Judicial Independence and Human Rights in Autocracies*

Mehdi Shadmehr[†] Raphael Boleslavsky[‡] Tom Ginsburg[§]

Abstract

Why do some autocracies empower their judiciaries to uphold human rights even though independent judiciaries can prevent the repression of opposition? We develop a theoretical framework to explain why dictatorships benefit from judiciaries that restrict the government's use of coercion, thereby addressing the contradiction between the function of independent judiciaries and their institutional origins. By granting a degree of judicial independence, the regime shapes how the public views the state's use of coercion. When the judiciary is more effective in preventing state repression, the public will have more confidence in the legitimacy of coercive acts that are not blocked by the judiciary. This shifts public opinion against the opposition in favor of the regime, reducing the public's incentive to support the opposition. Unlike propaganda and censorship that directly control the information that citizens receive, partially independent judiciaries enable autocracies to control how the public processes the informational content of coercion.

Keywords: Judicial Independence, Human Rights, Autocracy, Repression, Protest, Legitimate Coercion

Word Count: 9780

*This paper is the second part of a larger project with the title, "Institutions, Repression, and the Spread of Protest." We wish to thank Nemanja Antic, Dan Bernhardt, Brandice Canes-Wrone, Odilon Câmara, Charles Cameron, Christopher Cotton, Georgy Egorov, Ignacio Esponda, Justin Fox, Camilo Garcia-Jimeno, Soroush Ghazi, Sergei Guriev, Arvind Magesan, Adam Meirowitz, Steve Morris, Michael Peress, Nicola Persico, Andrea Prat, Carlo Prato, Kristopher Ramsay, Luis Rayo, Kevin Reffett, Kai Steverson, Milan Svolik, Scott Tyson, Dan Treisman, Léonard Wantchékon, and participants and discussants at EEA-ESEM, SIOE, APSA, MPSA, SPSA, and Miami Economic Theory conferences, and seminar audiences at the University of Utah, NYU-Abu Dhabi, FIU, Washington University, and Princeton.

†Harris School of Public Policy, University of Chicago, and Department of Economics, University of Calgary. E-mail: `mehdi@uchicago.edu`

‡Department of Economics, University of Miami. E-mail: `r.boleslavsky@miami.edu`

§University of Chicago Law School. E-mail: `tginsburg@uchicago.edu`

1 INTRODUCTION

Recent decades have witnessed a global expansion of judicial power, extending even to authoritarian regimes. Courts in some dictatorships now decide cases on an array of important topics, including human rights, and rule against the regime with some regularity. This phenomenon creates a puzzle. Why would dictators allow any degree of judicial independence in the area of human rights, which might result in more freedom for opposition figures?

In this paper, we argue that a partially independent judiciary can help make repression more effective, stabilizing the authoritarian regime. The mechanism we focus on is informational: By granting a degree of judicial independence, the regime can manipulate public opinion about the state and the opposition in favor of the state, stabilizing the regime by reducing the risk that repression leads to backlash public protests. Citizens recognize that some alternatives to the regime are worse: while an autocratic leader may be perceived as bad, the groups opposing him are not necessarily good. In this uncertain environment, the government's use of coercion against an opposition is informative about both the state and the opposition. Citizens perceive some acts of coercion by the state as "legitimate coercion" and others as illegitimate repression (Almond 1956; Mansbridge 2012, 2014). Our key insight is that the state can shape public opinion about the legitimacy of its coercive acts by committing itself to limit the repression of legitimate opposition. Thus, the state establishes a partially independent judiciary to restrict its own acts of repression to influence the public's perception of state coercion, bolstering public support for the regime.

Our argument paints partially independent judiciaries as instruments of social control in the spirit of Foucault: their effectiveness lies in manipulating how the people process the information content of repression (Foucault 1981). Unlike censorship or propaganda that directly affect the *information* available to the public, judicial independence affects the way that information is *processed* by the general public in indirect and subtle ways—which the judges themselves may not intend. Judges who sometimes have the authority to tie the state's hand can indirectly enable regimes to control how the public views acts of coercion, promoting the interests of the state (Fuller 1975).

The literature on judicial empowerment in authoritarian regimes has offered several explanations as to why authoritarian rulers would grant independence to courts (Moustafa 2014). First, independent judiciaries protect property rights and enforce contracts, promoting economic development. Second, they impose restrictions on a regime's successors, ensuring that outgoing elites will not face unfettered opposition rule and will have their personal interests preserved (Ginsburg

2003; Hirschl 2004; Epperly 2013). Third, independent judges can monitor the activities of low-level officials, mitigating principal-agent problems between the elite and the bureaucracy (Shapiro 1981; Verner 1984; Rosberg 1995; Peerenboom 2002; Ginsburg and Moustafa 2008). None of these accounts helps us to understand why courts in authoritarian regimes would ever protect human rights of regime opponents. By treating judiciaries in authoritarian regimes merely as government agents, or simply as window-dressing for the regime, conventional approaches treat any protection of human rights or judicial constraint on repression as an unintended byproduct of judicial independence, whose primary purpose lies elsewhere (Helmke and Rosenbluth 2009, p. 358). There is thus a tension between the usual accounts of the origins of judicial independence (which require the judiciary to contribute to autocratic stability) and its function in protecting human rights. In contrast, our positive account of judicial power integrates human rights protection into the calculus of the dictator, and shows how it can benefit the regime.

Our model begins with the fact that, in any country, there is typically a variety of opposition groups who make claims on the state. Some opposition groups demand changes that will benefit the public, while others demand harmful changes that will benefit only a narrow group, and may leave everyone else worse off. This is a shared feature of autocracies and democracies. Just as the democratic government in America has to contend with dangerous supremacist groups like the KKK, the autocratic government in Syria has to contend with dangerous religious groups like ISIS. Surely, some autocratic regimes feature a host of opposition groups who demand free and fair elections, or freedom of press and assembly. However, there is also another set of opposition groups who demand different forms of autocratic government that may be worse than the status quo. For example, under the Pahlavi regime in Iran, nationalists demanded democracy, while a group called the Mojahedin demanded the establishment of a government that combined Islam and Marxism to create a class-less society (Abrahamian 1982, 1989). Ali Shariati, an Iranian intellectual whose teachings influenced the Mojahedin and helped their recruitment, believed that a revolution's leader must not "be tempted by Western liberalism," and leave the revolution's fate to "the shaky hand of democracy." After the revolution, the leader must not establish a democracy; rather, he and those with proper ideological training should continue their "authentic revolutionary leadership" for "a few generations" until the people become ready (Shariati 1969, p. 171-177). Iran today is a society in which democracy is constrained by theocratic rulers.

The variety of opposition groups coupled with the difficulties in acquiring precise information means that the general public faces a difficult decision about whom to support in a conflict between an opposition group and the state. That is, when the state uses coercive force against an opposition group (i.e., when the state uses repression), should the public acquiesce or should it join the opposition's cause against the state? Without public support, the opposition has little chance against the coercive apparatus of the state. However, if the public joins the opposition, it will be difficult for the state to subdue the whole country. But the public's decision involves a complicated tradeoff. Siding with the opposition risks replacing a bad regime with a much worse one—it would be far worse to live under ISIS than under Assad. Not siding with the opposition (i.e., acquiescing to the state's repression) risks missing a chance for reforms—living under Mosaddegh, the democratic prime minister deposed by the US in 1953, would be much better than living under the Shah. The public's uncertainty about the opposition and the state makes information an essential element in the public's decision whether to support the opposition by joining their protest.

This fundamental uncertainty has largely been overlooked in the literature, which has focused on strategic uncertainty among the dissidents or the regime's supporters (Persson and Tabellini 2009; Bueno de Mesquita 2010; Shadmehr and Bernhardt 2011; Boix and Svolik 2013; Chen and Suen 2016; Rundlett and Svolik 2016; Tyson and Smith 2018; Shadmehr 2019). Moreover, the large literature on information manipulation in autocracies has not studied the informational content of the state's actions about its legitimacy (Egorov et al. 2009; Edmond 2013; Shadmehr and Bernhardt 2015; Barbera and Jackson 2016; Huang et al. 2019). Shadmehr and Boleslavsky (2019) depart from the literature by observing that, because state actors often know more than the general public about the state and the opposition, the very use of coercive force by the state may become informative about the nature of conflict between the state and the opposition. They investigate the implications of the information content of repression and explore the role of international intervention in that interaction. They do not, however, consider the effects of domestic institutions such as judiciaries in the interactions between the state, the opposition and the public. We use their model as our first benchmark (Benchmark 1).

The key insight of this paper is that, because the act of repression is informative about the nature of conflict, the state can benefit from ex ante committing to a repression policy that generates useful information. For example, by restricting its own ability to repress reform minded opposition, the state makes it more likely that the public views repression less negatively and acquiesces to it. The establishment of a partially independent judiciary does exactly that. Independent judiciaries tie the state's hands and restrict its use of repression—the extent of which depends on the degree of judicial independence. That is, partially independent judiciaries are institutions that act as commitment devices that enable the state to commit to a repression policy.

Our paper contributes to the literature on authoritarian institutions, which has to date primarily focused on parties and legislatures, military-civilian relationships, and power-sharing in dictatorships, providing rationales for their emergence (Acemoglu et al. 2008; Gandhi 2008; Cheibub et al. 2010; Svolik 2012, 2013; Boix and Svolik 2013; Geddes et al. 2014). But we have no account of why it would be in the interests of an authoritarian regime to set up judiciaries that sometimes protect dissidents. Indeed, the literature highlights how the judiciary protects dissidents, for example, by "refusing to prosecute people for exercising their human rights, and seeking to prosecute the police for abuses of power" (Hilbink 2012, p. 598-99). "Reform-minded judges may work to push the envelope through their judgments...in the direction of political reform" (Moustafa 2014, p. 288; Rosberg 1995, Ch. Four; Ip 2012). Thus, we model the degree of judicial independence as the likelihood with which the judiciary strikes down the government's attempt to repress legitimate opposition groups. When the judiciary succeeds in blocking the state's attempt to repress, it hurts the state in that particular interaction. However, the average gains in public opinion from knowing that the state cannot use arbitrary force may more than compensate for those losses. Our model formalizes these ideas and investigates the complex strategic considerations that arise.

To show the relevance of our model, we highlight how the logic of our model is consistent with Pereira's (2008) analysis of security courts in the military regimes of Brazil (1964-85), Chile (1973-1990), and Argentina (1976-83). While the Argentinian regime did not use security courts, both Chilean and Brazilian regimes did. In Brazil, security courts held public proceedings using civilian judges and gave defendants significant rights of defense—the acquittal rate in these courts were around 50%, a remarkably high level by any standard. Pereira argues that the courts in Brazil and Chile "helped marginalize and de-legitimate the opposition, [and] purchased some credibility for the regimes both at home and abroad" (p. 32). Indeed, these are the mechanisms that drive our results. Our model studies how and under what conditions courts can "de-legitimize the opposition," "purchase credibility for the regimes," and increase the longevity of the regime. As Pereira observes, "It is also plausible that the political trials in both Brazil and Chile helped consolidate and prolong authoritarian rule. It may not be a coincidence that the shortest-lived military regime of the three examined here, the Argentine, was also the one that engaged in the most extrajudicial repression" (p. 32).

Typical decision-makers in an autocracy, by definition, often do not represent the preferences of the public. For example, they may want to repress all opposition and maintain the status quo.

These bad types of officials design the persistent institutional structure of the country under the regime, including the judiciary. However, sometimes good officials whose preferences are congruent with the public may manage to rise to a high rank in the regime and form a government. Mohammad Khatami's presidency in Iran (1997-2005) may be an example. Our second benchmark model considers a setting in which the typical bad types can set up institutions that enable them to commit to a precise repression policy for the regime that determines the likelihood that the regime cracks down on each type of opposition. Thus, any time a bad type is in power, this repression policy goes into effect. When good types come to power, they may effectively dismantle these features behind the scenes, or at least temporarily put the policy on hold, and make their own decisions. In sum, our second Benchmark considers a setting in which the bad government has full commitment power, but the good government does not. When the public sees an instance of conflict between the government and the opposition, it does not know whether the government or the opposition is good or bad, but it does know the institutions of the regime and hence the repression policy that will be invoked if the bad government is in charge. This level of precise commitment is the best that the typical government of an autocracy (the bad government) can achieve. Such a precise and elaborate commitment power may seem unrealistic, but it provides a benchmark against which one can evaluate the effect of judiciaries.[1]

We show that by setting up a judiciary and optimally choosing a degree of judicial independence, the regime can generate the same outcomes as when the regime has the strong commitment power of the second Benchmark model. That is, instead of designing intricate institutions that commit the bad autocratic government to a likelihood of repressing each type of the opposition, the regime can design a judiciary with an optimal (for the regime) degree of judicial independence that determines the likelihood with which the judiciary blocks the repression of legitimate opposition.

Knowing that the judiciary can partially limit the repression of legitimate opposition, the public becomes more pessimistic about an opposition that the judiciary does not protect. In particular, the public reasons that perhaps the judiciary was not independent enough to protect that opposition group, but then again, maybe this was an opposition group for which the judiciary would approve

[1] From a theoretical perspective, our analysis treats commitment power in a novel way. Unlike the political delegation literature in which the uninformed principal commits to restrict the informed agent's decisions (Alonso and Matouschek 2008; Callander and Krehbiel 2014), in our model it is the informed agent—the bad government—who commits ex-ante to restrict *his own* subsequent choices. Moreover, our application requires that the typical state actors in the autocratic regime (i.e., the bad government) set up the institutional structure of the regime, which provide the basis of commitment power. This implies that only the bad government can commit, generating multiple equilibria that require appropriately designed refinements—in the Online Appendix we analyze the setting in which good and the bad government types can commit. That the informed agent (the government) also has a private conflict of interest, and that only one of the agent types can commit are among the new features of our analysis.

the use of coercion. As a result, the public becomes less inclined to side with the opposition and protest against the state's use of coercion. This, in turn, raises the good government's incentive to use coercion against bad opposition groups, reducing the public's incentive to protest even more. Thus, a complementary feedback loop emerges between the good government's incentives to coerce the bad opposition and the public's incentives to side with the government. This logic sheds light on why citizens who demand greater protection through law also tolerate severe repression of others. The People's Republic of China, for example, has responded to citizen demands with a set of legal institutions that are demonstrating greater independence and scope than at any time in the country's history (Zhang and Ginsburg 2019). At the same time, the Chinese Communist Party is engaged in mass repression of religious minorities like the Falun Gong and Uyghur Muslims in Xinjiang, and continues to hold trials of political dissidents. Ordinary citizens might observe this repression and come to doubt the good intentions of the government. In our account, increasing the capacity of courts is a useful device for distinguishing legitimate and illegitimate coercion. By helping the public to separate legitimate coercion from arbitrary abuses, the judiciary contributes to regime stability. Indeed, the World Justice Project, which surveys citizens in over 100 countries on many dimensions of the justice system, has shown increasingly positive perceptions of the Chinese judicial system. The courts are increasingly considered to be free of corruption, even as repression has worsened in recent years.

The restraint afforded to the regime by establishing a partially independent judiciary enables the typical bad government in the regime to exploit the strategic complementarity between the strategies of the occasional good government and the public, thereby reducing the likelihood of the public's protest and the endogenous costs of repression. This benefits the regime. We identify conditions under which a partially independent judiciary enables the bad government to raise repression and yet face a lower likelihood of public protest, or to maintain the same level of repression that it would without the restraints of the judiciary, but reduce the risk of public protest. The restraint on repression that is created by the judiciary is crucial: without it, the bad government would change behavior and increase repression, which would raise the public's incentives to protest, thereby upsetting the equilibrium. We show that the regime's decision to establish a partially independent judiciary is non-monotone in the popularity of the opposition. In contrast, the relationship between the regime's popularity and the degree of judicial independence is monotone. As the regime's popularity falls, the regime grants a higher degree of judicial independence. In fact, when the regime is sufficiently unpopular, it grants a high enough degree of judicial independence that the public never joins the protest following repression.

Although we focus on the origins of judicial independence and its function in autocracies, our paper also contributes to the literature on state repression and backlash protest. This literature studies the effect of democratic institutions (e.g., constraints on the executive) in reducing state repression, the effectiveness of repression in suppressing dissidents, and the potential for backlash protests instigated by repression (Poe and Tate 1994; Francisco 2004; Davenport 2007a, 2007b; Earl 2011; Siegel 2011; Aytaç and Stokes 2019). Our paper combines these branches of the literature by showing how partially independent judiciaries, which constrain the executive, interact with the informational frictions involved in backlash protests. It also shows that partially independent judiciaries can worsen human rights and repression. Repressive regimes can design judiciaries to take advantage of the public's information deficit, so that the very features that constrain the state also enable it to control how the public perceives the state's use of coercion, reducing the risk of backlash protest and stabilizing the regime.

Our paper proceeds in the next section by studying Benchmark 1, in which the government has no ex ante restriction in repressing the opposition. The following section analyzes Benchmark 2, in which the bad government has full commitment power to ex ante commit to a repression strategy—i.e., the probabilities to repress each type of opposition. We next develop and analyze our model of partially independent judiciaries. A Conclusion follows. Proofs are in an online appendix.

2 BENCHMARK MODEL 1: NO COMMITMENT

As a benchmark, we consider the model of Shadmehr and Boleslavsky (2019). There are two strategic players, a government and a public. There is also a non-strategic player, an opposition group who protests, demanding a set of social changes. The opposition has two types: a "good" opposition (type g) demands reforms that would benefit the public. A "bad" opposition (type b) demands reforms that would harm the public. The government has two types also, "good" (G) or "bad" (B). A good government prefers good reforms over the status quo and prefers the status quo over bad reforms. The bad government prefers good reforms to bad reforms, but prefers the status quo over all types of reform. Both government types derive a payoff from staying in power. The government observes the opposition's type, but the public does not observe the types of the government or the opposition. There is a common prior that the government is bad with probability $p \in (0, 1)$, and the opposition is bad with probability $q \in (0, 1)$.

The game proceeds as follows. First, nature chooses the types of the government and the opposition. The government observes the opposition's type, and then decides whether to concede to the opposition or repress him. If the government concedes, the game ends. If it represses, the public decides whether to protest.When the public protests, the government is replaced by the opposition, who implements its preferred reform. Otherwise no reform is implemented and the government remains in office. Payoffs are realized at the end of the game.

The public's payoff under the status quo is normalized to zero. If a good reform is implemented, the public's payoff rises to $b > 0$. If a bad reform is implemented, the public's payoff falls $-b < 0$. Thus, protesting has a downside: joining a *bad* opposition's protest is worse for the public than supporting the government. The government's payoff depends on whether or not it retains office, the terminal policy, and its type. If a government is toppled, then its payoff is normalized to zero. If the government remains in office without implementing any reforms, its payoff is 1. If the good government retains office by conceding to the good opposition's demands, its terminal payoff is $1 + \delta_g$; if it concedes to a bad opposition, its payoff is $1 - \delta_b$, where $0 < \delta_i$ and $\delta_b < 1$. In contrast, if the bad government concedes to the good opposition, its payoff is $1 - \alpha_g$; if it concedes to the bad opposition, its payoff is $1 - \alpha_b$, where $0 < \alpha_g < \alpha_b < 1$. We focus on the case where $\delta_b < \alpha_g$, so that the bad government's incentives to repress the good opposition is larger than the good government's incentives to repress the bad opposition. This captures the bad government's strong incentives to maintain the status quo.

Strategies and Equilibrium. The government's strategy is a quadruple, $(\rho_g^G, \rho_b^G, \rho_g^B, \rho_b^B) \in [0,1]^4$, where ρ_j^i is the probability with which the type $i \in \{G, B\}$ government represses the type $j \in \{g, b\}$ opposition. The public's strategy is a probability, $\pi \in [0,1]$, representing the probability of joining the protest following repression. The equilibrium concept is Perfect Bayesian subject to $D1$ refinement (Fudenberg and Tirole 2000, p. 452).

Analysis. Shadmehr and Boleslavsky (2019) characterize the equilibria of this model. We present their equilibrium characterization here. Figure 1 demonstrates.

Proposition 1 ***(Shadmehr & Boleslavsky 2019)*** *In equilibrium, $\rho_g^G = 0$ and $\rho_b^B = 1$. Further,*

- ***Low Protest:*** *If $q > 1/2$, there is a unique equilibrium with $\rho_b^G = 1$, $\rho_g^B = 1$, and $\pi = 0$.*
- ***High Protest:*** *If $q < \frac{p}{1+p}$, there is a unique equilibrium with $\rho_b^G = 0$, $\rho_g^B = \frac{q}{1-q}$, and $\pi = \alpha_g$.*
- ***Intermediate Protest:*** *If $\frac{p}{1+p} \leq q \leq 1/2$, both the above equilibria exist. In addition, there is an equilibrium with $\rho_b^G = \frac{p}{1-p}\frac{1-2q}{q}$, $\rho_g^B = 1$, and $\pi = \delta_b$.*

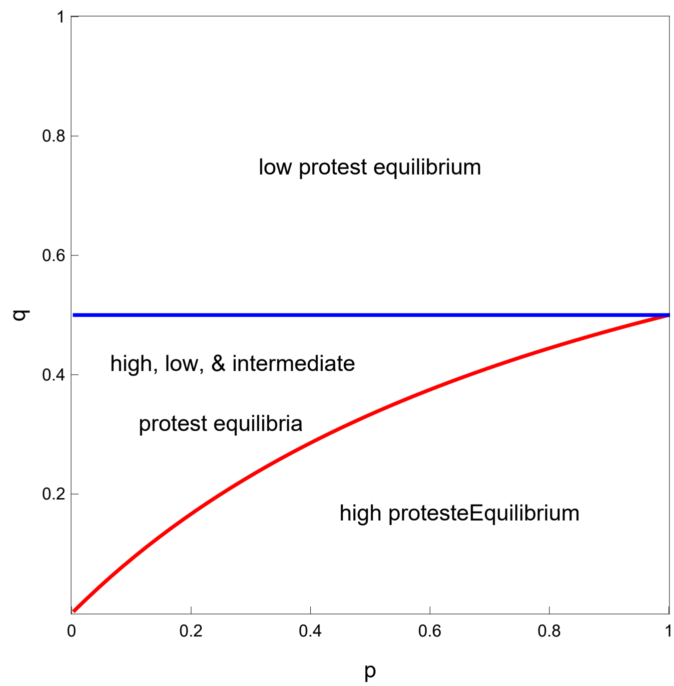

Figure 1: Equilibrium without Institutional Restrictions (Benchmark 1). The curve is $q(p) = \frac{p}{1+p}$. Proposition 1 describes the equilibrium.

When the opposition is sufficiently likely to be bad, the public does not want to support the opposition, and hence the government can repress without the worry of backlash protest. This constitutes the Low Protest Equilibrium. In contrast, when the opposition is sufficiently likely to be good, the public is so inclined to protest that the possibility of backlash protest completely deters the good government from repressing the bad opposition and even somewhat deters the more repression-oriented bad government from repressing the good opposition. This constitutes the High Protest Equilibrium. The Intermediate Protest equilibrium is the middle ground between these two, where the public is likely enough to protest that it deters the good government, but not the bad government.

3 BENCHMARK MODEL 2: FULL COMMITMENT

In the previous benchmark, the government can not ex ante commit itself to a repression policy, e.g., it can not restrain itself from repressing good opposition. In this section, we analyze a game in which the bad government commits to a repression strategy prior to the game of Benchmark 1. The bad government moves first, choosing a strategy (r_b, r_g), where r_i is the probability with which it will repress the type i opposition, $i \in \{b, g\}$. This choice is observable to all players. The rest of the game is identical to Benchmark 1. Next, nature moves, deciding what type of government will be in power. With probability p a bad government will be in power. With the remaining probability $1-p$ a good government will be in power. Next, the opposition protests. The opposition is bad with probability q and good with probability $1-q$. The public knows these probabilities, but it does not observe the types of the government or the opposition. Then, the government responds to the opposition's protest. If the government is bad, it represses the type i opposition with the probability r_i that was chosen earlier. If the government is good, it chooses the probability ρ_i^G with which it represses the type i opposition. If the government concedes to the opposition, the game ends. If the government represses the opposition, the public decides whether or not to join the opposition's protest. If the public does not join the protest, the government remains in power. If the public does join the protest, the government is replaced and the opposition's reforms are implemented.[2]

Analysis. The extensive form of this game has a continuum of subgames, each of which follows a particular choice of (r_b, r_g) by the bad government. Because the bad government has already committed to (r_b, r_g), the good government and the public are the strategic players in these subgames. The incentives of the good government and the public are similar to the Benchmark 1, except that, with commitment, they observe the bad government's strategy. Therefore, the public's best response is a mapping from the bad government's *observed* strategy (r_b, r_g) and the good government's *anticipated* strategy ρ_b^G into a protest probability.[3]

Lemma 1 *Given the strategy of the bad government* (r_b, r_g) *and the strategy of the good government* ρ_b^G*, the public's best response is:*

$$\pi(r_b, r_g; \rho_b^G) = \begin{cases} 1 & \text{if} \quad F(r_b, r_g) > K\rho_b^G \\ [0,1] & \text{if} \quad F(r_b, r_g) = K\rho_b^G \\ 0 & \text{if} \quad F(r_b, r_g) < K\rho_b^G \end{cases}$$

where $F(r_b, r_g) \equiv ((1-q)r_g - qr_b)b$ *and* $K \equiv \frac{1-p}{p}qb$.

Function $F(r_b, r_g)$ is the net expected payoff from protesting versus not protesting against a bad government. With probability qr_b, the bad government has repressed a bad opposition group, and protesting reduces the public's payoff by b. With probability $(1-q)r_g$, the bad government has repressed a good opposition group, and protesting raises the public's payoff by b. When $F(r_b, r_g) > K$, the public has a dominant strategy to always protest in the subgame, and hence

[2]In the Online Appendix, we analyze the case where both the good and the bad government can commit to a repression strategy.

[3]As before, $\rho_g^G = 0$ in equilibrium because the good government prefers good reforms to the status quo.

the good government will never repress in the unique equilibrium of the subgame. Similarly, when $F(r_b, r_g) < 0$, the public has a dominant strategy to never protest and the good government always represses the bad opposition. In contrast, when $0 \leq F(r_b, r_g) \leq K$ the public's best response depends on the good government's strategy, generating the potential for multiple equilibria.

Lemma 2 in the Online Appendix shows that when $0 < F(r_b, r_g) < K$, the subgame has three equilibria: in one the public does not protest, in another the public always protests, and in the third the public protests with probability δ_b. This creates the possibility that the public and the good government switch from one equilibrium to another following each (r_b, r_g). We impose the natural restriction that, for any (r_b, r_g) such that $0 < F(r_b, r_g) < K$, only one of the three equilibria is played in the subgame. In the text, we focus on the interesting case in which the protest probability is δ_b whenever $F(r_b, r_g) \in (0, K)$.[4] Thus, depending on the bad government's strategy, the public's equilibrium protest probability takes one of the three values of 0, δ_b, or 1. As the bad government represses the good opposition more, the likelihood that the protest spreads (weakly) increases as $F(r_b, r_g)$ moves from $F < 0$, where the public never protests, to $F \in (0, K)$, where it sometimes protests, to $F > K$, where the public always protests. Therefore, the bad government has a tradeoff between repressing the good opposition with a higher probability and facing a higher probability of protest by the public.

The bad government's ability to commit to a strategy creates another complication. When $F(r_b, r_g) = K$, a continuum of equilibria are possible in the subgame: any $\pi \in [0, \delta_b]$ with $\rho_b^G = 1$ can be part of the equilibrium—Lemma 2 in Online Appendix. Without commitment, these equilibria only arise in knife-edge cases. With commitment, however, the bad government may select a strategy with $F(r_b, r_g) = K$. Therefore, a refinement is needed to limit the set of possible equilibria. Using a similar logic to the trembling hand refinement, we show that the subgame with $\pi = \delta_b$ is uniquely selected when $F(r_b, r_g) = K$. In particular, we introduce stochastic shocks to the bad government's strategy, showing that as the support of the distribution of shocks vanishes, the equilibrium converges to one in which $\pi = \delta_b$ (see Theorem 1 in the Online Appendix). Given these equilibrium selections, Proposition 2 formally describes the equilibrium, and Figure 2 illustrates the four equilibrium regions that arise.

[4] Propositions 5 and 6 in the Online Appendix analyze the cases in which the protest probability is zero and one whenever $F(r_b, r_g) \in (0, K)$, establishing that no protest takes place in equilibrium.

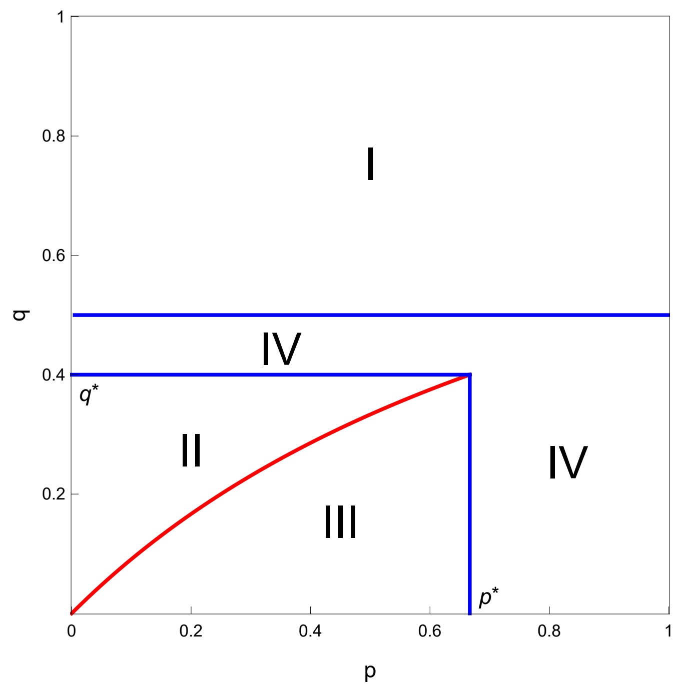

Figure 2: Equilibrium with Full Commitment. The red curve is $q(p) = \frac{p}{1+p}$. Proposition 2 describes the equilibrium.

Proposition 2 *Suppose the bad government can commit to a strategy strategy* (r_b, r_g)*. In equilibrium,*

1. *If* $\frac{1}{2} < q$*, then the strategies are identical to the Low Protest equilibrium (region I).*

2. *If* $\frac{p}{1+p} < q < q^*$*, strategies are identical to the Intermediate Protest equilibrium (region II).*

3. *If* $q < \frac{p}{1+p}$ *and* $p < p^*$*, then* $r_b = \rho_b^G = 1$*,* $r_g = \frac{q}{1-q}\frac{1}{p}$*, and* $\pi = \delta_b$ *(region III).*

4. *Otherwise,* $r_b = \rho_b^G = 1$*,* $r_g = \frac{q}{1-q}$*, and* $\pi = 0$ *(region IV).*

Moreover, $q^* = \frac{1}{2}(1 - \frac{\delta_b}{\alpha_g})$ *and* $p^* = \frac{\alpha_g - \delta_b}{\alpha_g + \delta_b}$.

By repressing the bad opposition more, the bad government gains directly from saving the concession costs, and indirectly from making the public update less negatively about the government and more negatively about the opposition. Therefore, the bad government always represses the bad opposition. However, when choosing how much to repress the good opposition, the bad government faces a tradeoff between the probabilities of repression and protest: as the discussion of Lemma 1 shows, increasing r_g can increase the protest probability from zero to δ_b, or from δ_b to 1. Clearly, the bad government does not repress so much that the public always protests following repression.

But it may trade off a lower probability of protest for a higher probability of repression. That is, the bad government's choice, *in equilibrium*, boils down to choosing between *two* thresholds of repressing the good opposition. Let $r_g = r_1$ be the low threshold and r_2 be the high one. The bad government can choose r_1 and eliminate the public's protest ($\pi = 0$), or it can choose $r_2 > r_1$ and risk the public's protest following repression with probability δ_b.

When $q > 1/2$, the likelihood of a bad opposition is sufficiently high that even if the bad government always represses, the public never protests.[5] When $\frac{p}{1+p} < q < 1/2$, if the bad government always represses, then the public protests with probability δ_b. Thus, the bad government has two potentially optimal choices: (1) repress both types of opposition with probability one ($r_2 = 1$) and face the public's protest with probability δ_b, or (2) limit the repression of the good opposition to $r_1 = \frac{q}{1-q} < 1$, ensuring that the public does not protest. When the opposition is relatively likely to be good ($q < q^* = (1/2)(1 - \frac{\delta_b}{\alpha_g})$), repressing the good opposition with a higher probability is more valuable to the bad government, and it chooses the first option (region II). Otherwise, $q^* < q < 1/2$, and it chooses the second option (in region IV).

In contrast, when $q < \frac{p}{1+p}$, the likelihood of a bad opposition is small enough that if the government were to always repress, the public would protest with probability one. Therefore, the bad government's equilibrium choices boil down to two:[6] (1) If it represses the good opposition with a smaller probability, $r_1 = \frac{q}{1-q}$, then the public does not protest in equilibrium. (2) If it represses the good opposition with a larger probability $r_2 = \frac{q}{1-q}\frac{1}{p}$, then the public protests with probability δ_b. As p increases, the public believes that the government is more likely to be bad. As a result, it becomes more inclined to join the opposition's protest, limiting the government's ability to repress the good opposition: $r_2(p)$ is decreasing in p. Therefore, when p is large (region IV), the government chooses option (1) in which the public does not protest. When p is smaller (region III), $r_2(p)$ is sufficiently large that the gain from raising the repression probability from r_1 to $r_2(p)$ offsets the corresponding increase in the protest probability from 0 to δ_b, and hence the government chooses option (2).

These outcomes are the best that a bad government can achieve with commitment power. Before we explore the effects of full commitment, we present our model of partially independent judiciaries. We show that the bad governments can achieve the same outcomes with the more limited form of commitment that they can get with a partially independent judiciary, which sometimes blocks the

[5] The thresholds $q = 1/2$ and $q = \frac{p}{1+p}$ determine the value of $F(1, 1, q)$, where we have made the dependency of F on q explicit. When $q > 1/2$, $F(1, 1, q) < 0$, and hence the bad government who chooses $r_g = 1$ faces $\pi = 0$. When $\frac{p}{1+p} < q < 1/2$, $F(1, 1, q) \in (0, K)$, and hence the bad government who chooses $r_g = 1$ faces $\pi = \delta_b$. When $q < \frac{p}{1+p}$, $F(1, 1, q) > K$, and hence the bad government who chooses $r_g = 1$ faces $\pi = 1$.

[6] These choices correspond to $F(1, r_1) = 0$ and $F(1, r_2) = K$.

repression of good opposition groups. Thus, the results obtained under full commitment in this section (Benchmark 2) can be obtained with a partially independent judiciary.

4 A MODEL OF PARTIALLY INDEPENDENT JUDICIARY

We conceptualize the essence of judicial independence in the area of human rights as the judiciary's ability to block the government's attempt at repressing legitimate opposition. The degree of judicial independence is captured by the extent to which the judiciary succeeds in blocking such government's attempts. Thus, we model the degree of judicial independence as a probability, $C \in [0,1]$, with which the judiciary blocks a government's attempt to repress a good opposition group. We develop a model to study whether and when autocratic regimes choose a positive degree of judicial independence.

We now describe the model. First, the bad government decides a degree of judicial independence $C \in [0,1]$. The subgame that follows is similar to our Benchmark 1. Nature moves, putting a bad government in power with probability $p \in (0,1)$, and putting a good government in power with the remaining probability $1-p$. Nature also determines the opposition's type, which will be bad with probability $q \in (0,1)$, and good with the remaining probability $1-q$. The government knows its own type and the opposition's type, but the public does not know either type. Next, the government in power chooses whether to repress the opposition or concede to it. If the government concedes, the game ends. If the government attempts to repress the opposition and is not blocked by the judiciary, the public observes repression and decides whether to protest. However, the government's attempt to repress a good opposition is blocked by the judiciary with probability C, in which case the government is forced to concede to the opposition and the game ends. The public knows the presence and the function of the judiciary, but does not observe whether a judiciary has attempted to block the government's repression. That is, when the public observes repression, it does not know whether the judiciary has attempted to block that the government's repression and failed, or the judiciary has not attempted to block the government.

Proposition 3 shows that if the bad government commits to a level of judicial independence, subject to an equilibrium selection that we describe in the Online Appendix, equilibrium outcomes are equivalent to the ones that arise if it commits to a repression strategy. That is, the game in which the bad government commits to a level of judicial independence C is outcome-equivalent to the game in Benchmark 2, in which the bad government can fully commit to a repression strategy (r_b, r_g).

Proposition 3 *Partially independent judiciaries can generate the same outcome as when the regime can fully commit to a precise repression strategy (Benchmark 2).*

To see the key intuition, note that in regions III and IV of Figure 2 (the two cases where the bad government's commitment power is relevant), the bad government supports an equilibrium with a low likelihood of protest by committing to repress the good opposition with probability $r_g < 1$. However, given the low likelihood of protest ($\pi < \alpha_g$), if the bad government's commitment power were removed, then the bad government would like to deviate by repressing the good opposition with probability one; absent other restrictions on the government's strategy, this would upset the equilibrium. Now, suppose that when the bad government's commitment power is removed, an independent judiciary is simultaneously introduced, which blocks the government's attempt to repress a good opposition with probability $C = 1 - r_g$. The bad government still *tries* to repress the good opposition with probability one, but it only succeeds with probability r_g. Therefore, the likelihood that the good opposition is repressed is identical to the equilibrium with precise commitment, the public's beliefs are unchanged, and the strategies of the public and good government continue to be best responses. Thus, the equilibrium that arises when the bad government commits to its strategy can be recreated by selecting a particular level of judicial independence.

To make this argument precise, one must also select the equilibrium in a consistent manner in both settings—recall that both settings feature multiple equilibria in the subgame that follows the bad government's choice. We describe this equilibrium selection in the Online Appendix. With this equilibrium selection, the equilibrium of the game in which the bad government commits to a repression strategy and the equilibrium of the game in which the bad government chooses the level of judicial independence are outcome-equivalent. That is, the probability that each type of government *successfully* represses each type of opposition and the probability that the public joins the protest are identical in both.

Proposition 3 allows to use our characterization in Benchmark 2 to study the origins and functioning of judicial independence. Corollary 1 demonstrates the sharp effect of designing a partially independent judiciary in preventing the spread of protest.

Corollary 1 *In equilibrium, the public does not protest upon repression if and only if the probability that the government is bad* <u>*or*</u> *the probability that the opposition is bad is sufficiently* <u>*large*</u>*. There exists $(p^*, q^*) \in (0,1)^2$ such that $\pi = 0$ in equilibrium if and only if $q > q^*$ or $p > p^*$.*

Proposition 2 and Figure 2 show that the bad government uses a partially idependent judiciary

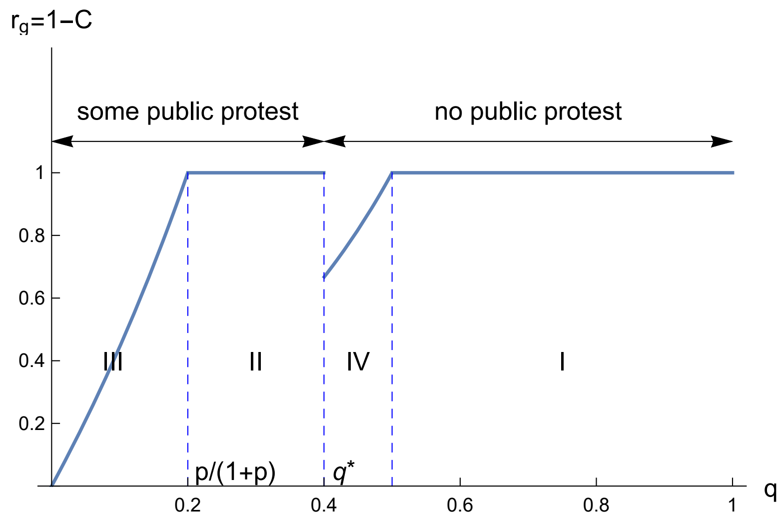

Figure 3: The bad government's equilibrium likelihood of repressing the good opposition as a function of the prior likelihood q that the opposition is bad—for a given $p < p^*$. A higher level of repression r_g corresponds to a lower level of judicial independence: $C = 1 - r_g$.

to limit its own repression in regions III and IV. To understand the effects of this endogenous restriction, we analyze how the government's equilibrium behavior varies with q, focusing on the interesting case when $p < p^*$. Figure 3 illustrates. An increases in q has two conflicting effects: (1) it reduces the bad government's ex-ante incentives to repress the good opposition because the opposition is less likely to be good, and (2) it reduces the public's incentives to protest following repression because it would be (i) less likely that a good opposition is repressed, and (ii) more likely that the public would be supporting a bad opposition.[7]

Recall that the bad government's equilibrium choices are effectively between a low likelihood r_1 of repressing the good opposition (high judicial independence, $C_1 = 1 - r_1$) and a high likelihood r_2 of repressing the good opposition (low judicial independence, $C_2 = 1 - r_2$). When q increases, both r_1 and r_2 rise (until r_2 reaches 1), but r_2 rises faster so that $r_2 - r_1$ also increases. The reason that r_2 rises faster is that when $r_g = r_1$, the public does not protest ($\pi = 0$), and hence the reduction in the likelihood that protest can lead to bad reforms (i.e., effect (ii)) is irrelevant. In contrast, when $r_g = r_2$, the public protests with probability $\pi = \delta_b$, and hence both (i) and (ii) contribute to reducing incentives to protest, and hence to raise r_2.

When q is small (region III), so that the good opposition is relatively likely, both $(r_2 - r_1)$ and the ex-ante value of repressing the good opposition are sufficiently large that it is worth it for the

[7]Although (i) and (ii) seem to be the flip sides of the same coin, the public incurs the associated costs of (ii) *only* if it protests with a positive probability.

bad government to risk protest and choose the high repression threshold, i.e., low level of judicial independence. As q increases, as long as $r_2 < 1$, $r_2 - r_1$ rises in region III until r_2 reaches 1 at the boundary of region II, $q = \frac{p}{1+p}$. Now, because $r_2 = 1$ and cannot rise any further, increases in q reduce $r_2 - r_1$ until the bad government's gains of raising repression from r_1 to $r_2 = 1$ become so small that it not worth to raise the likelihood of the public's protest from 0 to δ_b. This happens at the boundary $q = q^*$ of region IV, where the bad government switches from the higher threshold to the low threshold $r_1(q)$ (i.e., from low to high level of judicial independence $C_1(q) = 1 - r_1(q)$). From this point on, increases in q keep raising $r_1(q)$ until $r_1 = 1$ at the boundary $q = 1/2$ of region I. As Figure 3 shows, when $p < p^*$, both the bad government's likelihood of repression and whether or not it ex-ante limits repression are non-monotone in q. The bad government limits its repression when q is low (region III) or high (IV), but not when it is intermediate (region II) or very high (region I). The bad government established a partially independent judiciary when q is low (region III) or high (IV), but not when it is intermediate (region II) or very high (region I).

Corollary 2 *When $p < p^*$, the bad government's likelihood of repressing the good opposition is non-monotone in the prior likelihood q that the opposition is bad. Equivalently, when $p < p^*$, the degree of judicial independence C is non-monotone in the prior likelihood q that the opposition is bad. In contrast, the degree of of judicial independence C is increasing in the prior likelihood p that the government is bad.*

One may think that when a regime limits its own repression by establishing a partially independent judiciary, it must repress the good opposition *less often* in order to gain by manipulating the public's equilibrium beliefs, so that they protest less following repression. However, this argument does not take into account that sometimes good governments are in charge in autocratic regimes, and even though they may not be able to change the fundamental institutions of the regime, they do respond to the existing institutions such as partially independent judiciary. Knowing that the bad government cannot successfully repress the good opposition beyond a certain level due to the judiciary, the public updates less negatively about the government and more negatively about the opposition following repression, and hence it is less inclined to protest. This raises the good government's incentive to repress the bad opposition, which, in turn, further reduces the public's incentive to protest, and so on. It is this effect that allows the bad government, in region III, to repress *more* than what it would do absent the judiciary, and yet to face a lower probability of the public's protest.[8]

[8]From Propositions 1, when $q < \frac{p}{1+p}$, including region III, the bad government represses with probability

Corollary 3 *When $q < q_1(p)$ and $p < p^*$, the bad government represses the good opposition more often in the presence of a partially independent judiciary.*

Given the low likelihood that the public protests following repression, absent the judiciary, the bad government would raise repression and always repress the good opposition, thereby upsetting the equilibrium. The judiciary benefits the bad government by enabling it to support the equilibrium in which the good government always represses the bad opposition; the bad government leverages this by raising the repression of the good opposition. Similar strategic considerations arise in region IV. There, the bad government represses as much as it would do in the High Protest equilibrium absent the judiciary (in Benchmark 1), but eliminates the public's protest all together. To summarize, the bad government exploits its commitment power afforded to him by a partially independent judiciary in two distinct ways: in region III, it *raises* repression, and yet *lowers* the likelihood of the public's protest; and in region IV, it maintains the same level of repression (as in the High Protest equilibrium of Benchmark 1 without the judiciary), but eliminates the risk of the public's protest.

By creating an independent judiciary that limits repression of the good opposition, the bad government manipulates public opinion in its favor, reducing the public's incentive to protest. Consequently, the good government is more inclined to repress the bad opposition, further reinforcing the favorable shift in the public's beliefs. These effects reduce the endogenous cost of repression for the state, resulting in a lower likelihood of protest and a higher likelihood of repression in equilibrium. In particular in region III of Figure 2, the bad government introduces a judiciary with a relatively low degree of independence, reducing the equilibrium level of protest from $\pi = \alpha_g$ (in the absence of judiciary) to $\pi = \delta_b$ and increasing the likelihood with which it *successfully* represses the good opposition (see Corollary 8). In region IV, the bad government introduces a judiciary with a higher degree of independence, supporting an equilibrium in which the good government always represses the bad opposition and the public never protests.[9]

$\rho_g^B = \frac{q}{1-q}$ which is *less* than the level of repression $r_g = \frac{q}{1-q}\frac{1}{p}$ when there is a judiciary. Notably, absent the judiciary, the good government does not repress in equilibrium; in contrast, with the judicairy, the good government always represses the bad opposition.

[9]In some cases, the bad government prefers not to establish an independent judiciary that constrains repression; this happens when the opposition is very unpopular (region I) or when both opposition and government are popular but the government is relatively more popular than the opposition (region II).

5 CONCLUSION

Imprisonment, torture, killing, and disappearances are the staple of dictatorships. Yet, while some autocracies completely subdue their judiciary, others exhibit a degree of judicial independence as judges strike down government attempts to repress dissidents. This degree of judicial independence varies across countries. As discussed in the Introduction, in the Brazilian military regime, (security) courts struck down many government attempts to repress dissidents, but in Argentina, courts did little to protect dissidents (Pereira 2008). The degree of judicial independence can also vary over time. In Egypt, for example, Sadat came to power after Nasser's assassination by Islamists, at a time of deep unpopularity of the regime. Consistent with our analysis, judicial independence increased under Sadat. Even in the sphere of security, judges did not act as obedient agents of the executive. They continued to acquiesce to the repression of the Muslim Brotherhood, but protected other groups with more regularity. The Supreme Constitutional Court, created in 1979, began to rule against the government regularly, in some years striking down more laws than it upheld (Brown 2012).

Why do some dictatorships allow judicial independence, even though this independence enables judges to constrain the government's ability to repress dissidents? We suggest that by granting the courts the authority to constrain repression, the authoritarian regime manipulates public opinion about the merits of repression in its favor, thereby reducing backlash public protests and increasing the regime's stability. Our logic has two building blocks: an underlying mechanism and its implementation. The underlying mechanism is informational. By committing to repress legitimate dissidents less, authoritarian regimes shift public opinion against dissidents and in their own favor. When a government is less likely to repress legitimate dissidents, the public becomes more likely to believe that an act of coercion by the state is a "legitimate coercion" (Almond 1956; Mansbridge 2012, 2014). But how can the regime commit to restrain its repression of legitimate dissidents to implement this mechanism? The implementation of this informational mechanism is institutional: By establishing a partially independent judiciary the regime commits to constrain its own repression.

Our analysis suggests two main directions for future research. First, our informational mechanism takes public opinion about the regime and the dissidents as given, focusing on how it shapes the incentives for the regime to grant judicial independence. Second, our argument also takes as given that establishing an institution automatically makes it *functional*. As "democratization" is assumed to commit the elite to redistribute wealth in Acemoglu and Robinson (2001, 2006) or Boix (2003), our partially independent judiciary is assumed to commit the regime to restrict its repression of legitimate dissidents. However, as Hamilton points out in Federalist No. 78, the judiciary has a "natural feebleness" which places it in "continual jeopardy of being overpowered, awed, or influenced by its co-ordinate branches." The judiciary has neither the power of the purse nor the sword, "neither force nor will, but merely judgment." A dynamic extension of our analysis could address both of these issues simultaneously. In a dynamic setting, previous interactions endogenously influence current beliefs, allowing an analysis of the endogenous co-evolution of beliefs and judicial independence. Such a dynamic setting also allows one to investigate exactly how the judiciary generates commitment power for the ruler, by imposing direct or reputational costs on rulers who refuse to abide by judicial rulings. More broadly our paper points to the informational role of the judiciary in authoritarian regimes as a rich area for future study.

6 REFERENCES

Abrahamian, Ervand. 1982. *Iran Between Two Revolutions.* Princeton, NJ: Princeton University Press.

Abrahamian, Ervand. 1989. *The Iranian Mojahedin.* New Haven, CT: Yale University Press.

Acemoglu, Daron, and James A. Robinson. 2001. "A Theory of Political Transition." *American Economic Review* 91: 938-63.

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

Almond, Gabriel. 1956. "Comparative Political Systems." *Journal of Politics* 18: 391-409.

Alonso, Ricardo, and Niko Matouschek. 2008. "Optimal Delegation." *Review of Economic Studies* 75: 259-93.

Aytaç, Erdem, and Susan Stokes. 2019. *Why Bother?: Rethinking Participation in Elections and Protests.* New York, NY: Cambridge University Press.

Barbera, Salvador, and Matthew O. Jackson. 2016. "A Model of Protests, Revolution, and Information." Mimeo.

Boix, Carles. 2003. *Democracy and Redistribution.* New York: Cambridge University Press.

Boix, Carles, and Milan Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions and Power-Sharing in Dictatorships." *Journal of Politics* 75: 300-16.

Brown, Nathan J. 2012. *Egypt's Judges in a Revolutionary Age.* Carnegie Endowment for International Peace.

Bueno de Mesquita, Ethan. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review* 104: 446-66.

Callander, Steven, and Keith Krehbiel. 2014. "Gridlock and Delegation in a Changing World." *American Journal of Political Science* 58: 819-834.

Cheibub, Jose A., Jennifer Gandhi, and James R. Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143: 67-101.

Chen, Heng, and Wing Suen. 2016. "Falling Dominoes: A Theory of Rare Events and Crisis Contagion." *American Economic Journal: Microeconomics* 2016 8: 1-29.

Davenport, Christian. 2007a. "State Repression and Political Order." *Annual Review of Political Science* 10: 1-23.

Davenport, Christian. 2007b. *State Repression and the Domestic Democratic Peace.* New York, NY: Cambridge University Press.

Earl, Jennifer. 2011. "Political Repression: Iron Fists, Velvet Gloves, and Diffuse Control." *Annual Review Sociology* 37: 261-84.

Edmond, Chris. 2013. "Information Manipulation, Coordination and Regime Change." *Review of Economic Studies* 80: 1422-58.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103: 645-68.

Epperly, Brad. 2013. "The Provision of Insurance? Judicial Independence and the Post-Tenure Fate of Leaders." *Journal of Law and Courts* 1: 247-78.

Foucault, Michel. 2005 [1981]. *Hermeneutics of the Subject: Lectures at the College de France 1981-82*. Traslated by Graham Burchell. New York: Picador.

Francisco, Ronald. 2004. "After the Massacre: Mobilization in the Wake of Harsh Repression." *Mobilization* 9: 107-26 .

Fudenberg, Drew, and Jean Tirole. 2000. *Game Theory.* Cambridge, MA: MIT Press.

Fuller, Lon L. 1975. "Law as an Instrument of Social Control and Law as a Facilitation of Human Interaction." *BYU Law Review* 1: 89-96.

Gandhi, Jennifer. 2008. *Political Institutions Under Dictatorship.* New York, NY: Cambridge University Press.

Geddes, Barbara, Erica Frantz, and Joseph Wright. 2014. "Military Rule." *Annual Review of Political Science* 17: 147-62.

Ginsburg, Tom. 2003. *Judicial Review in New Democracies: Constitutional Courts in Asian Cases.* New York, NY: Cambridge Univ. Press

Ginsburg, Tom, and Tamir Moustafa. eds. 2008. *Rule by Law: The Politics of Courts in Authoritarian Regimes.* New York, NY: Cambridge University Press.

Helmke, Gretchen and Frances Rosenbluth. 2009. "Regimes and the Rule of Law: Judicial Independence in Comparative Perspective." *Annual Review of Political Science* 12: 345-66.

Hilbink, Lisa. 2012. "The Origins of Positive Judicial Independence." *World Politics* 64: 587-621.

Hirschl, Ran. 2004. *Towards Juristocracy: The Origins and Consequences of the New Constitutionalism.* Cambridge, MA: Harvard University Press.

Huang, Haifeng, Serra Boranbay-Akan, and Ling Huang. 2019. "Media, Protest Diffusion, and Authoritarian Resilience." *Political Science Research and Methods* 7: 23-42.

Ip, Eric C. 2012. "Judicial Review in China: A Positive Political Economy Analysis." *Review of Law & Economics* 8: 331-66.

Mansbridge, Jane. 2012. "On the Importance of Getting Things Done." *PS: Political Science & Politics* 45: 1-8.

Mansbridge, Jane. 2014. "What Is Political Science For?" *Perspectives on Politics* 12: 8-17.

Moustafa, Tamir. 2014. "Law and Courts in Authoritarian Regimes." *Annual Review of Law and Social Science* 10: 281-99.

Peerenboom, Randall. 2002. *China's Long March Toward Rule of Law.* Cambridge, UK: Cambridge University Press.

Pereira, Anthony W. 2008. "Of Judges and Generals: Security Courts under Authoritarian Regimes in Argentina, Brazil, and Chile." In *Rule by Law: The Politics of Courts in Authoritarian Regimes*, p. 23-57. Edited by Tom Ginsburg and Tamir Moustafa. New York: Cambridge University Press.

Persson, Torsten, and Guido Tabellini. 2009. "Democratic Capital: The Nexus of Political and Economic Change." *American Economic Journal: Macroeconomics* 1: 88-126.

Poe, Steven, and Neal Tate. 1994. "Repression of Human Rights to Personal Integrity in the 1980s: A Global Analysis." *American Political Science Review* 88: 853-72.

Rosberg, James H. 1995. "Roads to the Rule of Law: The Emergence of an Independent Judiciary in Contemporary Egypt.” PhD Thesis, Department of Political Science, MIT.

Rundlett, Ashlea, and Milan Svolik. 2016. “Deliver the Vote! Micromotives and Macrobehavior in Electoral Fraud.” *American Political Science Review* 110 (1): 180-97.

Shadmehr, Mehdi. 2019. “Investment in the Shadow of Conflict: Globalization, Capital Control, and State Repression.” *American Political Science Review. Forthcoming.*

Shadmehr, Mehdi, and Raphael Boleslavsky. 2019. “Institutions, Repression and the Spread of Protest.” Mimeo.

Shadmehr, Mehdi, and Dan Bernhardt. 2011. “Collective Action with Uncertain Payoffs: Coordination, Public Signals, and Punishment Dilemmas.” *American Political Science Review* 105: 829-51.

Shadmehr, Mehdi, and Dan Bernhardt. 2015. “State Censorship.” *American Economic Journal: Microeconomics* 7: 280-307.

Shariati, Ali. 1969 [1980]. *Ummat va Imāmat. Tihrān: Qalam.*

Shapiro, Martin. 1981. *Courts: A Comparative and Political Analysis.* Chicago, IL: University of Chicago Press.

Siegel, David. 2011. “When Does Repression Work? Collective Action in Social Networks.” *Journal of Politics* 73: 993-1010.

Svolik, Milan. 2012. *The Politics of Authoritarian Rule.* New York, NY: Cambridge Univ. Press.

Svolik, Milan. 2013. “Contracting on Violence: Moral Hazard in Authoritarian Repression and Military Intervention in Politics.” *Journal of Conflict Resolution* 57: 765-94.

Tyson, Scott, and Alastair Smith. 2018. “Dual-Layered Coordination and Political Instability: Repression, Cooptation, and the Role of Information.” *Journal of Politics* 80: 44-58.

Verner, Joel. 1984. “The Independence of Supreme Courts in Latin America: A Review of the Literature.” *Journal of Latin American Studies* 16: 463-506.

World Justice Project. Rule of Law Index 2014-2018. Available at https://worldjusticeproject.org

Zhang, Taisu and Tom Ginsburg. 2019. “Legality in Contemporary Chinese Politics.” *Virginia Journal of International Law. Forthcoming.*

Online Appendix: Proofs and Extensions

A PROOFS

We provide the proofs and extensions for a more general payoff structure of the game. In particular, under the status quo, the public's payoff is 0 if the government is good, and $-\beta \leq 0$ if the government is bad. If a good reform is implemented, the public's payoff rises to $\beta_g > 0$. If a bad reform is implemented the public's payoff falls to $-\beta_b < 0$, with $-\beta_b < -\beta$. The public's payoffs in the text is a special case of this generalized payoffs, where $\beta = 0$ and $\beta_g = \beta_b = b$.

We omit the proof of Proposition 1 because it corresponds to Proposition 2 in Shadmehr and Boleslavsky (2019) and its proof appears in that paper. Here, we provide the proofs for Lemma 1, Proposition 2 and Proposition 3 of this paper. We also present and prove a theorem regarding equilibrium selection. We omit the proofs of Corollaries 1-3 because they follow immediately from Propositions 2 and 3.

A.1 LEMMA 1

Lemma 1 *Given the strategy of the bad government* (r_b, r_g) *and the strategy of the good government* ρ_b^G*, the public's best response is:*

$$\pi(r_b, r_g; \rho_b^G) = \begin{cases} 1 & \text{if} \quad F(r_b, r_g) > K\rho_b^G \\ [0,1] & \text{if} \quad F(r_b, r_g) = K\rho_b^G \\ 0 & \text{if} \quad F(r_b, r_g) < K\rho_b^G \end{cases}$$

where $F(r_b, r_g) \equiv (1-q)(\beta_g + \beta)r_g - q(\beta_b - \beta)r_b$ *and* $K \equiv \frac{1-p}{p}q\beta_b$.

Proof of Lemma 1. Let p' be the public's posterior probability that the government is bad, and let q' be the public's posterior probability that the opposition is bad. The public's best response is:

$$\pi = \begin{cases} 1 & \text{if} \quad \beta_g(1-q') - \beta_b q' > -\beta p' \\ [0,1] & \text{if} \quad \beta_g(1-q') - \beta_b q' = -\beta p' \\ 0 & \text{if} \quad \beta_g(1-q') - \beta_b q' < -\beta p', \end{cases} \tag{1}$$

where the public's updated beliefs p' and q' depend on the government's strategy in equilibrium.

Moreover, recall that a good government never represses a good opposition, $\rho_g^G = 0$. Thus, Bayes rule implies:

$$(2) \qquad p' = \frac{p[qr_b + (1-q)r_g]}{p[qr_b + (1-q)r_g] + (1-p)q\rho_b^G} \qquad q' = \frac{q[(1-p)\rho_b^G + pr_b]}{q[(1-p)\rho_b^G + pr_b] + p(1-q)r_b}.$$

Substituting from the public's posterior belief, equation (2), into equation (1) gives the result. ■

A.2 PROPOSITION 2

Proposition 2 *Suppose the bad government can commit to a strategy* (r_b, r_g). *There exist an increasing curve,* $q_1(p) = \frac{(\beta+\beta_g)p}{\beta_b+\beta_g p}$, *and three constants,* $q_2 = \frac{\beta+\beta_g}{\beta_b+\beta_g}$, $q^* = q_2(1 - \frac{\delta_b}{\alpha_g})$, *and* $p^* = \frac{\beta_b(\alpha_g-\delta_b)}{\beta_b\alpha_g+\delta_b\beta_g}$, *with* $0 < q_1(p), q^* < q_2 < 1$, *such that, in equilibrium:*

1. *If* $q_2 < q$, *then the strategies are identical to the Low Protest equilibrium (region I).*
2. *If* $q_1(p) < q < q^*$, *strategies are identical to the Intermediate Protest equilibrium (region II).*
3. *If* $q < q_1(p)$ *and* $p < p^*$, *then* $r_b = \rho_b^G = 1$, $r_g = \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta+\beta_g)}$, *and* $\pi = \delta_b$ *(region III).*
4. *Otherwise,* $r_b = \rho_b^G = 1$, $r_g = \frac{q}{1-q}\frac{\beta_b-\beta}{\beta+\beta_g}$, *and* $\pi = 0$ *(region IV).*

First, we present four Lemmas, which we will use to prove Proposition 2.

Lemma 2 *Fix the bad government's strategy* (r_b, r_g). *The following characterizes the equilibria of the subgame:*

1. *If* $F(r_b, r_g) < 0$, *then the unique equilibrium of the subgame has* $\pi = 0$ *and* $\rho_b^G = 1$.
2. *If* $F(r_b, r_g) > K$, *then the unique equilibrium of the subgame has* $\pi = 1$ *and* $\rho_b^G = 0$.
3. *If* $0 < F(r_b, r_g) < K$, *then the equilibria described in (1) and (2) both exist. In addition, there is an equilibrium of the subgame in which* $\pi = \delta_b$ *and* $\rho_b^G = pF(r_b, r_g)/((1-p)q\beta_b)$.
4. *If* $F(r_b, r_g) = 0$, *then the equilibrium described in (1) exists. In addition, a continuum of equilibria exist in which* $\pi \in [\delta_b, 1]$ *and* $\rho_b^G = 0$.
5. *If* $F(r_b, r_g) = K$, *then the equilibrium described in (2) exists. In addition, a continuum of equilibria exist in which* $\pi \in [0, \delta_b]$ *and* $\rho_b^G = 1$.

Proof. A government's expected payoff from repression is $1 - \pi$. However, if the government concedes, its payoff depends on both its type and the opposition's type. Because the good government prefers a good reform to the status quo ($\delta_g > 0$), it always concedes to the good opposition, $\rho_g^G = 0$. Moreover, recall that the good government's payoff from conceding to a bad opposition is $1 - \delta_b$. Thus, the best response of the good government is:

$$\rho_b^G = \begin{cases} 1 & ;\ \pi < \delta_b \\ [0,1] & ;\ \pi = \delta_b \\ 0 & ;\ \pi > \delta_b. \end{cases} \tag{3}$$

(1) If (r_b, r_g) is such that $F(r_b, r_g) < 0$, then $F(r_b, r_g) < K\rho_b^G$ for any $\rho_b^G \in [0,1]$. Hence, in the subgame $\pi = 0$ for any $\rho_b^G \in [0,1]$. Because $\pi = 0$, equation (3) requires that $\rho_b^G = 1$.

(2) If (r_b, r_g) is such that $F(r_b, r_g) > K$, then $F(r_b, r_g) > K\rho_b^G$ for any $\rho_b^G \in [0,1]$. Hence, in the subgame $\pi = 1$ for any $\rho_b^G \in [0,1]$. Because $\pi = 1$, equation (3) requires that $\rho_b^G = 0$.

(3) Suppose (r_b, r_g) is such that $F(r_b, r_g) \in (0, K)$. If $\rho_b^G = 1$, then $\pi = 0$ from Lemma 1; conversely, $\rho_b^G = 1$ is consistent with equation (3) when $\pi = 0$. If $\rho_b^G = 0$, then $\pi = 1$ from Lemma 1; conversely, $\rho_b^G = 0$ is consistent with equation (3) when $\pi = 1$. If $\rho_b^G \in (0,1)$, then equation (3) implies that $\pi = \delta_b$; conversely, Lemma 1 implies that if $\rho_b^G = F(r_b, r_g)/K \in (0,1)$, then $\pi = \delta_b$ is a best response.

(4) Suppose (r_b, r_g) is such that $F(r_b, r_g) = 0$. If $\rho_b^G \in (0,1]$, then $\pi = 0$ from Lemma 1; conversely, when $\pi = 0$, equation (3) implies that $\rho_b^G = 1$. If $\rho_b^G = 0$, then Lemma 1 implies that $\pi = [0,1]$; conversely, equation (3) implies that $\rho_b^G = 0$ is a best response if and only if $\pi \in [\delta_b, 1]$.

(5) Suppose (r_b, r_g) is such that $F(r_b, r_g) = K$. If $\rho_b^G \in [0,1)$, then $\pi = 1$ from Lemma 1; conversely, when $\pi = 1$, equation (3) implies that $\rho_b^G = 0$. If $\rho_b^G = 1$, then Lemma 1 implies that $\pi = [0,1]$; conversely, equation (3) implies that $\rho_b^G = 1$ is a best respose if and only if $\pi \in [0, \delta_b]$. ■

Lemma 3 *If $q > q_2$, then there is a unique equilibrium in which $r_b = r_g = \rho_b^G = 1$ and $\pi = 0$.*

Proof. If $q > q_2$, then $F(1,1) < 0$, which implies $\pi = 0$ from Lemma 2. Thus, $\rho_b^G = 1$. The bad government's payoff is $B(1,1,0) = 1$ which is strictly larger than $B(r_b, r_g, \pi)$ for any $(r_b, r_g) \neq (1,1)$. ■

Equilibrium Selection. Next, we impose the equilibrium selection ES1, and make the observation ES2:

ES1. If $F(r_b, r_g) = K$, then in the equilibrium of the subgame we have $\pi = \delta_b$ and $\rho_b^G = 1$.

ES2. If $F(r_b, r_g) = 0$, then in the equilibrium of the subgame we have $\pi = 0$ and $\rho_b^G = 1$.

ES1 is an equilibrium selection that is justified using a refinement similar to trembling hand in Theorem 1. ES2 is justified in Lemma 6.

Lemma 4 *1. If $q < q_2$, then $R_0 \equiv (1, \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta})$ is the unique strategy that maximizes the bad government's expected payoff among all (r_b, r_g) for which $F(r_b, r_g) \leq 0$, and the associated payoff is $B_0 \equiv 1 - \alpha_g(1 - q/q_2)$.*

2. If $q_1(p) \leq q < q_2$, then $R_1 \equiv (1, 1)$ is the unique strategy that maximizes the bad government's expected payoff among all (r_b, r_g) for which $0 < F(r_b, r_g) \leq K$, and the associated payoff is $B_1 \equiv 1 - \delta_b$.

3. If $q < q_1(p)$, then $R_2 \equiv (1, \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta_g + \beta)})$ is the unique strategy that maximizes the bad government's expected payoff among (r_b, r_g) for which $0 < F(r_b, r_g) \leq K$, and the associated payoff is:

$$B_2 \equiv 1 - \alpha_g \left(1 - \frac{q}{q_2}\right) + \frac{q(\beta_b(\alpha_g - \delta_b) - p(\beta_b \alpha_g + \delta_b \beta_g)}{p(\beta + \beta_g)}.$$

Proof. Let $B(r_b, r_g, \pi)$ be the bad government's payoff, if it remains in power, from (r_b, r_g) that induces protest probability π in the equilibrium of the subgame:

$$(4) \qquad B(r_b, r_g, \pi) = q\ [r_b(1-\pi) + (1-r_b)(1-\alpha_b)] + (1-q)\ [r_g(1-\pi) + (1-r_g)(1-\alpha_g)],$$

so that the bad government's ex-ante payoff is $pB(r_b, r_g, \pi)$.

1. If $F(r_b, r_g) \leq 0$ and $q < q_2$, then $\pi = 0$ in the equilibrium of the subgame. From (4), the payoff of such a strategy is:

$$B(r_b, r_g, 0) = q(r_b + (1-r_b)(1-\alpha_b)) + (1-q)(r_g + (1-r_g)(1-\alpha_g)).$$

Thus, the government's problem becomes:

$$\max_{(r_b, r_g) \in [0,1]^2} B(r_b, r_g, 0) \ \text{ s.t. } \ F(r_b, r_g) \leq 0.$$

$B(r_b, r_g, 0)$ is increasing in both r_b and r_g. Because $F(r_b, r_g)$ is decreasing in r_b, we must have $r_b = 1$ at the optimum. Because $q < q_2$, $F(1, 1) > 0$, and hence $r_b = r_g = 1$ is not feasible. Because

$F(r_b, r_g)$ is increasing in r_g, we must have $F(1, r_g) = 0$ at the optimum r_g. Finally, $F(1, r_g) = 0$ implies $r_g = \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}$. B_0 is derived from substituting $(r_b, r_g) = (1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta})$ into $B(r_b, r_g, 0)$.

If $0 < F(r_b, r_g) \le K$, then $\pi = \delta_b$ in the equilibrium of the subgame. From (4), the bad government's payoff from any such strategy is:

$$B(r_b, r_g, \delta_b) = q\ [r_b(1-\delta_b) + (1-r_b)(1-\alpha_b)] + (1-q)\ [r_g(1-\delta_b) + (1-r_g)(1-\alpha_g)].$$

Thus, the government's problem in parts 2 and 3 becomes:

$$\max_{(r_b,r_g)\in[0,1]^2} B(r_b, r_g, \delta_b)\ \ s.t.\ \ 0 < F(r_b, r_g) \le K.$$

Because $\delta_b < \alpha_g < \alpha_b$, $B(r_b, r_g, \delta_b)$ is increasing in both r_b and r_g.

2. If $q_1(p) \le q < q_2$, then $0 < F(1,1) \le K$, and hence, $r_b = r_g = 1$ is feasible. Because $B(r_b, r_g, \delta_b)$ is increasing in r_b and r_g, $r_b = r_g = 1$ is the bad government's optimal choice. B_1 is derived by substituting $(r_b, r_g) = (1,1)$ into $B(r_b, r_g, \delta_b)$.

3. When $q < q_1(p)$, $F(1,1) > K$, and hence the constraint $F(r_b, r_g) = K$ binds. Because F is decreasing in r_b, we must have $r_b = 1$ at the optimum. Then, the optimal r_g is derived from $F(r_b = 1, r_g) = K$. ■

Lemma 5 *In equilibrium, the payoff of selecting any (r_b, r_g) such that $F(r_b, r_g) > K$ is smaller than B_0.*

Proof. If $F(r_b, r_g) > K$, then $\pi = 1$, and hence $B(r_b, r_g, 1) = q(1-r_b)(1-\alpha_b) + (1-q)(1-r_g)(1-\alpha_g)$. The bad government can benefit by deviating to $(r_b, r_g) = (0,0)$, so that $F(r_b, r_g) = 0$, and its expected payoff becomes $B(0,0,0) = q(1-\alpha_b) + (1-q)(1-\alpha_g) > B(r_b, r_g, 1)$. Because $F(0,0) = 0$, but $R_0 \neq (0,0)$ is optimal among $F(r_b, r_g) \le 0$ it must be that $B(0,0,0) < B_0$. Summarizing, If $F(r_b, r_g) > K$, then $B(r_b, r_g, 1) < B(0,0,0) < B_0$. ■

Proof of Proposition 2. Lemma 3 establishes part 1. Thus, we focus on $q < q_2$ in the rest of the proof.

Suppose $q_1(p) < q < q_2$. If $B_0 > B_1$, then Lemma 4 implies that R_0 dominates any strategy for which $F(r_b, r_g) \le K$, and Lemma 5 implies that R_0 dominates any strategy for which $F(r_b, r_g) > K$. Hence, if $B_0 > B_1$, then $R_0 = (1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta})$ dominates all $(r_b, r_g) \in [0,1]^2$, and it is the bad government's equilibrium choice. If $B_1 > B_0$, then Lemma 4 implies that R_1 dominates any strategy for which $F(r_b, r_g) \leq K$. Because $B_1 > B_0$, Lemma 5 implies that R_1 dominates any strategy for which $F(r_b, r_g) > K$. Hence, if $B_1 > B_0$ then $R_1 = (1,1)$ dominates all $(r_b, r_g) \in [0,1]^2$, and it is the bad government's equilibrium choice. Using B_0 and B_1 from Lemma 4,

$$B_1 > B_0 \text{ if and only if } q < q^* \equiv q_2\left(1 - \frac{\delta_b}{\alpha_g}\right).$$

Thus, if $q_1(p) < q < q^*$, in equilibrium, $(r_b, r_g) = (1,1)$, $\rho_b^G = \frac{F(1,1)}{K} \in (0,1)$, and $\pi = \delta_b$. If $q^* < q < q_2$, in equilibrium, $(r_b, r_g) = (1, \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta})$, $\rho_b^G = 1$, and $\pi = 0$.

Suppose $q < q_1(p)$. If $B_0 > B_2$, then Lemma 4 implies that R_0 dominates any strategy for which $F(r_b, r_g) \leq K$, and Lemma 5 implies that R_0 dominates any strategy for which $F(r_b, r_g) > K$. Hence, if $B_0 > B_2$, then $R_0 = (1, \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta})$ dominates all $(r_b, r_g) \in [0,1]^2$, and it is the bad government's equilibrium choice. If $B_2 > B_0$, then Lemma 4 implies that R_2 dominates any strategy for which $F(r_b, r_g) \leq K$. Because $B_2 > B_0$, Lemma 5 implies that R_2 dominates any strategy for which $F(r_b, r_g) > K$. Hence, if $B_2 > B_0$ then $R_2 = (1, \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta_g + \beta)})$ dominates all $(r_b, r_g) \in [0,1]^2$, and it is the bad government's equilibrium choice. Using B_0 and B_2 from Lemma 4,

$$B_2 > B_0 \text{ if and only if } p < p^* \equiv \frac{\beta_b(\alpha_g - \delta_b)}{\beta_b \alpha_g + \delta_b \beta_g}.$$

Thus, if $p < p^*$ *and* $q < q_1(p)$, in equilibrium, $(r_b, r_g) = (1, \frac{p}{1-p}\frac{\beta_b - p\beta}{p(\beta + \beta_g)})$, $\rho_b^G = 1$, and $\pi = \delta_b$. If $q < q_1(p)$ *and* $p > p^*$, in equilibrium, $(r_b, r_g) = (1, \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta})$, $\rho_b^G = 1$, and $\pi = 0$. ■

A.2.1 Equilibrium Selection

In this section, we present two results which justify our equilibrium selections.

ES1. From Lemma 2 point 5, when $F(r_b, r_g) = K$, a continuum of equilibria are possible in the subgame: any $\pi \in [0, \delta_b]$ can be part of the equilibrium of the subgame. In the text, we focus on the equilibrium of the subgame in which $\pi = \delta_b$. Here, we show that whenever the bad government's equilibrium strategy (r_b, r_g) has $F(r_b, r_g) = K$, the subgame with $\pi = \delta_b$ is uniquely selected by a simple refinement. In particular, we introduce stochastic shocks to the bad government's strategy, showing that as the support of the distribution of shocks vanishes, the equilibrium with the shocks converges to the one in which $\pi = \delta_b$.

Suppose that when the bad government commits to a strategy (r_b, r_g), the probability with which the type i opposition is actually repressed is a random variable $R(r_i) \equiv \max\{\min\{r_i + \nu_i, 1\}, 0\}$, where ν_i's are iid continuous random variables with support $[-\epsilon, \epsilon]$. Let $\hat{r}_i$ be the realization of the random variable $R(r_i)$. Let function $\pi^*(\hat{r}_b, \hat{r}_g)$ represent the protest probability in the equilibrium of the subgame following $(\hat{r}_b, \hat{r}_g)$:

$$\pi^*(\hat{r}_b, \hat{r}_g) = \begin{cases} 1 & \text{if} \quad F(\hat{r}_b, \hat{r}_g) > K \\ \delta_b & \text{if} \quad 0 < F(\hat{r}_b, \hat{r}_g) < K \\ 0 & \text{if} \quad F(\hat{r}_b, \hat{r}_g) < 0 \end{cases}$$

From Lemma 2, if $F(\hat{r}_b, \hat{r}_g) > K$, then $\pi^*(\hat{r}_b, \hat{r}_g) = 1$, and if $F(\hat{r}_b, \hat{r}_g) < 0$, then $\pi^*(\hat{r}_b, \hat{r}_g) = 0$. In Proposition 2, we focus on the case in which $0 < F(\hat{r}_b, \hat{r}_g) < K$ implies $\pi = \delta_b$, one of the three options that follows from Lemma 2 (the others are considered in Propositions 5, 6). Because ν_is are independent and have no mass points, for any choice of the bad government (r_b, r_g), the probability that the realizations are such that $F(\hat{r}_b, \hat{r}_g) = K$ or $F(\hat{r}_b, \hat{r}_g) = 0$ is zero.

For a given ϵ, the bad government's problem is:

$$\max_{(r_b, r_g)} \; E[\; B(R(r_b), R(r_g), \pi^*(R(r_b), R(r_g)))\;],$$

where the expectation is over (ν_b, ν_g). Let $(r_b^*(\epsilon), r_g^*(\epsilon))$ be the maximand(s) and $B^*(\epsilon)$ be the maximum value. From Proposition 2, when $q < q_1(p)$ and $p < \frac{\beta_b(\alpha_g - \delta_b)}{\beta_b \alpha_g + \delta_b \beta_g}$, absent trembles, the bad government's equilibrium choice is $(r_b^*, r_g^*) \equiv (1, \frac{q}{1-q} \frac{\beta_b - p\beta}{p(\beta + \beta_g)})$, yielding the bad government's payoff $B^* \equiv B(r_b^*, r_g^*, \delta_b)$. This is the only instance in which the bad government's equilibrium choice is such that $F(r_b, r_g) = K$.

Theorem 1 *Suppose $q < q_1(p)$ and $p < \frac{\beta_b(\alpha_g - \delta_b)}{\beta_b \alpha_g + \delta_b \beta_g}$, so that the bad government's choice in the absence of stochastic shocks is (r_b^*, r_g^*). As the support of the distribution of the shocks shrinks to zero:*

1. *The bad government's payoff converges to its payoff in the absence of shocks:* $\lim_{\epsilon \to 0} B^*(\epsilon) = B^*$

2. *The protest probability converges to δ_b:* $\lim_{\epsilon \to 0} \Pr\{\pi^*(R(r_b^*(\epsilon)), R(r_g^*(\epsilon))) = \delta_b\} = 1.$

3. *The bad government's strategy converges to (r_b^*, r_g^*):* $\lim_{\epsilon \to 0} r_i^*(\epsilon) = r_i^*$ *for $i \in \{b, g\}$.*

Proof. (1) Because (r_b^*, r_g^*) is optimal for the bad government in the absence of trembles, $B^* > B(\hat{r}_b, \hat{r}_g, \pi^*(\hat{r}_b, \hat{r}_g))$ for all possible realizations $(\hat{r}_b, \hat{r}_g) \neq (r_b^*, r_g^*)$, and hence

$$B^* > B^*(\epsilon). \tag{5}$$

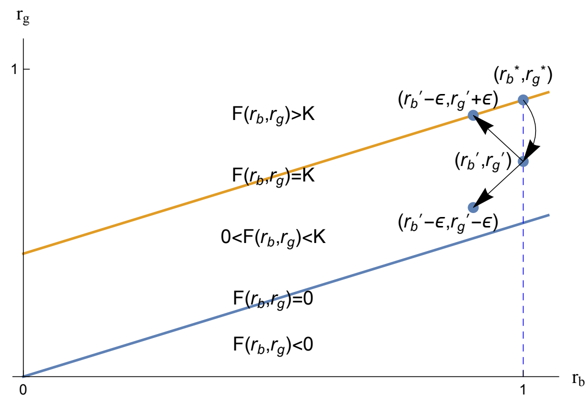

Figure 4: The nature of the deviation from (r_b^*, r_g^*) to (r_b', r_g').

Recall that $(r_b^*, r_g^*) = (1, \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta+\beta_g)})$, and consider an alternative strategy for the bad government:

$$(r_b', r_g') \equiv (r_b^*, r_g^* - \epsilon(1 + \frac{q}{1-q}\frac{\beta_b - \beta}{\beta + \beta_g})),$$

so that $F(r_b' - \epsilon, r_g' + \epsilon) = K$ (see Figure 4). For sufficiently small ϵ, the monotonicity properties of F imply:

$$0 < F(1, r_g' - \epsilon) < F(R(r_b'), R(r_g')) < F(r_b' - \epsilon, r_g' + \epsilon) = K.$$

That is, if the bad government chooses (r_b', r_g'), then for any realization of shocks $0 < F(R(r_b'), R(r_g')) < K$, and hence $\Pr\{\pi^*(R(r_b'), R(r_g')) = \delta_b\} = 1$. Thus, the government's expected payoff from (r_b', r_g') is $E[B(R(r_b'), R(r_g'), \delta_b)]$. Let $k \equiv 2 + \frac{q}{1-q}\frac{\beta_b - \beta}{\beta + \beta_g}$, so that $B(r_b^* - \epsilon, r_g^* - k\epsilon, \delta_b) = B(r_b' - \epsilon, r_g' - \epsilon, \delta_b)$. Then,

$$B(r_b^* - \epsilon, r_g^* - k\epsilon, \delta_b) = B(r_b' - \epsilon, r_g' - \epsilon, \delta_b) < E[B(R_b(r_b'), R_g(r_g'), \delta_b)] \le B^*(\epsilon) < B^*.$$

The first inequality follows from monotonicity properties of B, the second inequality follows from optimality of $B^*(\epsilon)$, and the third is (5). From continuity of B, $\lim_{\epsilon\to 0} B(r_b^* - \epsilon, r_g^* - k\epsilon, \delta_b) = B^*$, and hence $\lim_{\epsilon\to 0} B^*(\epsilon) = B^*$.

(2) For simplicity, denote random variable $R(r_i^*(\epsilon))$ by R_i^*. When the bad government chooses $(r_b^*(\epsilon), r_g^*(\epsilon))$ its payoff is:

$$\begin{aligned} B^*(\epsilon) = {} & \Pr\{\pi^*(R_b^*, R_g^*) = 1\} E[B(R_b^*, R_g^*, 1) | \pi^*(R_b^*, R_g^*) = 1] \\ & + \Pr\{\pi^*(R_b^*, R_g^*) = \delta_b\} E[B(R_b^*, R_g^*, \delta_b) | \pi^*(R_b^*, R_g^*) = \delta_b] \\ & + \Pr\{\pi^*(R_b^*, R_g^*) = 0\} E[B(R_b^*, R_g^*, 0) | \pi^*(R_b^*, R_g^*) = 0]. \end{aligned} \tag{6}$$

If $\pi^*(\hat{r}_b, \hat{r}_g) = 1$, then $B(\hat{r}_b, \hat{r}_g, 1) < B(0, 0, 1)$; If $\pi^*(\hat{r}_b, \hat{r}_g) = \delta_b$, then $B(\hat{r}_b, \hat{r}_g, \delta_b) < B(r_b^*, r_b^*, \delta_b) =$

B^*; If $\pi^*(\hat{r}_b, \hat{r}_g) = 0$, then $B(\hat{r}_b, \hat{r}_g, 0) < B(1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}, 0) < B^*$. Substituting these into (6) yields:

$$B^*(\epsilon) < \Pr\{\pi^*(R_b^*, R_g^*) = 1\}B(0,0,1) + \Pr\{\pi^*(R_b^*, R_g^*) = \delta_b\}B^* + \Pr\{\pi^*(R_b^*, R_g^*) = 0\}B(1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}, 0).$$

Rearranging the right hand side yields:

$$B^*(\epsilon) < B^* - \Pr\{\pi^*(R_b^*, R_g^*) = 1\}(B^* - B(0,0,1)) - \Pr\{\pi^*(R_b^*, R_g^*) = 0\}(B^* - B(1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}, 0)).$$

Taking the limit of both sides yields:

$$\lim_{\epsilon\to 0} B(\epsilon) \leq B^* - (B^* - B(0,0,1))\lim_{\epsilon\to 0}\Pr\{\pi^*(R_b^*, R_g^*) = 1\} - \lim_{\epsilon\to 0}\Pr\{\pi^*(R_b^*, R_g^*) = 0\}(B^* - B(1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}, 0)).$$

From part (1), $\lim_{\epsilon\to 0} B(\epsilon) = B^*$. Because $B(0,0,1) < B^*$ and $B(1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}) < B^*$, we must have:

$$\lim_{\epsilon\to 0}\Pr\{\pi^*(R_b^*, R_g^*) = 1\} = 0 \qquad \lim_{\epsilon\to 0}\Pr\{\pi^*(R_b^*, R_g^*) = 0\} = 0.$$

From part (1), $\lim_{\epsilon\to 0} B^*(\epsilon) = B^*$. From part (2), $\lim_{\epsilon\to 0}\Pr\{\pi^*(R_b^*, R_g^*) = \delta_b\} = 1$. Thus, (6) implies $\lim_{\epsilon\to 0} E[B(R_b^*, R_g^*, \delta_b)|\pi^*(R_b^*, R_g^*) = \delta_b] = B^*$. By continuity, $\lim_{\epsilon\to 0} r_i^*(\epsilon) = r_i^*$. ■

ES2 is justified by the following lemma.

Lemma 6 *Suppose that (1) $q_2(1-\delta_b/\alpha_g) < q < q_2$, so that $B_0 > B_1$, or (2) $q < q_1(p)$ and $p > p^*$, so that $B_0 > B_2$. Moreover, suppose ES2 is violated, so that in the subgame following the bad government's choice of (r_b, r_g) for which $F(r_b, r_g) = 0$ the protest probability is $\pi > 0$. Under these conditions no equilibrium exists.*

Proof. Consider the bad government's maximization problem over the region where $F(r_b, r_g) \leq 0$:

$$(7) \quad \max_{(r_b,r_g)\in[0,1]^2} B(r_b, r_g, \pi(r_b, r_g)) \quad \text{subject to} \quad F(r_b, r_g) \leq 0, \quad \text{and } \pi(r_b, r_g) = \pi \text{ if } F(r_b, r_g) = 0.$$

Lemma 4 implies that if $\pi = 0$, then the solution is R_0, generating payoff B_0. However, if $\pi > 0$, choosing R_0 does not deliver payoff B_0, because the bad government's payoff function is decreasing in the protest probability and $\pi > 0$ at R_0. Thus, the government's payoff cannot exceed B_0. Consider the choice of $(r_b', r_g') = (1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta} - \epsilon)$ for small ϵ. Because $F(1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}) = 0$ and F is increasing in r_g, $F(r_b', r_g') < 0$. Hence, following the bad government's choice of (r_b', r_g'), the protest probability is zero. Therefore, the bad government's expected payoff of this choice is $B(r_b', r_g', 0) = B_0 - \epsilon\alpha_g(1-q)$. Therefore, as ϵ approaches zero, the bad government's payoff approaches B_0, but no strategy delivers payoff B_0.

From Lemma 5, B_0 is larger than the payoff of any (r_b, r_g) for which $F(r_b, r_g) > K$. Moreover, Proposition 2 implies that, under conditions (1) or (2), B_0 is larger than the payoff of any (r_b, r_g) for which $0 < F(r_b, r_g) \leq K$. We showed a sequence of strategies with $F(r_b, r_g) < 0$ delivers a payoff that approaches B_0. Hence, optimization (7) has no solution, and no equilibrium exists. ■

A.3 PROPOSITION 3

Proposition 3 *Partially independent judiciaries can generate the same outcome as when the regime can fully commit to a precise repression strategy (Benchmark 2).*

Terminology. To ease the exposition, we introduce the following terminology. We refer to the game in which the bad government can commit to a repression strategy (r_b, r_g) as *the game with precise commitment*. The game in which the government's attempt to repress the good opposition is blocked with an exogenous probability is called *the game with exogenous C*. We refer to the game in which the bad government can ex ante choose this probability as *the game with endogenous C*. With this terminology, Proposition 3 concerns the equilibria of the game with endogenous C. The equilibria of the game with precise commitment are presented in Proposition 2 and the equilibria of the game with exogenous C are presented in Shadmehr and Boleslavsky (2019, Proposition 3). An equilibrium of the game with precise commitment is *outcome-equivalent* to an equilibrium of the game with exogenous C or an equilibrium of the game with endogenous C if the probability that each type of government successfully represses each type of opposition, and the probability that the public protests following repression are identical.

A.3.1 Equilibrium Selection

The subgame following the bad government's choice of C is identical to the game with exogenous C, and the equilibria of this subgame correspond to the equilibria of the game characterized in Proposition 3 of Shadmehr and Boleslavsky (2019). We adapt and present their proposition here:

Proposition 4 ***(Shadmehr and Boleslavsky 2019)*** *Suppose that the government's attempt to repress a good opposition is blocked with an exogenous probability $C \in (0,1)$. In equilibrium, $\rho_g^G = 0$ and $\rho_b^B = 1$. For a given C, there exists an increasing function $g(p) = \frac{(1-C)(1+p)}{1+(1-C)p} \in (0,1)$ such that:*

- *High Protest Equilibrium with Exogenous C: If $q < g(p)q_1(p)$, then a unique equilibrium exists in which $\rho_b^G = 0$, $\rho_g^B = \frac{\beta_b - \beta}{\beta_g + \beta}\frac{q}{1-q}\frac{1}{1-C}$, and $\pi = \alpha_g$.*

- *Low Protest Equilibrium with Exogenous C: If $q > g(1)q_2$, then a unique equilibrium exists in which $\rho_b^G = 1$, $\rho_g^B = 1$, and $\pi = 0$.*

- *Intermediate Protest Equilibrium with Exogenous C: If $g(p)\ q_1(p) \leq q \leq g(1)\ q_2$, then, in addition to the above two equilibria, there exists an equilibrium in which:*

$$\rho_b^G = \frac{p}{1-p}\Big(\frac{(\beta+\beta_g)-(\beta_b+\beta_g)q}{\beta_b q} - \frac{1-q}{q}\frac{\beta+\beta_g}{\beta_b}C\Big),\ \rho_g^B = 1,\ \text{and } \pi = \delta_b.$$

For $q \in [q_1(p)g(p), q_2 g(1)]$, multiple equilibria exist in the subgame. We consider the following selection in these cases: if $q = q_2 g(1)$, then we select the *low protest equilibrium with exogenous C*, and if $q \in [q_1(p)g(p), q_2 g(1))$, then we select the *intermediate protest equilibrium with exogenous C*. This equilibrium selection is consistent with the equilibrium selection in Benchmark 2 in the following sense. In Benchmark 2, equilibrium multiplicity arises when the bad government's choice of (r_b, r_g) satisfies $F(r_b, r_g) \in [0, K]$, and we select the equilibrium with $\pi = \delta_b$ when $F(r_b, r_g) \in (0, K]$, and the equilibrium with $\pi = 0$ when $F(r_b, r_g) = 0$. When multiple equilibria arise in the subgame following the government's choice of judicial independence, one of those equilibria features $\pi = \alpha_g$. This equilibrium is inconsistent with our preceding selection in which $\pi = \delta_b$ or $\pi = 0$, and hence we do not select it. In the other equilibria, the bad opposition is repressed with probability 1, and the good opposition is repressed successfully with probability $1-C$. Thus, we select the equilibrium with $\pi = 0$ whenever $F(1, 1-C) = 0$ and the equilibrium with $\pi = \delta_b$ whenever $F(1, 1-C) \in (0, K]$.

A.3.2 Proof

Lemma 7 *Fix a choice of C in the game with endogenous C. There exists an (r_b, r_g) such that the bad government's payoff of selecting (r_b, r_g) in the game with precise commitment is weakly greater than the bad government's payoff of selecting C in the game with endogenous C.*

Proof. We divide the proof into three cases: $q \in (0, q_1(p)g(p))$, $q \in [q_1(p)g(p), q_2 g(1))$, and $q \in [q_2 g(1), 1)$, where we recognize that $g(p)$ depends on C.

Case I. Suppose C is such that $q < q_1(p)g(p)$, that is,

$$q < q_1(p)\ \left(1 - C\frac{\beta_b - p\beta}{\beta_b - p\beta + (1-C)p(\beta+\beta_g)}\right) \Leftrightarrow C < 1 - \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta+\beta_g)}.$$

Following this choice of C, from Proposition 4, the *high protest equilibrium with exogenous C* is the unique equilibrium of the subgame, and it generates an expected payoff of $\hat{B}(C) = 1 - \alpha_g$. Suppose that in the game with precise commitment, the bad government chooses $(r_b, r_g) = (1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta})$, so that $F(r_b, r_g) = 0$. Then, in the subgame following $(r_b, r_g) = (1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta})$, we have $\pi = 0$. Hence, from equation (4), the bad government's expected payoff is $B(1, \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}, 0) = q + (1-q)\ [\frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta} + (1 - \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta})(1-\alpha_g)] = 1 - \alpha_g + \frac{q}{q_2}\alpha_g > \hat{B}(C)$.

Case II. Suppose C is such that $q_1(p)g(p) \leq q < q_2 g(1)$, that is,

$$q_1(p)\ \left(1 - C\frac{\beta_b - p\beta}{\beta_b - p\beta + (1-C)p(\beta+\beta_g)}\right) \leq q < q_2\ \left(1 - C\frac{\beta_b - \beta}{\beta_b - \beta + (1-C)(\beta+\beta_g)}\right) \Leftrightarrow$$
$$1 - \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta+\beta_g)} \leq C < 1 - \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}.$$

Following this choice of C, from Proposition 4, the *intermediate protest equilibrium with exogenous* C is selected in the subgame, and it generates an expected payoff of $\hat{B}(C) = q(1-\delta_b) + (1-q)\ [(1-C)(1-\delta_b) + C(1-\alpha_g)]$. Suppose that in the game with precise commitment, the bad government chooses $(r_b, r_g) = (1, 1-C)$. To determine which equilibrium arises in the subgame following this choice, we calculate $F(1, 1-C) = (1-q)(\beta_g+\beta)(1-C) - q(\beta_b - \beta)$. Note that

$$F(1, 1-C) > (1-q)(\beta_g+\beta)\frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta} - q(\beta_b-\beta) = 0,$$
$$F(1, 1-C) \leq (1-q)(\beta_g+\beta)\frac{q}{1-q}\frac{\beta_b-p\beta}{p(\beta_g+\beta)} - q(\beta_b-\beta) = \frac{1-p}{p}q\beta_b = K,$$

where the last equality follows from the definition of K in Lemma 1. Hence, $\pi = \delta_b$ in the subgame following the choice of $(r_b, r_g) = (1, 1-C)$, and hence, from equation (4), the bad government's payoff is $B(1, 1-C, \delta_b) = q(1-\delta_b) + (1-q)\ [(1-C)(1-\delta_b) + C(1-\alpha_g)] = \hat{B}(C)$.

Case III. Suppose C is such that $q \geq q_2 g(1)$, that is,

$$q \geq q_2\ \left(1 - C\frac{\beta_b-\beta}{\beta_b-\beta+(1-C)(\beta+\beta_g)}\right) \ \Leftrightarrow\ C \geq 1 - \frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}.$$

Following this choice of C, from Proposition 4, the *low protest equilibrium with exogenous* C is the unique equilibrium of the subgame,[10] and generates an expected payoff of $\hat{B}(C) = q + (1-q)(1-C)$. Suppose that in the game with precise commitment, the bad government chooses $(r_b, r_g) = (1, 1-C)$. To determine which equilibrium arises in the subgame following this choice, we calculate $F(1, 1-C) = (1-q)(\beta_g+\beta)(1-C) - q(\beta_b-\beta)$. Note that

$$F(1, 1-C) \leq (1-q)(\beta_g+\beta)\frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta} - q(\beta_b-\beta) = 0,$$

[10]If $q = q_2 g(1)$, this equilibrium is selected by our equilibrium selection described above.

and hence, $\pi = 0$ in the subgame following $(r_b, r_g) = (1, 1 - C)$. Thus, from equation (4), the bad government's payoff is $B(1, 1 - C, 0) = q + (1 - q)(1 - C) = \hat{B}(C)$. ■

Lemma 8 *Fix a pair of priors (p, q). Let (r_b, r_g) be the equilibrium strategy of the bad government in the game with precise commitment. In the game with endogenous C, suppose the bad government chooses $C = 1 - r_g$. Then, the equilibrium of the subgame following this choice of C is outcome-equivalent to the equilibrium of game with precise commitment.*

Proof. Proposition 2 characterizes the equilibrium choices of (r_b, r_g). First, consider case 1 of Proposition 2, where $(r_b, r_g) = (1, 1)$ in equilibrium. Let $C = 0$. Then, both in the equilibrium of the game with endogenous C and in the equilibrium of the game with precise commitment, the good opposition is always successfully repressed by the bad government, the bad opposition is always repressed by both types of the government, and the public never protests.

Second, consider case 2 in Proposition 2, which corresponds to $q_1(p) < q < q^* < q_2$. Let $C = 1 - r_g = 0$, and observe that, $C = 0$ implies $g(p) = 1$, so that $[q_1(p)g(p), q_2 g(1)) = [q_1(p), q_2]$. Given our equilibrium selection, when $q \in [q_1(p)g(p), q_2 g(1)) = [q_1(p), q_2)$, the *intermediate protest equilibrium with exogenous C* (from Proposition 3) is selected. For $C = 0$, this equilibrium is identical to the *intermediate protest equilibrium* (from Proposition 1) that obtains in the game with precise commitment.

Third, consider case 3 in Proposition 2, where $r_b = 1$ and $r_g = \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta_g+\beta)}$ in equilibrium. Let

$$C = 1 - \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta_g + \beta)}. \tag{8}$$

Observe that this choice of C implies $q = q_1(p)g(p)$. Given our equilibrium selection, with a C such that $q = q_1(p)g(p)$, the equilibrium of the subgame is the *intermediate protest equilibrium with exogenous C* in which $\pi = \delta_b$, $\rho_g^B = 1$, $\rho_b^B = 1$, $\rho_g^G = 0$, and

$$\rho_b^G = \frac{p}{1-p}\Big(\frac{(\beta + \beta_g) - (\beta_b + \beta_g)q}{\beta_b q} - \frac{1-q}{q}\frac{\beta + \beta_g}{\beta_b}\, C\Big). \tag{9}$$

Substituting from equation (8) to (9) yields $\rho_b^G = 1$. Moreover, in the game with endogenous C, the probability that repression against the good opposition succeeds is $\rho_g^B \times (1 - C) = \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta_g+\beta)}$, which is the same as the equilibrium level of r_g in the game with precise commitment. Thus, these equilibria are outcome-equivalent.

Fourth, consider case 4 in Proposition 2, where $r_b = 1$ and $r_g = \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g+\beta}$ in equilibrium. Let

$C = 1 - \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}$. Observe that this choice of C implies $q = q_2 g(1)$. Given our equilibrium selection, with a C such that $q = q_2 g(1)$, the equilibrium of the subgame is the *low protest equilibrium with exogenous* C in which $\rho_g^G = 0$, $\rho_b^G = \rho_g^B = \rho_b^B = 1$ and $\pi = 0$. In this equilibrium, $\rho_g^B = 1$, and hence the probability that repression against the good opposition succeeds is $\rho_g^B \times (1 - C) = r_g$. Hence, this equilibrium is outcome-equivalent to equilibrium of the game with precise commitment.

■

Lemma 9 *The equilibrium of the game with endogenous C is generically unique.*

Proof. In the game with endogenous C, the bad government's payoff of selecting C is

$$\hat{B}(C) = q(1 - \pi(C)) + (1 - q)\ [(1 - C)\rho_g^B(C)(1 - \pi) + (1 - (1 - C)\rho_g^B(C))(1 - \alpha_g)],$$

where $\pi(C)$ and $\rho_g^B(C)$ depend on the equilibrium of the subgame following the choice of C.

Consider $q \in (0, q_1(p))$. First, note that

$$q \geq q_2 g(1) \Leftrightarrow C \geq C_H \equiv 1 - \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}, \quad \text{and} \quad C_H \in (0, 1).$$

Second, note that

$$q \geq q_1(p)g(p) \Leftrightarrow C \geq C_L \equiv 1 - \frac{q}{1-q}\frac{\beta_b - p\beta}{p(\beta_g + \beta)}, \quad \text{and} \quad C_L \in (0, C_H).$$

It follows that (1) for $C < C_L$, the *high protest equilibrium with exogenous* C is unique in the subgame following the bad government's initial choice, generating $\pi(C) = \alpha_g$ and $\rho_g^B(C) = \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}\frac{1}{1-C}$, (2) for $C_L \leq C < C_H$, the *intermediate protest equilibrium with exogenous* C is selected in the subgame following the bad government's initial choice, generating $\pi(C) = \delta_b$ and $\rho_g^B(C) = 1$, and (3) for $C \geq C_H$, *low protest equilibrium with exogenous* C is selected, generating $\pi(C) = 0$ and $\rho_g^B(C) = 1$. Hence, for $C < C_L$, the bad government's payoff is $\hat{B}(C) = 1 - \alpha_g$, which does not depend on C. At $C = C_L$, the protest probability $\pi(C)$ jumps down from α_g to δ_b, and it is δ_b for all $C \in [C_L, C_H)$; similarly, $\rho_g^B(C)$ jumps up from $\frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}\frac{1}{1-C}$ to 1, and it is 1 for $C \geq C_L$. Hence, $\hat{B}(C)$ has an upward jump discontinuity at $C = C_L$ at which it is right continuous, and for $C \in [C_L, C_H)$, $\hat{B}(C)$ is a decreasing linear function of C. At $C = C_H$, the protest probability $\pi(C)$ jumps down from δ_b to 0, and it is 0 for all $C \geq C_H$. Hence, $\hat{B}(C)$ has an upward jump discontinuity at $C = C_H$ at which it is right continuous, and for $C \geq C_H$, $\hat{B}(C)$ is a decreasing linear function of C. Hence, only $C = 0$, $C = C_L$, or $C = C_H$ could be optimal. Next, note that $\hat{B}(0)$ does not depend on q, and $\hat{B}(C_H)$ are both linear functions of q with different slopes, and hence each of the following equations holds for no more than a single value of q: $\hat{B}(0) = \hat{B}(C_H)$, $\hat{B}(0) = \hat{B}(C_L)$, or $\hat{B}(C_L) = \hat{B}(C_H)$. Hence, outside of knife-edge cases, the equilibrium value of C is unique. The remaining cases, $q > q_2$ and $q \in (q_1(p), q_2)$, are analogous and simpler. ■

Proof of Proposition 3. Lemma 7 implies that the bad government's equilibrium payoff in the game with precise commitment is an upper bound for its equilibrium payoff in the game with endogenous C. Lemma 8 shows that for each (p, q), the bad government can always select a C to achieve this upper bound. Thus, such a C must be an optimal choice. Moreover, Lemma 8 also shows that when the bad government selects this optimal C, the equilibrium of the game is outcome-equivalent to the equilibrium of the game with precise commitment. Finally, Lemma 9 show that the optimal C is unique. ■

B EXTENSION: OTHER CASES WITH COMMITMENT

Because the subgame that follows a choice of (r_b, r_g) has multiple equilibria when $F(r_b, r_g) \in (0, K)$, the equilibrium of the full game depends on which of the three possible equilibria is anticipated in the subgame. In Proposition 2 we focus on the equilibrium with $\pi = \delta_b$. In the following two propositions, we characterize the equilibrium when $\pi = 0$ and $\pi = 1$. In both cases, the public does no protest in equilibrium.

Proposition 5 *Suppose that when $F(r_b, r_g) \in (0, K)$, the equilibrium of the subgame has $\pi = 0$. In equilibrium, the public never protests upon observing repression, the good government and the bad government always repress the bad opposition, and (1) if $q > q_1(p)$, then the bad government also always represses the good opposition, but (2) if $q < q_1(p)$, it represses the good opposition with a positive probability less than one.*

Proof. Lemma 3 establishes part 1 for $q > q_2 > q_1(p)$. If $q_1(p) < q < q_2$, then $F(1, 1) \in (0, K)$. Hence, with $r_b = r_g = 1$, $\pi = 0$ in the equilibrium of the subgame. The monotonicity of $B(r_b, r_g, \pi)$ implies that $r_b = r_g = 1$ must be the bad government's equilibrium choice.

Suppose that $q < q_1(p)$. First, we show that if $F(r_b, r_g) = K$ in equilibrium, then we must have $\pi = 0$ in the equilibrium of the subgame. Suppose not, i.e., $\pi > 0$. Because $F(r_b, r_g) = K > 0$, we have $r_g > 0$, and hence r_g can be reduced. If the bad government slightly decrease r_g by ϵ, then $0 < F(r_b, r_g - \epsilon) < K$, hence $\pi = 0$ in the equilibrium of the subgame, and hence the bad government gains by such a deviation: $B(r_b, r_g - \epsilon, 0) - B(r_b, r_g, \pi) = \pi((1 - q)r_g + qr_b) - \epsilon\alpha_g(1 - q) > 0$ for sufficiently small ϵ.

In addition, any combination of (r_b, r_g) for which $F(r_b, r_g) > K$ is dominated by $r_b = r_g = 0$, and cannot be the bad government's equilibrium choice. If $F(r_b, r_g) > K$, then $\pi = 1$, and hence $B(r_b, r_g, 1) = q(1 - r_b)(1 - \alpha_b) + (1 - q)(1 - r_g)(1 - \alpha_g)$. The bad government can benefit by deviating to $(r_b, r_g) = (0, 0)$, so that $F(r_b, r_g) = 0$, and its expected payoff becomes $B(0, 0, 0) = q(1 - \alpha_b) + (1 - q)(1 - \alpha_g) > B(r_b, r_g, 1)$.

Therefore, the bad government's equilibrium choice solves the following maximization problem:

$$\max_{(r_b, r_g) \in [0,1]^2} B(r_b, r_g, 0) \quad \text{s.t.} \quad F(r_b, r_g) \leq K.$$

B is increasing in r_b and r_g, and F is increasing in r_g and decreasing in r_b. Thus, at the optimum $r_b = 1$ and r_g satisfies $F(1, r_g) = K$. ■

Proposition 6 *Suppose that when $0 < F(r_b, r_g) \leq K$, the equilibrium of the subgame has $\pi = 1$. In equilibrium, the public never protests upon observing repression, the good government and the bad government always repress the bad opposition, and (1) if $q > q_2$, then the bad government also always represses the good opposition, but (2) if $q < q_2$, it represses the good opposition with a positive probability less than one.*

Proof. Lemma 3 establishes part 1. We focus on $q < q_2$. There is no equilibrium in which $F(r_b, r_g) > 0$ because if $F(r_b, r_g) > 0$, then $\pi = 1$, and $B(r_b, r_g, 1) < B(0, 0, 0)$. Next, suppose that if $F(r_b, r_g) = 0$, then $\pi = 0$ in the equilibrium of the subgame. Therefore, the bad government's equilibrium choice becomes:

$$\max_{(r_b, r_g) \in [0,1]^2} B(r_b, r_g, 0) \;\; \text{s.t.} \;\; F(r_b, r_g) \leq 0.$$

B is increasing in r_b and r_g, and F is increasing in r_g and decreasing in r_b. Thus, at the optimum, $r_b = 1$ and r_g satisfies $F(1, r_g) = K$.

Finally, we show that no equilibrium exists if $\pi > 0$ in the subgame that follows the bad government's strategy (r_b, r_g) such that $F(r_b, r_g) = 0$. Because $F(r_b, r_g) = 0$, we have either $r_b > 0$ and $r_g > 0$ or $r_b = r_g = 0$. If $r_g > 0$, then r_g can be reduced. If the bad government slightly decrease r_g by ϵ, then $F(r_b, r_g - \epsilon) < 0$, hence $\pi = 0$ in the equilibrium of the subgame, and hence the bad government gains by such a deviation. Similarly, if $r_b = r_g = 0$, then r_b can be increased. If the bad government slightly increases r_b by ϵ, then $F(r_b + \epsilon, r_g - \epsilon) < 0$, hence $\pi = 0$ in the equilibrium of the subgame, and hence the bad government gains by such a deviation. ■