H. 1991. Sovereignty and the mixed constitution: Bodin and his critics. In *The Cambridge History of Political Thought: 1450-1700*, ed. J.H. Burns and Mark Goldie. New York: Cambridge University Press.

Govindan, Srihari and Robert Wilson. 2009. "On Forward Induction." *Econometrica* 77(1):1–28.

Halbertal, Moshe and Stephen Holmes. 2017. *The Beginning of Politics: Power in the Biblical Book of Samuel*. Princeton: Princeton University Press.

Halpern, Baruch. 1981. *The Constitution of the Monarchy in Israel*. The Netherlands: Brill.

Herzog, Tamar. 2018. *A Short History of European Law*. Cambridge: Harvard University Press.

Hincmar. 1980. On the Governance of the Palace. In *The History of Feudalism*, ed. David Herlihy. London: Macmillan.

King, P.D. 1988. The Barbarian Kingdoms. In *The Cambridge History of Medieval Political Thought: c.350–c.1450*, ed. J.H. Burns. Cambridge: Cambridge University Press p. 123–154.

Kingdon, Robert M. 1991. Calvinism and Resistance Theory, 1550–1580. In *The Cambridge History of Political Thought: 1450-1700*, ed. J.H. Burns and Mark Goldie. New York: Cambridge University Press.

Klosko, George. 2012. *History of Political Theory: An Introduction. Volume I: Ancient and Medieval. Second Edition*. Oxford: Oxford University Press.

Llord, Howell A. 1991. Constitutionalism. In *The Cambridge History of Political Thought: 1450-1700*, ed. J.H. Burns and Mark Goldie. New York: Cambridge University Press.

Locke, John. 2003. *Two Treatises of Government and A Letter Concerning Toleration*. New Haven: Yale University Press.

Mace, George. 1979. *Locke, Hobbes, and the Federalist Papers: An Essay on the Genesis of the American Political Heritage*. Carbondale: Southern Illinois University Press.

Machiavelli, Niccolo. 1903. *Discourses on Livy*. New York: Charles Scribner's Sons.

Maddicott, John R. 2012. *The Origins of the English Parliament, 924-1327*. New York: Oxford University Press.

Marsilius of Padua. 2005. *Marsilius of Padua: The Defender of the Peace*. Cambridge Texts in the History of Political Thought Cambridge: Cambridge University Press.

Melamed, Abraham. 2011. Aristotle's *Politics* in Medieval and Renaissance Jewish Political Thought. In *Well Begun is Only Half Done: Tracing Aristotle's Political Ideas in Medieval Arabic, Syriac, Byzantine, and Jewish Sources*, ed. Vasileios Syross. Tempe: Arizona Center for Medieval and Renaissance Studies pp. 145–186.

Myers, Alec Reginald. 1975. *Parliaments and Estates in Europe to 1789*. San Diego: Harcourt Brace Jovanovich.

Pantos, Aliki and Sarah Semple. 2004. *Assembly Places and Practices in Medieval Europe*. London: Four Courts Press.

Pennington, Kenneth. 2008. The Growth of Church Law. In *The Cambridge History of Christianity, Volume 2: Constantine to c.600*, ed. Augustine Casiday and Frederick W. Norris. Cambridge: Cambridge University Press pp. 386–402.

Pirie, Fernanda. 2021. *The Rule of Laws: A 4,000 Year Quest to Order the World*. New York: Basic Books.

Plato. 2016. *Plato: Laws*. Cambridge Texts in the History of Political Thought Cambridge: Cambridge University Press.

Plutarch. 1914. *Lives, Volume I: Theseus and Romulus. Lycurgus and Numa. Solon and Publicola*. Translated by Bernadotte Perrin. Loeb Classical Library 46 Cambridge: Harvard University Press.

Polybius. 2010. *The Histories*. Oxford World's Classics New York: Oxford University Press.

Ryan, Alan. 2012. *On Politics: A History of Political Thought from Herodotus to the Present*. New York: W.W. Norton Co.

Sinclair, Thomas Alan. 2012. *A History of Greek Political Thought*. New York: Routledge.

Skinner, Quentin. 1978. *The Foundations of Modern Political Thought, Vol. 1: The Renaissance*. New York: Cambridge University Press.

Smith, Sir Thomas. 1982. *De Republica Anglorum*. New York: Cambridge University Press.

Sommerville, Johann P. 1999. *Royalists and Patriots: Politics and Ideology in England, 1603-1640*. New York: Addison Wesley Longman.

Tacitus. 1999. *Tacitus: Germania*. Clarendon Ancient History Series Oxford: Oxford University Press.

Teegarden, David. 2013. *Death to Tyrants!: Ancient Greek Democracy and the Struggle against Tyranny*. Princeton: Princeton University Press.

Tierney, Brian. 2008. *Religion, Law and the Growth of Constitutional Thought, 1150-1650*. New York: Cambridge University Press.

Vile, Maurice John Crawley. 1967. *Constitutionalism and the Separation of Powers*. Oxford: Clarendon Press.

von Fritz, Kurt. 1954. *The Theory of the Mixed Constitution in Antiquity: A Critical Analysis of Polybius' Political Ideas*. New York: Columbia University Press.

Waley, Daniel and Trevor Dean. 2010. *The Italian City-Republics, Fourth Edition*. London: Routledge.

Walzer, Michael. 2012. *In God's Shadow: Politics in the Hebrew Bible*. New Haven: Yale University Press.

Wickham, Christopher. 2017. "Consensus and Assemblies in the Romano-Germanic Kingdoms." *Vorträge und Forschungen* 82:389–426.

Wills, Garry. 1981. *Explaining America: The Federalist*. New York: Doubleday.