International Pressure,
State Repression and the Spread of Protest

Abstract

We analyze strategic interactions between a state that decides whether to repress activists and the general public that decides whether to protest following repression. The public would like to support activists who demand beneficial reforms, but it is uncertain about both the merit of the activists' demands and the intentions behind the state's repression. Multiple equilibria arise, suggesting an important role for social norms, which provides a rationale for the conflicting empirical findings on the determinants of repression and reform. We show that international pressure, which directly reduces the state's ability to repress, can indirectly increase repression by shifting the public's belief in favor of the state, thereby reducing its incentive to protest. To protect legitimate activists or promote positive reforms, international pressure must be sufficiently strong. Lukewarm international commitments at best achieve nothing, and at worst crowd out domestic checks on repression, generating the *opposite* of their intended effects.

Keywords: Repression, Protest, Legitimate Coercion, Social Norms, International Pressure.

1   INTRODUCTION

On September 8, 1978, Iranian security forces fired on protesters in Tehran, killing many demonstrators. Soon after, protests continued in larger numbers, and strikes swept the country, culminating in the 1979 Iranian Revolution (Abrahamian 1982). Iran is not an isolated case. Martin (2007) provides several case studies of the spread of protest following repression, and Francisco (2004) quantifies the magnitude of such spreads in many more cases—see Earl (2011) for a review.[1] The theoretical literature has focused on the deterrence effect of repression, ignoring the empirical evidence that the public sometimes joins the protest following repression. This paper studies the subtle interactions between the state's incentive to repress activists and the public's incentive to join the protest, exploring the role of international pressure in shaping the interaction between the state and the general public. What factors influence the likelihood that protest spreads following repression, and how does this potential for the spread of protest affect the state's response to dissent? How do the efforts of international institutions to protect legitimate dissidents affect the public's incentives to protest? Can such efforts generate the opposite of their intended effects?

We consider contexts in which a group of activists has initiated a protest and put forth demands, and the state must decide whether to concede to their demands or repress them. We refer to the state's use of coercive force as repression, which includes imprisonment, killing, or other punishments.[2] Because a main responsibility of the state is to protect its citizens against harm by transgressors, the public recognizes that the use of coercive force

---

[1]Other examples abound. On March 21, 1960, South African security forces fired on demonstrators in Sharpeville, killing unarmed civilians. Subsequently, protests and strikes spread throughout the country. In Ukraine, following the repression of protesters on November 30, 2013, the scale of protest expanded dramatically, culminating in the ouster of President Yanukovich. Notably, in a survey of over a thousand protesters in Independence Square, conducted on December 7-8, about 70% of respondents chose repression of November 30 as a main reason for joining the protest (Survey conducted by the Kiev International Institute of Sociology and the Ilko Kucheriv Democratic Initiative Foundation. Available at: http://www.kiis.com.ua/?lang=eng&cat=reports&id=216&page=11). Similarly, in a survey of protesters in Gezi Park protest in Turkey between June 6-8, 2013, about half of over 4400 respondents indicated that they joined the protests after observing government's repression (KONDA 2014, p. 20); see Aytaç et al. (2018) for a detailed analysis of Gezi protests). See also Wood (2003, p. 233-4) and Aytaç and Stokes (2018).

[2]This definition of repression is consistent with the notion of repression used in theoretical models (see the literature review), and with the sociology literature on repression. For example, Tilly (1978) defines repression as any action by the state that "raises the contender's cost of collective action" (p. 100).

may be legitimate.[3] However, both the activists who protest and the states that repress them claim that their actions are in the public's best interests. Given the difficulties of obtaining precise information, the public remains uncertain about the nature of the activists' demands and the intentions of the state.[4] This often-ignored uncertainty is at the core of the interaction between the state and citizens in this paper.

This paper develops a model of strategic interactions between a government that must decide whether to repress a group of activists or concede to their demands, and a bystander citizen, representing the general public, who must decide whether to join the activists' protest upon observing repression. There are two types of activists: good and bad. The good activists' demands (if implemented) are beneficial to the public, while the bad activists' demands are harmful. Similarly, there are two types of governments: good and bad. Both types of government prefer to stay in power, but they differ in their preferences for reform. The good government's preferences for change are aligned with the public's: the good government prefers the good activists' reforms over the status quo and the status quo over the bad activists' reforms. However, the bad government prefers the status quo over both beneficial and harmful reforms. The government observes the activists' type. In contrast, the public does not observe the types of the activists or the government. Therefore, upon observing repression, the public cannot distinguish whether a bad government has repressed good activists—blocking beneficial reforms—or the government (of either type) has repressed bad activists—a necessary action that protects the public.

The bystander citizen's uncertainty about the types of the government and activists underlies the citizen's fundamental tradeoff: by supporting activists and toppling the government, he risks implementing harmful social changes, but by supporting the government's decision to repress activists, he risks blocking beneficial reforms. In turn, the potential for

---

[3]This view resonates with Weber's notion of "the legitimate use of physical force," or "legitimate coercion" as appears in Almond (1956) and Mansbridge (2012, 2014). The central role of coercive force in government has long been emphasized by statesmen. When Washington sent militia to suppress the Whiskey Rebellion in 1794, Hamilton explained that "government can never [be] said to be established until some signal display has manifested its power of military coercion" (Wood 2009, p. 111).

[4]In a survey of Turkish citizens conducted on July 6-7, 2013, with over 2600 participants from 28 provinces, over 55% of the respondents believed that "the [Gezi Park] protests are a plot set up by foreign conspirators which resent Turkey's development" is a "right" or "absolutely right" statement, while about 27% believed that it is "wrong" or "absolutely wrong" (KONDA 2014, p. 59).

the bystander's protest creates an endogenous cost of repression that drives the government's fundamental tradeoff: by repressing the activists, the government prevents costly social changes, but it risks public protests that would topple it. However, by conceding to the activists' demands, the government pays a cost, but it avoids public protests. These two fundamental tradeoffs are linked through the bystander's posterior beliefs about the types of the government and the activists upon observing repression. Because good and bad governments have different incentives to repress different types of activists, observing repression is informative about both the type of the government and the type of the activists.

We show that the bystander's Bayesian updating—the formation of public opinion—has two key features. First, repression is bad news about *both* the government and the activists: upon observing repression, the bystander citizen updates negatively about both the government and the activists. Second, when the good government represses bad activists more often, the bystander updates more negatively about the activists and less negatively about the government.[5] This reduces the public's incentives to join the activists' protest, lowering the (endogenous) cost of repression for the government, thereby raising the good government's incentive to repress bad activists. The strategic complementarity between the public's incentive to protest and the good government's incentive to concede generates the potential for multiple equilibria. In fact, we show that there can be three equilibria, which can be ranked according to their level of repression.

The multiplicity of equilibria suggests that social norms play a key role in determining the state's response to dissent and the public's response to repression by shaping the expectations of the public and the state of each other's behavior.[6] This suggests an explanation for the conflicting empirical findings on the determinants of repression. For example, while some studies suggest that higher income per capita reduces the likelihood of repression (Mitchell and McCormick 1988; Henderson 1991; Poe and Tate 1994), others have not found a significant correlation (Gandhi 2008; Conrad and Moore 2010; Shadmehr and Haschke 2016; Davenport (2007) provides a review). Our analysis shows that reductions in income can

---

[5]When the bad government is more likely to repress good activists, the opposite updating happens.

[6]This interpretation is in line with the literature that interprets equilibrium selection as a reflection of social norms (Kreps 1990; Burke and Young 2011; Myerson 1991 p. 113-114; Postlewaite 2011; Acemoglu and Jackson 2015; Young 2015).

increase or decrease repression depending on the social norms that select the equilibrium. Because social norms likely vary across countries, this result suggests a rationale for these conflicting empirical findings.

More generally, our analysis provides a counterpoint to the common theme in the repression literature that higher income reduces repression by reducing the incentives to protest (Mitchell and McCormick 1988; Henderson 1991; Poe and Tate 1994; Bueno de Mesquita et al. 2005). As Davenport (2007, p. 14) highlights, a common rationale of the literature is that "fewer resources enhance the need for coercive behavior by increasing societal grievances." However, this approach does not distinguish between the *initiation* of protest and its *spread*. If higher income makes the public more hesitant to protest following repression, this reduces the endogenous cost of repression for the state, *increasing* the state's incentive to repress the activists.

Our analysis exhibits three main points of departure from the theoretical literature on protests and revolutions.[7] (1) The literature assumes that the public knows whether the state's use of coercion is legitimate or illegitimate. In contrast, our theoretical framework emphasizes that the public remains uncertain about whether the state's use of coercion was legitimate, meant to protect the public, or illegitimate, aimed to preserve the state's vested interests. A key implication is that acts of coercion are informative to the public about the intentions of the state and its challengers. (2) The literature focuses on the deterrence effect of repression, modeling repression as state actions that either directly or indirectly reduce the likelihood of revolution.[8] In contrast, in our analysis, repression can be followed by a spread

---

[7]Topics studied in this literature include: coordination (Bueno de Mesquita 2010; Shadmehr & Bernhardt 2011; Tyson & Smith 2014), leaders and their tactics (Bueno de Mesquita 2010, 2013; Wantchékon & García-Ponce 2014; Morris & Shadmehr 2017; Shadmehr & Bernhardt 2019), the role of media (Egorov et al. 2009; Edmond 2013; Shadmehr & Bernhardt 2015; Guriev & Treisman 2015; Barbera & Jackson 2016; Huang et al. 2018), the effect of elections (Little 2012; Egorov & Sonin 2015; Lou and Rozenas 2018), signaling strength (Ginkel & Smith 1999), and contagion (Chen & Suen 2016).

[8]In Acemoglu & Robinson (2001), Boix (2003), Besley & Persson (2011), Svolik (2013), and Shadmehr (2014), repression directly reduces the likelihood of revolution. In Persson and Tabellini (2009), Bueno de Mesquita (2010), Shadmehr & Bernhardt (2011), Boix & Svolik (2013), Casper & Tyson (2014), Guriev & Treisman (2015), and Egorov & Sonin (2015), repression indirectly reduces the likelihood of revolution by increasing the cost of dissent. Siegel (2011) studies how the structure of social networks determines whether and when citizens' anger and fear of repression leads to backlash protests—see also Shadmehr (2014, Appendix A) and Bueno de Mesquita & Dickson (2007). In Guriev & Treisman (2015), a dictator's repression of the elite, when observed, immediately reveals his incompetence and leads to his removal—see also Egorov & Sonin (2015).

of protest that further destabilizes the regime. Consistent with numerous historical incidents and empirical evidence, repression does not necessarily deter revolution: it may also spark it, creating an endogenous cost of repression for the regime. (3) The literature has extensively studied complementarities that arise due to coordination aspects of revolution. In contrast, our analysis identifies novel strategic complementarities between the public's incentive to protest and the (good) government's incentive to concede. These complementarities drive the multiplicity of equilibria and the role of social norms.

In our initial analysis, the only force that prevents bad governments from repressing activists is the endogenous threat of protest by the public. However, governments often face additional restrictions on the repression of legitimate activists which are imposed by the international community. Foreign governments and international institutions can sometimes prevent the repression of legitimate dissidents by using a variety of instruments such as suspending aid, terminating trade benefits, making exclusive memberships contingent on observation of human rights, or threatening military intervention (Simmons 2009; Hafner-Burton 2013; Magesan 2013). We investigate the impact of such restrictions using our model.

Even though international pressure sometimes directly blocks the repression of good activists, we show that the overall likelihood that good activists are repressed may *increase*. This result is driven by the effect of international pressure on the bystander's updating: because international pressure sometimes prevents the bad government from repressing good activists, the bystander citizen updates less negatively about the government and more negatively about the activists upon observing repression. By shifting the bystander's beliefs in favor of the government, international pressure reduces the bystander's incentive to protest,[9] thereby leading the bad government to *attempt* repression of good activists more often. Thus, the informational effect of international pressure dampens the general public's response, which can increase the overall likelihood of repression.[10] Though Simmons

---

[9]This logic is consistent with Hollyer and Rosendorff's (2011) empirical evidence that the overall level of protest can fall following the participation in human rights treaties.

[10]These effects of international interventions have surface similarities with the effects of government grants on charities. For example, in Andreoni and Payne (2003), government grants (government intervention) reduce private donations by generating free-riding among contributors and by decreasing the charity's fund-raising efforts. Nevertheless, government grants always achieve their intended effect, increasing the charity's overall budget. In contrast, in our analysis international intervention reduces the bystander's incentive to protest through its effect on beliefs, and it can result in more repression of legitimate activists,

(2009, p. 135-55) argues that international human rights treaties provide domestic activists with opportunities "to mobilize for human rights," the empirical evidence on this is mixed. Human rights treaties sometimes seem to reduce and sometimes seem to raise repression (Hathaway 2002; Hafner-Burton and Tsutsui 2005; Smith-Cannoy 2012; Magesan 2013). As Hafner-Burton (2012) argues, "there is a troubling and recurrent finding that participation in some treaties correlates with worse human rights behavior" (p. 280). Our paper provides a novel rationale for why international pressure in general, and human rights treaties in particular, may lead to an increase in repression: by altering the nature of updating (public opinion) following repression in favor of the state, international pressure can mitigate the public's incentive to protest, thereby reducing the domestic costs of repression to the state.

Moreover, we show that even when the overall likelihood that good activists are repressed falls, the likelihood that good reforms are implemented can also fall. Therefore, the international community may face a tradeoff between protecting legitimate dissidents and promoting good reforms. These results generate a sharp and simple policy implication: in order to protect good activists, international pressure must be sufficiently strong; otherwise, at best it achieves nothing, and at worst it generates the *opposite* of its intended effect.

Before we proceed, we emphasize a key scope condition of our model. Our focus is on non-democratic settings in which the public is uncertain about the intentions of the state and dissidents, and a primary goal of the state is to avoid turning public opinion against itself to prevent the spread of protest from activists to the general public. When the public knows that the state is using coercion to block beneficial reforms of legitimate dissidents, the state's primary concern switches to other strategic considerations such as hampering coordination or signaling its strength, which have been the focus of the literature.

We next presents the model. Section 3 discusses our modeling assumptions, including the information structure, and their relation with the formal literature. We analyze the model in Sections 4 and 5. Section 6 investigates the role of international pressure. A conclusion follows. Proofs are in Online Appendix I.

---

achieving the opposite of its intended effect.

2 MODEL

We consider a game with two strategic players: a ruler and a bystander citizen. In addition, there is a non-strategic activist who protests, demanding that the ruler implements a set of social changes. The activist is one of two possible types: a "good" activist (type g) demands reforms that would benefit the bystander (relative to the status quo) if implemented, while reforms demanded by a "bad" activist (type b) would hurt the bystander. The ruler is also of two possible types, "good" (G) or "bad" (B). Like the bystander, a good ruler prefers good reforms over the status quo and prefers the status quo over bad reforms. The bad ruler prefers good reforms to bad reforms, but prefers the status quo to reforms of either type. Both types of ruler receive a private benefit from being in office. The ruler observes the activist's type, but the bystander does not observe either the type of the ruler or the type of the activist. Under the common prior, the ruler is bad with probability $p \in (0,1)$, and the activist is bad with probability $q \in (0,1)$.[11]

The game proceeds as follows. First, nature chooses the ruler's type and the activist's type. The ruler observes the activist's type, and then decides whether to concede to the activist or repress him. If the ruler concedes to the activist, the game ends. If the ruler represses, the bystander citizen decides whether to protest.[12] When the bystander protests, the current ruler is removed from office and is replaced by the activist, who implements his preferred reform. Otherwise no reform is implemented and the ruler retains power. Payoffs are realized at the end of the game.

The bystander's payoff is the sum of two components. The first component is the valence of the leader who is in power at the end of the game (depending on the outcome, this is either the initial ruler or the activist). The second component is the value of the terminal policy. Naturally, the public prefers that a good leader is in power at the end. The valence of a good leader is normalized to zero, while the valence of a bad leader is $-\beta \leq 0$. Similarly,

---

[11]The main empirical implications of our model do not require the researcher to know the specific values of p and q. However, it is worth emphasizing that p and q are essentially public opinion parameters that can be estimated; for example, surveys are used to gauge the public's views of the government and opposition groups even in authoritarian regimes (Cammett 2011; Jamal 2012; Beissinger et al. 2015).

[12]In our model the bystander always learns that the ruler repressed the activist. The analysis could be routinely extended to incorporate censorship, whereby the bystander observes repression with some probability. Provided that the degree of censorship is not too large, our results extend with no substantive changes.

Outcome	Bystander's Payoff	Ruler's Payoff
Status Quo under Bad Ruler	$-\beta$	1
Status Quo under Good Ruler	0	1
Ruler Replaced by Bad Activist	$-\beta_b$	0
Ruler Replaced by Good Activist	β_g	0
Bad Ruler Concedes to Bad Activist	$-\beta_b$	$1 - \alpha_b$
Good Ruler Concedes to Bad Activist	$-(\beta_b - \beta)$	$1 - \delta_b$
Bad Ruler Concedes to Good Activist	$\beta_g - \beta$	$1 - \alpha_g$
Good Ruler Concedes to Good Activist	β_g	$1 + \delta_g$

Table 1: The bystander and ruler payoffs for each final outcome. Parameter restrictions: $\beta_g > 0$, $\beta_b > \beta \geq 0$, $\delta_i > 0$, $\delta_b < 1$, $0 < \alpha_g < \alpha_b < 1$.

the bystander's value for the status quo policy is normalized to zero, the value for a good reform is $\beta_g > 0$, and the value for a bad reform is $-(\beta_b - \beta) < 0$. We summarize the bystander payoffs following all possible game outcomes in Table 1.[13] Because a bad reform is strictly harmful for the bystander ($\beta_b > \beta$), protesting always has a downside: joining a *bad* activist's protest is worse for the bystander than supporting the ruler, even if the initial ruler is bad. Furthermore, because the status quo under the good ruler is weakly better than under a bad ruler, the net loss of replacing a good ruler with a bad activist (β_b) is at least as large as the net loss of replacing the bad ruler with the bad activist ($\beta_b - \beta$). Thus, the potential downside of protesting against a good ruler is at least as large as the potential downside of protesting against a bad ruler.

The ruler's payoff depends on whether or not he retains office, the terminal policy, and his type. If a ruler is removed from office (no matter by whom), then his payoff is normalized to zero. If the ruler maintains office without implementing any reforms, the ruler receives an office rent of 1. A good ruler prefers a good reform to the status quo, but prefers the status quo to a bad reform. Hence, if the good ruler retains office by conceding to the good activist's demands, his terminal payoff is $1 + \delta_g$; if he concedes to a bad activist, his payoff is $1 - \delta_b$, where $0 < \delta_i$ and $\delta_b < 1$. Meanwhile, a bad ruler prefers the status quo to either reform, but prefers good reforms to bad reforms. Hence, if the bad ruler concedes to the good activist, his

---

[13]For example, supporting a good activist following repression generates a beneficial reform, worth β_g, and leaves a good leader in power, and hence, the overall payoff of supporting a good activist is β_g. Meanwhile, supporting a bad activist following repression generates a harmful social change, reducing the bystander's by payoff by $\beta_b - \beta$, and leaves a bad leader in power, further reducing the bystander's payoff by β. Hence, the overall payoff of supporting a bad activist is $-(\beta_b - \beta) - \beta = -\beta_b$.

payoff is $1 - \alpha_g$; if he concedes to the bad activist, his 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 ruler's incentives to repress the good activist is larger than the good ruler's incentives to repress the bad activist.[14]

In Section 5, we modify the model to investigate the effects of international pressure on the bad ruler's ability to repress good activists, deriving conditions under which such interventions are counter-productive.

3 DISCUSSION OF THE MODEL

Before proceeding to the analysis, we discuss some of our modeling choices in more detail.

Information structure. We assume that the ruler is better informed than the bystander about the activist: the ruler observes the activist's type, but the bystander can only make inferences about it. This assumption is based on two observations. First, the ruler has access to resources (e.g., intelligence agencies) that gather and process information about the goals and preferences of the activists, which are inaccessible to the general public. Second, the general public may have difficulties in learning about the activists' true intentions, since rulers who use repression claim that they do so to protect citizens against harmful dissidents. Indeed, in their study on the collapse of democracies, Levitsky and Ziblatt (2018) state that "In just about every case of democratic breakdown we studied, would be authoritarians...have justified their consolidation of power by labeling their opponents as an existential threat" (p. 106).

For example, in protests proceeding the April 2013 Venezuelan presidential election, officials called the protesters "the reactionary, criminal and murderous right wing that is run by Henrique Capriles" (Vyas and Gonzalez 2013).[15] When protests broke out again in February 2014 in Venezuela, Delcy Rodriguez, the minister of information stated that the protesters "are not students, they are violent gangs. They are executing a plan with the goal of a civil war in Venezuela" (Minaya 2014a). "Mr. Maduro accused what he called 'fascist leaders'

---

[14]We focus on $\delta_b < \alpha_g$ because it captures the bad ruler's strong incentives to maintain the status quo, and because it leads to richer strategic interactions and more interesting results than the opposite case. Results for the case of $\delta_b > \alpha_g$ are available upon request.

[15]Similar statements were made by Iranian officials about the supporters of Mousavi who were protesting against election fraud following the 2009 Iranian presidential election.

financed by the U.S. of using highly trained teams to topple his socialist government from power...he charged that the demonstrators were trying 'to fill the country with violence and to create a spiral of hatred among our people.' He said his foes were hoping to generate chaos to justify a foreign military intervention. 'In Venezuela, they're applying the format of a coup d'état,' he said" (Minaya 2014b).

To some citizens, these statements are dictators' cliches, but they are believable arguments to others. For example, "Danny Ojeda, 44, who works in distant Guarico state but came with other pro-government workers, said he agreed with the late president's [Chavez's] assertions that the U.S. government was behind Venezuela's troubles, a claim also made by Mr. Maduro. 'There's an economic coup against our country, but Maduro has the support to overcome it,' he said" (Forero 2014). Faced with opposing claims from the government and activists, many ordinary citizens remain uncertain about whom to believe and support.

Indeed, sometimes the rulers' statements contain some truth. In June 1975, a group of seminary students gathered in the Fayzyah Seminary School in Qom, Iran, demonstrating against the Pahlavi regime and praising Khomeini who was in exile. The regime responded with repression, beating and arresting the protesters. The Shah argued that the protest was the result of "the unholy alliance of black reactionist[s] and stateless Reds" (Kurzman 2003, p. 289). By "black reactionists" the Shah meant religious fanatics whose goal was to establish a theocracy, and by "state-less reds" he meant Soviet-backed communists who wanted to establish a communist state. We now know that they were not communists, but a subset of those protesters did want to establish a theocratic state in Iran although they were not explicit about it at the time.

Why no protest following concessions? We view the bystander citizen as a follower who may join an already existing protest but does not initiate one. This view is consistent with the robust finding in the social movements literature that sustaining protest activities takes significant resources and planning that require activists (Gamson 1975; McCarthy and Zald 1977; Tilly 1978, 2004; Tarrow 1998; McAdam 1999; McAdam, Tarrow, and Tilly 2001). Spontaneous protests occasionally occur, but they typically end quickly without any policy change. Indeed, many *seemingly* spontaneous movements are based on complex networks of

organizations and committed activists (Morris 1984; Diani and McAdam 2003; Khatib and Lust 2014). Therefore, in our model, when the ruler concedes to the activist's demands, the movement ends, and the ruler retains power.[16]

Repression by the good ruler. In line with the literature, in our model, the state (regime or ruler) can either repress the activists or concede to them, e.g., redistribution or democratization in Acemoglu and Robinson's (2001, 2006) framework or Boix's (2003). That is, a protest in these models is not an obscure demonstration on a street corner that the state can simply ignore. Rather, protest refers to an organized movement that will achieve its goals unless stopped by coercive force. With this notion of protest in mind, consider a minority of religious activists who protest against the state and attempt to implement religious laws against the preferences of the majority, or a fascist group who protests and attempts to impose its racist views on the society. A good ruler prevents this small minority from imposing their preferences on the majority by using coercive means, e.g., arresting and imprisoning the activists. This does *not* imply that the good government persecutes this group for their views. Rather, repression in this context means that the government uses its coercive means to prevent these groups from imposing their religious laws or racist policies on the majority.

What is bad about the bad ruler? Our analysis highlights two ways in which bad rulers are harmful to the general public. First, a bad ruler's preference for reform is not aligned with the public's. Provided he can maintain power, a bad ruler prefers to block reforms that the public finds beneficial. For example, a bad ruler may grant monopolies to enrich himself and his cronies at the cost of public welfare. If activists demand that the monopoly protections are revoked, the bad ruler would inherently prefer that the reform is blocked, even though it is beneficial to the public. This feature of the bad ruler is captured in the model by his payoffs (α_g, α_b). Given that he maintains power, the bad ruler prefers the status quo over both beneficial and harmful reforms, $(1 > 1 - \alpha_g > 1 - \alpha_b)$, while the good ruler prefers good reforms to the status quo, and the status quo over bad reforms. Second, a bad ruler's incompetence, corruption, and rent seeking behaviors harm the general public, independent of the ruler's

---

[16]This feature of our model captures an important aspect of the nature of interactions between the state, activists, and the public in many authoritarian settings, but this feature is less suitable for democratic settings. In democratic settings, the public can vote the ruler out of office regardless of his policy choices.

11

specific policy choices. In other words, the valence of a bad ruler is worse than the valence of a good ruler. This feature of the bad ruler is captured by the valence parameter $\beta \geq 0$, which is the cost paid by the public whenever a bad ruler is in power at the end of the game.

We view these two features of a bad ruler as complementary: an incompetent or corrupt ruler is more heavily invested in the status quo and is therefore also more-likely to be resistant to reforms. Nevertheless, it is important to point out that the direct harm that the bad ruler inflicts on the general public is not fundamental to the strategic interaction we consider. Indeed, our parametric restriction $\beta \geq 0$ includes $\beta = 0$ as a special case, and our analysis applies directly to the case in which the bad ruler's valence is identical to the good ruler's. In contrast, the bad ruler's inherent preference to block beneficial reforms underlies the information content of repression, which is fundamental to our analysis.

Non-strategic Activist. We have assumed that the activist always protests to simplify the exposition. Our results readily extend when the activist's payoffs from receiving his demands are sufficiently high. Then, the activist always decides to protest in the low and intermediate repression equilibria in which the ruler may concede or the bystander citizen may protest, and hence these equilibria and their characteristics remain unchanged. Moreover, our modeling choice of having the activist always protest reflects the observation that some activists protest even when they are sure to be repressed as a matter of principle or as a strategy to gain publicity and raise future recruitment.

Direct Costs of Protest and Repression. We have abstracted from direct protest and repression costs to simplify analysis. Clearly, adding known costs does not change our results qualitatively. When costs are private knowledge, equilibria are in pure strategies. However, the tradeoffs and strategic complementarities that underlie our results (including multiple equilibria) also arise with private costs, while the analysis becomes more cumbersome.

Relationship to Democratic Settings. Our model has been developed for authoritarian settings. However, to further highlight its contrasts with the literature, it is useful to describe how it would map into democratic settings. Consider our base model, and relabel "ruler" as "incumbent politician", "activist" as "challenger", and "bystander citizen/public" as a "median voter" who is uncertain about the policy that maximizes his welfare. The challenger

proposes a policy and commits to implement it if elected. There are two types of (policy proposals by the) challenger: the good policy improves the median voter's payoff over the status quo, the bad one worsens it. There are also two types of incumbent: the good/congruent incumbent shares the median voter's policy preference, while the bad/incongruent incumbent is biased toward the status quo. The incumbent knows the challenger's type, but the median voter is uncertain about the types of the incumbent and the challenger—he does not know whether the challenger's policy improves or worsens the status quo. After the challenger announces his platform, the incumbent has two options: if he preemptively implements the challenger's policy, he ensures that he will retain office—the median voter is assumed to re-elect him, a feature of the model that fits authoritarian settings, but becomes problematic in democratic settings. If he maintains the status quo, then the median voter updates his beliefs, and votes accordingly. Beyond the base model, the exogenous restrictions, imposed by international institutions, that we consider are difficult to interpret in democratic contexts.

Relationship to the Political Agency Literature. From a theoretical perspective, the structure of our initial model, *without* international pressure, is related to the political economy literature on the interaction between informed politicians and uninformed voters. This literature investigates topics such as the pandering of politicians to voters' opinions (Canes-Wrone et al. 2001), politicians' incentives to acquire policy expertise (Prato and Strulovici 2013), optimal institutional design (Maskin and Tirole 2004; Fox and Stephenson 2011), government transparency (Fox 2007), or populism (Acemoglu et al. 2013). In this literature, the authority over actions rests solely with the decision maker. In contrast, in our model, both the ruler and the bystander share authority over the policy that is implemented. In particular, the ruler can concede to the activists, which is costly but always allows him to remain in power. However, if the ruler represses, the bystander can join the protest, topple the regime and change the policy. Therefore, in contrast to this literature, updating about the activist's type (the state of the world) enters into the bystander's calculations. This, in turn, underlies the strategic complementarities that arise between the actions of the good ruler and the bystander, a key force in our analysis.

A more distant literature studies the interactions between an uninformed principal and an informed expert with an uncertain bias in a cheap talk context (Sobel 1985; Morris 2001).

13

The expert sends a cheap talk message to the principal who then decides which action to take. In addition to the differences delineated above, unlike the cheap talk literature in which advice is free, in our model, different actions for the government are costly. Indeed, the government's cost of concession depends both on its own type and the activist's.

4 PRELIMINARY ANALYSIS

Strategies. If the ruler represses the activist, the bystander must decide whether to join the protest. The bystander strategy is a probability, $\pi \in [0, 1]$, representing the probability of joining the protest. The ruler'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 $i \in \{G, B\}$ ruler represses the type $j \in \{g, b\}$ activist.

Protest Strategy. When deciding whether to join the protest, the bystander faces a trade-off: support the ruler and possibly prevent the implementation of beneficial changes, or support the activist and risk the implementation of bad changes. The bystander's decision depends on his (updated) belief that the activist is bad, q', and on his (updated) belief that the ruler is bad, p'. The bystander's expected payoff from protesting is $\beta_g(1 - q') - \beta_b q'$: with probability q' the activist is bad, and the bystander receives $-\beta_b$; with the remaining probability $1 - q'$ the activist is good, and the bystander receives β_g. If the bystander does not protest, his expected payoff is $-\beta p'$: with probability p' the ruler is bad, and the bystander receives $-\beta$; with the remaining probability the ruler is good, and the bystander receives his payoff that is normalized to 0. Therefore, the bystander's best response is:

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

where the bystander's updated beliefs p' and q' depend on the ruler's strategy in equilibrium.

Repression Strategy. If the bystander never joined the protest, a good ruler would repress a bad activist, and a bad ruler would repress either type of the activist. What complicates the ruler's decision is that the bystander may join the activist's protest following repression,

in which case he is removed from office. When the ruler represses, he is deposed whenever the bystander joins the protest, which happens with probability π. Therefore, the ruler's expected payoff from repression is $1 - \pi$. However, if the ruler concedes, his payoff depends on both his type and the activist's type. Because the good ruler prefers a good reform to the status quo ($\delta_g > 0$), he always concedes to the good activist, $\rho_g^G = 0$. Recall that the good ruler's payoff of conceding to a bad activist is $1 - \delta_b$, and the bad ruler's payoff of conceding to a type $i \in \{g, b\}$ activist is $1 - \alpha_i$. Therefore, the ruler's best response is:

$$
(2) \qquad \rho_b^G = \begin{cases} 1 & \text{if } \pi < \delta_b \\ [0, 1] & \text{if } \pi = \delta_b \\ 0 & \text{if } \pi > \delta_b. \end{cases} \qquad \rho_i^B = \begin{cases} 1 & \text{if } \pi < \alpha_i \\ [0, 1] & \text{if } \pi = \alpha_i \\ 0 & \text{if } \pi > \alpha_i. \end{cases}
$$

The ruler's strategy weighs the benefit of repressing reforms that he does not favor with the cost of inciting protest by the bystander.

Remark 1. We focus on the case in which the ruler is removed from power whenever the bystander joins the activist's protest. Our analysis is virtually unchanged if, instead, the ruler is removed from power with some probability $\theta \in (0, 1)$ in this event. In particular, equation (1), the bystander's best response, is not affected by the probability θ. To see this, note that the bystander's expected payoff of joining the protest is $\theta(\beta_g(1 - q') - \beta_b q') + (1 - \theta)(-\beta p')$, while the expected payoff of not protesting is $-\beta p'$. Comparing these gives an identical best response to (1). The only change occurs in (2). The ruler's payoff of conceding is unaffected, but the ruler's expected payoff of repression is $1 - \theta\pi$. Thus, accounting for the possibility that the ruler retains power with probability $1 - \theta$ when the bystander protests simply requires us to scale up the parameters $(\delta_b, \alpha_g, \alpha_b)$ by a factor $1/\theta$.

Remark 2. There is always an equilibrium in which no repression takes place ($\rho_j^i = 0$). In any such equilibrium, the bystander joins the protest with sufficiently high probability upon observing repression, but this information set is off-the-equilibrium path. This behavior deters the ruler from repressing, but it is supported by the bystander's off-the-equilibrium-path beliefs that if the ruler represses, then with a high probability he must be a bad ruler

repressing a good activist. However, we show in Online Appendix I that this equilibrium does not satisfy the D1 criterion for equilibrium selection (Fudenberg and Tirole 2000, p. 452). To see why, note that the bad ruler dislikes conceding to the bad activist more than conceding to a good one, and the bad ruler is therefore more inclined to repress the bad activist. Hence, the D1 restriction on off-the-path beliefs rules out the possibility that the bystander believes that the bad ruler may have repressed the good activist. We therefore consider equilibria in which repression takes place with a positive probability.

If the bad ruler never repressed the good activist, then repression would imply that the activist must be the bad type. Consequently, the bystander would never join the protest. However, if the bystander never joins the protest, then the bad ruler would deviate by repressing the good activist. Therefore, in any equilibrium the bad ruler represses the good activist with positive probability ($\rho_g^B > 0$). Because the bad ruler dislikes the bad activist more than the good activist ($\alpha_b > \alpha_g$), if the bad ruler represses the good activist with positive probability, then he always represses the bad activist ($\rho_b^B = 1$). Moreover, because the bad ruler represses both types of activist with positive probability, the bystander remains uncertain about the activist's type following repression. Lemma 1 summarizes these observations.

Lemma 1 *In equilibrium, the good ruler never represses the good activist, but the bad ruler represses the good activist with a positive probability, and always represses the bad activist: $\rho_g^G = 0$, $\rho_g^B > 0$, $\rho_b^B = 1$. Moreover, repression never fully reveals the activist's type.*

The lemma states that whenever the type of ruler and activist match, the equilibrium behavior of the ruler is in line with the bystander's ideal: $\rho_g^G = 0$ and $\rho_b^B = 1$. It also establishes that a distortion from this ideal strategy is a part of every equilibrium: the bad ruler represses the good activist with a positive probability, $\rho_g^B > 0$. Lemma 1 does not say anything about the behavior of the good ruler toward the bad activist, ρ_b^G, which will be characterized in Proposition 2.

5 EQUILIBRIUM

If the bystander knew the types of the ruler and activist, he would join the protest if and only if the bad ruler was repressing a good activist. However, when deciding whether to join the protest, the bystander is unsure whether the good ruler repressed the bad activist, the bad ruler repressed the bad activist, or the bad ruler repressed the good activist. Thus, he faces a tradeoff: by supporting the ruler, he risks blocking beneficial changes, but by supporting the activist, he risks implementing harmful changes. The bystander uses all the available information to update his beliefs about the ruler and the activist. In particular, using Bayes' Rule:

$$p' = \Pr(\text{bad ruler}|\text{repression}) = \frac{\Pr(\text{repression} \cap \text{bad ruler})}{\Pr(\text{repression})}.$$

The bad ruler always represses the bad activist and represses the good activist with probability ρ_g^B. The probability of repression and a bad ruler is therefore $p[q + (1-q)\rho_g^B]$. The bad ruler is not the only one who represses; the good ruler also represses the bad activist with probability ρ_b^G. Thus, the probability of repression is $p[q + (1 - q)\rho_g^B] + (1 - p)q\rho_b^G$. Similar calculations for updating beliefs about the activist show:

$$(3) \qquad p' = \frac{p[q + (1 - q)\rho_g^B]}{p[q + (1 - q)\rho_g^B] + (1 - p)q\rho_b^G} \qquad q' = \frac{q[(1 - p)\rho_b^G + p]}{q[(1 - p)\rho_b^G + p] + p(1 - q)\rho_g^B}.$$

Proposition 1 highlights the key aspects of the bystander's updating.

Proposition 1 *In any equilibrium, repression causes the bystander to update negatively about <u>both</u> the ruler and the activist: $q' > q$ and $p' > p$. Moreover, holding the bad ruler's strategy fixed, when the good ruler represses the bad activist more often, the bystander updates less negatively about the ruler and more negatively about the activist: $\frac{\partial p'}{\partial \rho_b^G} < 0 < \frac{\partial q'}{\partial \rho_b^G}$. In contrast, holding the good ruler's strategy fixed, when the bad ruler represses the good activist more often, the bystander updates are the opposite: $\frac{\partial q'}{\partial \rho_g^B} < 0 < \frac{\partial p'}{\partial \rho_g^B}$.*

Proposition 1 has two implications. When the bad ruler represses the good activist more, the bystander's incentives to protest increase. However, when the good ruler represses

the bad activist more, the bystander's incentives to protest *fall*. Further, as the bystander protests more, the incentives of both the good and bad ruler to repress also fall. These underlying forces drive the equilibrium behavior.

Proposition 2 *In equilibrium, the good ruler never represses a good activist and the bad ruler always represses a bad activist, $\rho_g^G = 0$ and $\rho_b^B = 1$. There exists an increasing curve $q_1(p) \equiv \frac{(\beta + \beta_g)p}{\beta_b + \beta_g p}$ and a constant $q_2 \equiv \frac{\beta + \beta_g}{\beta_b + \beta_g}$ with $0 < q_1(p) < q_2 < 1$, such that:*

- ***Low Repression Equilibrium:*** *When the prior likelihood that the activist is bad is low, $q < q_1(p)$, a unique equilibrium exists. The good ruler never represses the bad activist, the bad ruler represses the good activist with a positive probability less than one, and upon observing repression, the bystander protests with a positive probability less than one: $\rho_b^G = 0$, $\rho_g^B = \frac{\beta_b - \beta}{\beta_g + \beta} \frac{q}{1-q}$, and $\pi = \alpha_g$.*

- ***High Repression Equilibrium:*** *When the prior likelihood that the activist is bad is high, $q > q_2$, a unique equilibrium exists. The good ruler always represses the bad activist, $\rho_b^G = 1$, the bad ruler always represses the good activist, $\rho_g^B = 1$, and the bystander never protests upon observing repression, $\pi = 0$.*

- ***Intermediate Repression Equilibrium:*** *When $q_1(p) \le q \le q_2$, the high repression and low repression equilibrium described above both exist. In addition, an equilibrium exists in which the good ruler represses the bad activist with a positive probability less than 1, the bad ruler always represses the good activist, and upon observing repression, the bystander protests with a positive probability less than 1:*

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

Figure 1 illustrates. To see the intuition, consider a simplified version of the game in which the good ruler is a non-strategic player who always represses the bad activist, i.e., $\rho_b^G = 1$. The probability with which the bad ruler represses the good activist ρ_g^B is (in equilibrium) decreasing in the bystander's likelihood of protest π. Because the bystander's
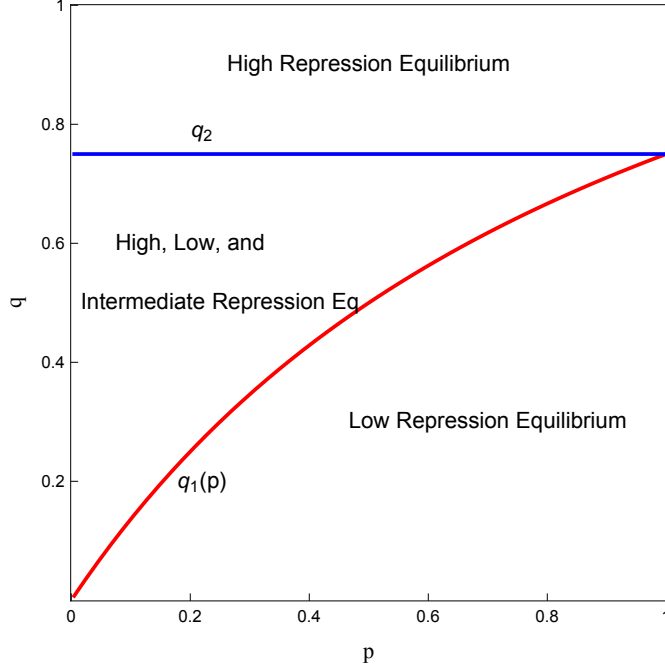
18

Figure 1: Equilibrium without restrictions.

best response π is increasing in the bad ruler's strategy ρ_g^B, a unique equilibrium exists. Next, consider a different simplified version of the game in which the bad ruler is a non-strategic player who always represses the good activist, i.e., $\rho_g^B = 1$. The good ruler's best response ρ_b^G is decreasing in the bystander's strategy π. However, unlike the previous case, the bystander's best response π is also decreasing in the good ruler's strategy ρ_b^G: when it is more likely that repression was carried out by the good ruler against the bad activist, the bystander has less incentive to protest.[17]

This structure of best responses allows for multiple equilibria because best responses can cross multiple times. In fact, it is easy to show that this simplified game has three equilibria. In one equilibrium, the bystander never protests and the good ruler always represses the bad activist. This equilibrium exists whenever $q > q_1(p)$. In another equilibrium, the bystander protests with a higher probability, $\pi = \delta_b$, and the good ruler represses the bad activist with a lower probability, $\rho_b^G = \frac{p}{1-p} \frac{(\beta + \beta_g) - (\beta_b + \beta_g)q}{\beta_b q}$. This equilibrium exists whenever $q_2 > q > q_1(p)$. In the third equilibrium, the bystander always protests and the good ruler

---

[17]Reverse the ordering of the good ruler's strategy to focus on the likelihood that the good ruler concedes to the bad activist: $1 - \rho_b^G$. Then, holding the bad ruler's strategy fixed, the best responses of the good ruler and the bystander, $1 - \rho_b^G(\pi)$ and $\pi(1 - \rho_b^G)$, are both increasing, i.e., they feature strategic complementarity.

never represses the bad activist. The first two equilibria remain even when the bad ruler acts strategically: in both of them $\pi < \alpha_g$, so it is the bad ruler's best response to always repress the good activist. However, the third equilibrium must be modified. If the bystander always protests upon observing repression, then the strategic bad ruler does not repress in equilibrium (while the non-strategic bad ruler always represses). The low repression equilibrium features the same behavior by the good ruler, but modifies the bad ruler's and the bystander's behavior to account for the bad ruler's strategic response. This logic reveals that the multiplicity of equilibria stems from the nature of the bystander's updated beliefs and the subtle interactions between the bystander and the good ruler.

Strategic forces that generate multiple equilibria suggest that social norms play a critical role in the interactions between citizens and the state. By influencing the public's and government's expectations of each other's behavior, social norms determine which equilibrium arises. These social norms have important implications, not only because equilibria exhibit different levels of repression, but also because the levels of repression in different equilibria respond differently to changes in the environment. Let $R \equiv p(q\rho_b^B + (1-q)\rho_g^B) + (1-p)(q\rho_b^G + (1-q)\rho_g^G)$ be the ex-ante expected level of repression in equilibrium.

Corollary 1 *When the public expects the status quo under a bad ruler to be worse, the expected level of repression decreases in the low repression equilibrium ($\frac{\partial R}{\partial \beta} < 0$), but it increases in the intermediate repression equilibrium ($\frac{\partial R}{\partial \beta} > 0$).*

When the status quo under the bad government is worse, the public has more incentive to protest. To offset this extra incentive to revolt, in equilibrium, either the bad ruler must repress the good activist less, or the good ruler has to repress the bad activist more. In the low repression equilibrium, the former occurs, while in the intermediate repression equilibrium, the latter occurs. Put differently, in accordance with the social norms that select the low repression equilibrium, the bystander expects the bad ruler to repress the good activist less to offset the extra incentive to revolt; similarly, in accordance with the social norms that select the intermediate repression equilibrium, the bystander expects the good ruler to repress the bad activist more to offset the extra incentive to revolt. Moreover, when the

20

public expects the status quo under a bad ruler to be worse, multiple equilibria arise in a larger area of the parameter space: $\frac{\partial[q_2 - q_1(p)]}{\partial\beta} > 0$.

To analyze the empirical implications of these results, consider the relationship between income and repression. Let G be a country's per capita GDP, and suppose that under a good regime, the bystander citizen receives G, but a bad regime secretly diverts d for its private consumption, leaving the bystander with $G - d$. Then, when the bystander's utility is concave, increases in G render the status quo under a bad ruler less harmful to the bystander: $\beta \equiv u(G) - u(G - d)$ and $\frac{\partial\beta}{\partial G} < 0$. Therefore, increases in income are associated with decreases in β, which increase repression in the low repression equilibrium, but decrease repression in the intermediate repression equilibrium.

These results suggest an explanation for the conflicting empirical findings on the relationships between income and repression and violence. For example, although some studies have found that higher income per capita reduces the likelihood of repression (Mitchell and McCormick 1988; Henderson 1991; Poe and Tate 1994), others have not found a significant correlation (Gandhi 2008; Conrad and Moore 2010; Shadmehr and Haschke 2016; see Davenport (2007) for a review).

Corollary 1 shows that increases in income can raise repression in one country, but reduce it in another, depending on the social norms that govern the state's response to dissent and the public's response to repression. Social norms can vary across countries, making the empirical results sensitive to the countries in the data set and the time span of the study, creating the potential for conflicting empirical findings.[18] In Online Appendix II, we discuss how recent developments in the estimation of games with multiple equilibria can be used to address these empirical challenges.

The preceding results are based on equilibrium multiplicity and contrasting comparative statics in low and intermediate repression equilibria. But the low repression equilibrium is unique in the particularly relevant subset of parameters where the ruler is relatively likely

---

[18]For example, consider the simplest case of a cross-section linear regression of repression on income and some controls. If all the countries in the data set (i.e., countries for which data is available) are in the low (intermediate) repression equilibrium, then one expects the coefficient for income to be negative (positive). But if some countries are in one equilibrium and others in another equilibrium, then the estimate of the coefficient can be of either sign, depending on the number of countries and the variances of income and repression in each group.

to be bad and the activist is relatively likely to be good. In the rest of this section, we focus on the low repression equilibrium and analyze its empirical implications. First, we highlight the contrasting effects of higher income on repression in our model and the literature. We then apply the insights to provide an alternative explanation for the conflicting results in the empirical literature on repression.

In our model, higher income reduces the public's incentive to protest following repression. This reduces the endogenous cost of repression, and in the low repression equilibrium, it *increases* the likelihood that the state represses the activist's initial protest (Corollary 1). This result provides a counterpoint to the common theme in the repression literature that higher income reduces repression by decreasing the likelihood of unrest. For example, Bueno de Mesquita et al. (2005, p. 447) argue that "when a state's level of economic development is low, citizens have a greater incentive to resort to conflict in order to improve their lot, and they have less to lose by doing so than when development is higher" (see also Mitchell and McCormick (1988), Henderson (1991), or Poe and Tate (1994)). In his review of the literature, Davenport (2007, p. 14) emphasizes a common rationale in the literature that "fewer resources enhance the need for coercive behavior by increasing societal grievances." However, the literature does not distinguish between the *initiation* of protest and its *spread*, ignoring the state's strategic response to changes in the likelihood that protest spreads following repression. Via the informational channel identified in our model, or the standard "opportunity costs" channel in the literature, an increase in income reduces the likelihood that protest spreads following repression. This reduces the endogenous cost of repression for the state, *increasing* the state's incentive to repress the activists.

This insight suggests an alternative explanation for the conflicting empirical findings on repression that only relies on the low repression equilibrium. To focus on the novel aspects of our model, we have assumed that the activist always protests. For purposes of this argument, assume that the activists are less likely to *initiate* protest as income increases. Then, an increase in income has two effects: (1) it reduces the likelihood that a protest is initiated, and (2) it reduces the likelihood that an initiated protest spreads following repression via both the standard and the informational channel. The first effect reduces the likelihood of repression by mitigating the state's need to repress the activist's protest; but the second

effect (identified in our model) raises the likelihood that the state uses repression by reducing its endogenous cost. Therefore, the overall effect remains ambiguous, providing another potential explanation for the conflicting empirical results in the literature.

6   LIMITS OF INTERNATIONAL PRESSURE

Our analysis so far presumes that the only force that prevents bad governments from repressing legitimate activists is the endogenous threat of protest by the public. However, international pressure sometimes makes the repression of legitimate activists prohibitively costly. Foreign governments and international institutions can pressure a regime into reducing repression by threatening to cut foreign aid, terminate trade benefits, or suspend membership in international organizations.[19]   U.S. pressure on Guatemala in the 1990s and European Union pressure on Turkey are two classic examples (Hafner-Burton 2013, Ch. 8). Moreover, regimes may inherit human rights treaties that sometimes make it too costly to repress legitimate activists, or they may join such treaties to receive foreign aid (Magesan 2013). Whether inherited or the result of a choice (which we do not explicitly model), international pressure can prevent regimes from repressing legitimate activists (Simmons 2009; Simmons and Danner 2010). However, mounting sufficient pressure requires significant resources which may be difficult to muster, especially when international actors have incentives to free-ride.[20] Thus, the international community may fail to prevent repression of legitimate activists.

To explore the effects of international pressure, we consider the following modification of our base model. Suppose that when a ruler attempts to repress the good activist, the attempt fails with (an exogenous) probability $C \in (0, 1)$ due to international pressure, in which case the ruler is forced to concede to the good activist's demands. With the remaining probability, $1 - C$, the ruler's attempt to repress the good activist succeeds. The bystander knows the likelihood C of a successful intervention, but he does not observe whether an

---

[19]Hörner, Morelli, and Squintani (2015) study the role of mediation in resolving international disputes.

[20]For example, regarding the enforcement of human rights treaties, Simmons (2009) argues that the members of international community "have incentive to ignore violations, either because they are essentially unaffected by practice elsewhere, or because other foreign policy objectives swamp the concerns they have in a particular case, or because they hope someone else will pay the costs of enforcement" (p. 126).

intervention has occurred.

The possibility that the bad ruler's repression of the good activist can fail changes how the bystander updates his beliefs from (3) to:

(4)
$$p' = \frac{p[q + (1-q)(1-C)\rho_g^B]}{p[q + (1-q)(1-C)\rho_g^B] + (1-p)q\rho_b^G} \qquad q' = \frac{q[(1-p)\rho_b^G + p]}{q[(1-p)\rho_b^G + p] + p(1-q)(1-C)\rho_g^B}.$$

When the bystander observes repression, he only assigns a probability $(1-C) \times [p(1-q)\rho_g^B]$ that a bad ruler has repressed a good activist. That is, holding the ruler's strategy fixed, increases in the strength of international pressure (i.e. increases in C) mean that repression is less likely to be carried out by a bad ruler against a good activist. Therefore, the bystander updates less negatively about the ruler and more negatively about the activist, shifting the bystander's beliefs in the ruler's favor. This effect has important implications for the equilibria, which are characterized in Proposition 3.

Proposition 3 *Suppose that the ruler's attempt to repress a good activist is blocked with 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 $g(p) \in (0,1)$ such that:*

- **Low Repression Equilibrium with International Pressure:** *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 $\pi = \alpha_g$.*

- **High Repression Equilibrium with International Pressure:** *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 Repression Equilibrium with International Pressure:** *If $g(p)\,q_1(p) \le q \le g(1)\,q_2$, then, in addition to the above two equilibria, there exists an equilibrium in which:*

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

The structure of equilibria is similar to our base model. There can be three equilibria which are ranked according to their repression levels. Because international pressure shifts
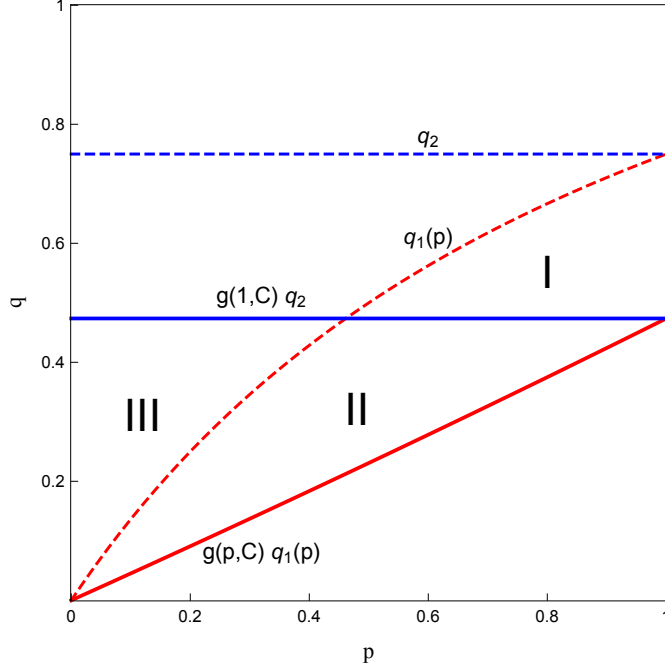
Figure 2: Equilibrium with international pressure. Region I: $g(1,C)q_2 < q < q_1(p)$. Region II: $g(p,C)q_1(p) < q < \min\{q_1(p), g(1,C)q_2\}$. Region III: $q_1(p) < q < g(1,C)q_2$.

the bystander's beliefs in the ruler's favor, international pressure reduces the bystander's incentive to join the protest, leading the bad ruler to attempt repression of the good activist more often. As Figure 2 illustrates, the set of parameters in which the high repression equilibrium arises expands from $q > q_1(p)$ in our base model to $q > g(p,C)\ q_1(p)$, where we have made the dependence of g on C explicit. Without international pressure, in regions I and II the low repression equilibrium is unique, and the bad ruler represses the good activist with probability $\rho_g^B = \frac{\beta_b - \beta}{\beta_g + \beta}\frac{q}{1-q}$. With international pressure, the high repression equilibrium is unique in region I, and all three equilibria arise in region II. In the high and intermediate repression equilibria with international pressure, the bad ruler always *attempts* to repress the good activist, succeeding with probability $1 - C$. Notably, the likelihood that the good activist is successfully repressed is *identical* in the low repression equilibrium with international pressure and in the unique equilibrium without international pressure.

Protection of Good Activists and Implementation of Good Reforms. Critically, even though under some conditions international pressure protects the good activist and raises the likelihood that good reforms are implemented, it can also generate the opposite

effects. This is most clear in region II of Figure 2, where the low repression equilibrium is unique without international pressure, but the high and intermediate repression equilibria arise with international pressure.

Corollary 2 *In region II, the likelihood that the good activist is repressed and good reforms are blocked is higher in the intermediate and high repression equilibria with international pressure than in the unique equilibrium without international pressure.*

In Corollary 2, we highlight region II because the uniqueness of the low repression equilibrium without international pressure allows for a sharp comparison. In essence, in region II, international pressure makes it possible to sustain equilibria in which the good activist is more likely to be repressed successfully, even though the international community blocks the ruler's repression attempt some of the time.[21] In particular, the likelihood that the good activist is repressed in the intermediate or high repression equilibrium with international pressure is strictly larger than that likelihood in the unique equilibrium without international pressure, $1 - C > \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}$.[22] This result resonates with a "troubling and recurrent finding" of the empirical literature, "that participation in some international treaties correlates with worse human rights behavior" (Hafner-Burton 2012, p. 280).

Of course, international pressure can also be unambiguously helpful. In particular, international pressure reduces the likelihood that the good activist is repressed above the solid horizontal line in Figure 2 where $q > g(1, C)q_2$. In particular, in region I, $1 - C < \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}$. In contrast, it does not affect this likelihood under the solid curve, where $q < g(p, C)q_1(p)$. However, even when international pressure reduces the overall likelihood that the good activist is repressed, it comes with a cost.

Corollary 3 *There exists $\hat{\alpha}_g \in (0, C)$ such that the likelihood that good reforms are blocked is higher with international pressure in region I <u>if and only if</u> $\alpha_g > \hat{\alpha}_g$.*

---

[21]In region III multiple equilibria arise, whether or not there is international pressure; the overall likelihood that the good activist is repressed is also higher in the intermediate and high repression equilibria with international pressure than in the low repression equilibrium without pressure. Therefore, even in region III it is also possible that international pressure results in a higher likelihood that a good activist is repressed.

[22]These are the likelihoods conditional on the bad ruler being in power and the activist being the good type; the corresponding ex ante likelihoods are obtained by multiplying both sides by $p(1 - q)$, which then cancel.

Corollary 3 implies that when bad rulers sufficiently dislike good reforms (so that α_g is high) in countries with relatively unpopular rulers (relatively high p) and moderately unpopular activists ($q \in (g(1, C)q_2, q_1(p))$), international pressure necessarily hinders good reforms by reducing the public's incentives to protest—even as it reduces the overall likelihood that good activists are repressed. To see the intuition, note that in region I, international pressure changes the unique equilibrium in three ways: the likelihood that the public protests following repression drops from α_g to 0, the likelihood that the bad ruler attempts to repress the good activist increases from $\frac{q}{1-q}\frac{\beta_b-\beta}{\beta_g+\beta}$ to 1, and the likelihood that the international community intervenes to prevent repression of the good activist increases from 0 to C. Because international pressure shifts the public's beliefs in favor of the ruler, it completely displaces domestic checks on repression generated by the potential for public protest, which in turn induces the bad ruler to try to repress the good activist with probability one. Therefore, with international pressure, a good reform can only be implemented due to international pressure that forces the bad ruler to concede. In contrast, in the unique (low repression) equilibrium without international pressure, a beneficial reform can be implemented *either* through a direct concession by the ruler or by a public protest. When α_g is large enough, to counter the bad ruler's large incentives to repress, the bystander protests so often that the overall likelihood of good reforms is higher without international pressure.

Strength of International Pressure (Magnitude of C). Our preceding results investigate the effects of a particular level of international pressure on different societies (with different primitive parameters). But international institutions and foreign governments can exert different *degrees* of pressure on a particular government. Thus, we study how different degrees of international pressure affect the same society. We focus on cases in which the good activists are more likely to arise ($q < q_1(p)$), where the uniqueness of equilibrium without international pressure permits sharper comparisons. Let $R_g(C)$ be the equilibrium likelihood that the good activist is successfully repressed by the bad ruler, given international pressure of degree C.[23]

---

[23]That is, if ρ_g^B comes from the bad ruler's equilibrium strategy, then $R_g(C) \equiv p(1-q)\rho_g^B(1-C)$.
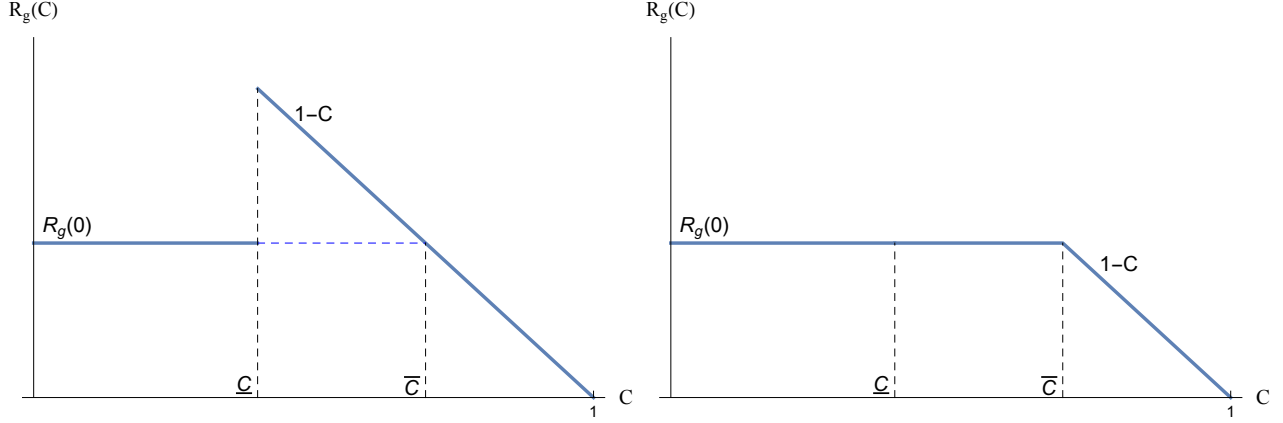
Figure 3: The likelihood that the good activist is repressed by the bad ruler $R_g(C)$ as a function of the strength of international pressure C. Right panel depicts the case in which the low repression equilibrium with international pressure is selected in region II. Left panel depicts the case in which the intermediate or high repression equilibrium with international pressure is selected in region II.

Proposition 4 *Consider a society with parameters (p, q) such that $q < q_1(p)$. International pressure decreases the likelihood that the good activist is repressed if and only if it is sufficiently strong. When it is weaker, it can __increase__ the likelihood that the good activist is repressed. Formally, there exist $0 < \underline{C} < \overline{C} < 1$ such that: (i) $R_g(C) < R_g(0)$ if and only if $C > \overline{C}$, and (ii) if $C \in (\underline{C}, \overline{C})$ and either the intermediate or high repression equilibrium is selected, then $R_g(C) > R_g(0)$.*

Figure 3 illustrates the results. The logic builds on Corollaries 2 and 3. As international pressure becomes stronger (i.e., as C increases) the lower boundaries of regions I and II, $g(1, C)q_2$ and $g(p, C)q_1(p)$, smoothly shift downward. Therefore, when C is small ($C < \underline{C}$), a society with $q < q_1(p)$ remains below region II and in the low repression equilibrium with international pressure, in which the overall likelihood that the good activist is repressed is unaffected by international pressure. As international pressure increases moderately to $C \in (\underline{C}, \overline{C})$, the society becomes part of region II where all three equilibria are possible. If the low repression equilibrium with international pressure is selected, again international pressure has no effect. However, if the intermediate or high repression equilibrium is selected, then the likelihood that the good activist is repressed is strictly higher than the likelihood

without international pressure, $1 - C > \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}$. In this case, the strategic effect of international pressure—the reduction in the public's incentives to protest—dominates its direct effect. Finally, as international pressure rises even further $(C > \overline{C})$, the society becomes part of region I, in which $1 - C < \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}$. In this case, the direct effect of international pressure dominates, and the overall likelihood that the good activist is repressed falls.[24]

These results have an immediate policy consequence. Small levels of international pressure either have no effect or an adverse effect. Thus, international pressure should be used only if it is sufficiently forceful $C > \overline{C} = 1 - \frac{\beta_b - \beta}{\beta_g + \beta}\frac{q}{1-q}$. Moreover, because $\frac{\partial \overline{C}}{\partial q} < 0 < \frac{\partial \overline{C}}{\partial \beta}$, the minimum degree of international pressure that renders it beneficial is higher when (i) the good activists are likely to arise so that the public is less suspicious of the activists (q is lower), or (ii) when the status quo under the bad ruler is worse (β is higher, corresponding to lower income in our interpretation in Corollary 1). This suggests that situations in which international pressure appears most necessary are also more prone to generate ineffective or adverse results.

7   CONCLUSION

"The seed of revolution is repression," once said Woodrow Wilson.[25] While the state's repression of activists may terminate their movement, it may also spur the general public to join the activists, leading to revolution. We develop a theoretical framework to study the interactions between the state's incentive to repress activists and the public's incentive to join the protest, and investigate its empirical implications.

Conceptual Framework: The literature has focused on settings in which citizens oppose, in thought if not in action, the state's use of coercion against dissidents. This framework presumes that all forms of state coercion are illegitimate. However, the state's raison

---

[24]The likelihood that good reforms are blocked depends on C in a similar way, except that when $C > \overline{C}$, international pressure increases the likelihood that good reforms are blocked if and only if $\alpha_g > \hat{\alpha}_g$.

[25]In his 7th annual message to the congress in 1919, he wrote: "I would call your attention to the widespread condition of political restlessness in our body politic.... Broadly, they arise...from the machinations of passionate and malevolent agitators.... The only way to keep men from agitating against grievances is to remove the grievances. An unwillingness even to discuss these matters produces only dissatisfaction and gives comfort to the extreme elements in our country which endeavor to stir up disturbances *in order to provoke governments to embark upon a course of retaliation and repression. The seed of revolution is repression* [our emphasis]."

d'être is to exercise legitimate coercion to deliver public good. An essential form of public good is to protect citizens from activists and agitators who aim to impose their harmful policies or ideological dispositions on the general public, e.g., ISIS in Syria. Critically, given the difficulties of acquiring precise information about the nature of the dissidents' demands and the intentions of the state, citizens remain uncertain about whether or not a particular instance of state coercion against dissidents was legitimate. Thus, repression can help maintain the status quo, but it can also change the public's opinion about the Leviathan and its discontents: it is informative about the nature of the beast and its challengers. Our paper provides a framework that integrates this fundamental uncertainty into the interactions between citizens and the state.

We explore three key empirical implications of this conceptual framework:

Backlash Protests: Public opinion about the state and dissidents changes following the state's use of coercion. When the general public believes that coercion was likely to have been illegitimate, backlash protest may follow repression. This result integrates the empirical regularity that repression can lead to the spread of protest, creating an endogenous cost of repression for the state.

Social Norms: Our analysis identifies a novel strategic complementarity between the state's decision to abuse its coercive force and the general public's decision to support dissidents. This complementarity generates multiple equilibria and underscores the importance of the social norms that determine the state's and citizens' expectations of each others' behavior. Depending on the prevailing social norms, variations in the environment (e.g., income) can generate opposing effects on the likelihood of repression and protest. Moreover, this framework allows one to distinguish between factors that contribute to the initiation of protests and those that contribute to the spread of protest. For example, higher income can reduce the likelihood of protest initiation, but it can also reduce the likelihood that protest spreads. This, in turn, reduces the endogenous cost of repression for the state, increasing the state's incentive to repress.

International Intervention: International pressure aimed to protect legitimate dissidents from state repression can have the opposite of its intended effect. Because interna-

tional pressure reduces the probability of illegitimate coercion, public opinion responds less adversely to the state's use of coercion, thereby reducing the public's incentive to protest. Because the mechanism is informational, this crowding-out of the domestic response can dominate the direct effect of international pressure, thereby raising the overall level of repression. A simple policy implication is that lukewarm international pressure is at best ineffective, and at worst, increases repression by displacing domestic checks on the state.

We end by highlighting three directions for future research. First, our analysis is in a static setting, where the prior beliefs are exogenously given from past interactions and events. One can endogenize these beliefs by analyzing a long-term interaction, in which citizens' beliefs in each period are affected by the state's past behavior. Second, our analysis assumes only one form of coercion. However, coercive force can be exercised in various manners that may be correlated with the nature of the state. For example, using rubber versus metal bullets on aggressive protestors, or blocking demonstrators by riot shields versus using pepper stray or tear gas may have different effects on public opinion. It may be easier to disperse protesters if the state uses an iron fist, but such an excessive use of force will also create a critical question for the general public: what kind of state is likely to use that kind of coercion? How the public answers this question may be critical for the state's collapse or survival. Finally, our analysis treats the general public as a unitary actor. Disaggregating the general public into a mass of citizens with different preferences or beliefs might yield additional insight into the information content of repression and its link to backlash protest.

8   REFERENCES

Aytaç, Erdem, Luis Schiumerini, and Susan Stokes. 2018. "Why Do People Join Backlash Protests? Lessons from Turkey." *Journal of Conflict Resolution* 62: 1205-28.

Aytaç, Erdem, and Susan Stokes. 2018 (forthcoming). *Why Bother?: Rethinking Participation in Elections and Protests.* New York, NY: Cambridge University Press.

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

Acemoglu, Daron, and Matthew O. Jackson. 2015. "History, Expectations, and Leadership in the Evolution of Social Norms." *Review of Economic Studies* 82: 1-34.

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.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2013. "A Political Theory of Populism." *Quarterly Journal of Economics* 128: 771-805.

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

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

Beissinger, Mark, Amaney A Jamal, and Kevin Mazur. 2015. "Explaining Divergent Revolutionary Coalitions: Regime Strategies and the Structuring of Participation in the Tunisian and Egyptian Revolutions." *Comparative Politics* 48: 1-24.

Besley, Timothy, and Torsten Persson. 2011. "The Logic of Political Violence." *Quarterly Journal of Economics* 126: 1411-45.

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.

Bueno de Mesquita, Bruce, George W. Downs, Alastair Smith, and Feryal Marie Cherif. 2005. "Thinking Inside the Box: A Closer Look at Democracy and Human Rights." *International Studies Quarterly* 49: 439-57.

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

Bueno de Mesquita, Ethan. 2013. "Rebel Tactics." *Journal of Political Economy* 121: 323-57.

Bueno de Mesquita, Ethan, Eric S. Dickson. 2007. "The Propaganda of the Deed: Terrorism, Counterterrorism, and Mobilization." *American Journal of Political Science* 51: 364-81.

Burke, Mary, and Peyton Young. 2011. "Social Norms." *Handbook of Social Economics, Volume 1A*, edited by Jess Benhabib, Alberto Bisin, and Matthew Jackson. Amsterdam: North-Holland.

Cammett, Melani C. 2011. "Partisan Activism and Access to Welfare in Lebanon." *Studies in Comparative International Development* 46: 70-97.

Canes-Wrone, Brandice, Michael Herron, and Kenneth Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45: 532-50.

Casper, Brett, and Scott Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'etat." *Journal of Politics* 76: 548-64.

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

Conrad, Courtenay, and Will Moore. 2010. "What Stops the Torture?" *American Journal of Political Science* 54: 459-76.

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

Diani, Mario, and Doug McAdam. 2003. *Social Movements and Networks: Relational Approach to Collective Action.* New York: Oxford 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, and Konstantin Sonin. 2015. "Incumbency Advantage in Non-Democracies." Mimeo.

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.

Forero, Juan. 2014. "Venezuela Divided One Year After Chavez's Death; Supporters Com-

memorate Hugo Chavez while Opponents Protest." *Wall Street Journal (Online)*, Mar 05.

Fox, Justin. 2007. "Government Transparency and Policymaking." *Public Choice* 131: 23-44.

Fox, Justin, and Matthew Stephenson. 2011. "Judicial Review as a Response to Political Posturing." *American Political Science Review* 105: 397-414.

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.

Gamson, William A. 1975. *The Strategy of Social Protest.* Homewood, IL: Dorsey.

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

Ginkel, John, and Alastair Smith. 1999. "So You Say You Want a Revolution: A Game Theoretic Explanation of Revolution in Repressive Regimes." *Journal of Conflict Resolution* 43: 291-316.

Guriev, Sergei, and Daniel Treisman. 2015. "How Modern Dictators Survive: An Informational Theory of the New Authoritarianism." Mimeo.

Hafner-Burton, Emilie. 2012. "International Regimes for Human Rights." *Annual Review of Political Science* 15: 265-86.

Hafner-Burton, Emilie. 2013. *Making Human Rights a Reality.* Princeton, NJ: Princeton University Press.

Hafner-Burton, Emilie, and Kiyoteru Tsutsui. 2005. "Human Rights in a Globalizing World: The Paradox of Empty Promises." *American Journal of Sociology* 110: 1373-411.

Hathaway, Oona. 2002. "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111: 1935-2042.

Henderson, Conway. 1991. "Conditions Affecting the Use of Political Repression." *Journal of Conflict Resolution* 35: 120-42.

Hollyer, James, and Peter Rosendorff. 2011. "Why Do Authoritarian Regimes Sign the Convention Against Torture? Signaling, Domestic Politics, and Non-Compliance." *Quar-*

terly Journal of Political Science 6: 275-327.

Hörner, Johannes, Massimo Morelli, and Francesco Squintani. 2015. "Mediation and Peace." *Review of Economic Studies* 82: 1483-1501.

Huang, Haifeng, Serra Boranbay, and Ling Huang. 2018. "Media, Protest Diffusion, and Authoritarian Resilience." *Political Science Research and Methods. Forthcoming.*

Jamal, Amaney A. 2012. *Of Empires and Citizens: Pro-American Democracy or No Democracy at All?* Princeton, NJ: Princeton University Press.

Khatib, Lina, and Ellen Lust. 2014. *Taking to the Streets: The Transformation of Arab Activism.* Baltimore, MD: Johns Hopkins University Press.

KONDA. 2014. "GEZI REPORT: Public Perception of the Gezi Protests. Who Were the People at Gezi Park?" Available at: http://konda.com.tr/en/raporlar/KONDA_Gezi_Report.pdf

Kreps, David M. 1990. "Corporate Culture and Economic Theory," in J. E. Alt and K. A. Shepsle, eds., *Perspectives on Positive Political Economy*, Cambridge, England: Cambridge University Press, 1990.

Kurzman, Charles. 2003. "The Qum Protests and the Coming of the Iranian Revolution, 1975 and 1978." *Social Science History* 27: 287-325.

Levitsky, Steven, and Daniel Ziblatt. 2018. *How Democracies Die.* New York, NY: Crown.

Little, Andrew. 2012. "Elections, Fraud, and Election Monitoring in the Shadow of Revolution." *Quarterly Journal of Political Science* 7: 249-83.

Lou, Zhaotian, and Arturas Rozenas. 2018. "The Election Monitor's Curse." *American Journal of Political Science* 62: 148-60.

Magesan, Arvind. 2013. "Human Rights Treaty Ratification of Aid Receiving Countries." *World Development* 45: 175-88.

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.

Martin, Brian. 2007. *Justice Ignited: The Dynamics of Backfire.* Lanham, MD: Rowman & Littlefield.

Maskin, Eric, and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94: 1034-54.

McAdam, Doug. 1999 *Political Process and the Development of Black Insurgency, 1930-1970.* Chicago, IL: University of Chicago Press.

McAdam, Doug, Sydney Tarrow, and Charles Tilly. 2001. *Dynamics of Contention.* New York: Cambridge University Press.

McCarthy John D., and Mayer N. Zald. 1977. "Resource Mobilization and Social Movements: A Partial Theory." *American Journal of Sociology* 82: 1212-41.

Minaya, Ezequiel. 2014a. "Venezuelan Opposition Leader Says He Will Risk Arrest; Leopoldo Lopez Says He Plans to March with Antigovernment Protesters Tuesday." *Wall Street Journal (Online)*, Feb 16.

Minaya, Ezequiel. 2014b. "Protests Against Venezuela's Government Escalate; Thousands Burned Tires, Cars as Security Forces Fought to Control the Streets in Caracas, Other Cities." *Wall Street Journal (Online)*, Feb 20.

Mitchell, Neil, and James McCormick. 1988. "Economic and Political Explanations of Human Rights Violations." *World Politics* 40: 476-98.

Morris, Aldon D. 1984. *The Origins of the Civil Rights Movement: Black Communities Organizing for Change.* New York: The Free Press.

Morris, Stephen. 2001. "Political Correctness." *Journal of Political Economy* 109: 231-65.

Morris, Stephen, and Mehdi Shadmehr. 2017. "Inspiring Regime Change." Mimeo.

Myerson, Roger. 1991. *Game Theory: Analysis of Conflict.* Boston, MA: Harvard 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.

Postlewaite, Andrew. 2011. "Social Norms and Preferences." *Handbook of Social Economics, Volume 1A,* edited by J. Benhabib, A. Bisin, and M. Jackson. Amsterdam: North-Holland.

Prato, Carlo and Bruno Strulovici. 2013. "Direct Democracy and Responsibility Substitution." Mimeo.

Simmons, Beth. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics.* New York, NY: Cambridge University Press.

Simmons Beth, and Allison Danner. 2010. "Credible Commitments and the International Criminal Court." *International Organization* 64: 225-56.

Shadmehr, Mehdi. 2014. "Mobilization, Repression, and Revolution: Grievances and Opportunities in Contentious Politics." *Journal of Politics* 76: 621-35.

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.

Shadmehr, Mehdi, and Dan Bernhardt. 2019. "Vanguards in Revolution." *Games and Economic Behavior* 115: 146-66.

Shadmehr, Mehdi, and Peter Haschke. 2016. "Youth, Revolution and Repression." *Economic Inquiry* 54: 778-93.

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

Smith-Cannoy, Heather. 2012. *Insincere Commitments: Human Rights Treaties, Abusive States, and Citizen Activism.* Washington DC: Georgetown University Press

Sobel, Joel. 1985. "A Theory of Credibility." *Review of Economic Studies* 52: 557-73.

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

Tarrow, Sydney. 1998. *Power in Movement: Social Movements and Contentious Politics.* New York, NY: Cambridge University Press.

Tilly, Charles. 1978. *From Mobilization to Revolution.* Reading, MA: Addison-Wesley.

Tilly, Charles. 2004. *Social Movements: 1768-2004.* Boulder, CO: Paradigm.

Tyson, Scott, and Alastair Smith. 2014. "Two-Sided Coordination: Competition in Collective Action." Mimeo.

Vyas, Kejal and Angel Gonzalez. 2013. "Venezuela Opens Investigation of Opposition Leader." *Wall Street Journal (Online)*, Apr 25.

Wantchékon, Léonard, and Omar García-Ponce. 2014. "Critical Junctures: Independence Movements and Democracy in Africa." Mimeo.

Wood, Elisabeth J. 2003. *Insurgent Collective Action and Civil War in El Salvador.* New York: Cambridge University Press.

Wood, Gordon. 2009. *Empire of Liberty: A History of the Early Republic, 1789-1815.* New York, NY: Oxford University Press.

Young, Peyton. 2015. "The Evolution of Social Norms." *Annual Review of Economics 7: 359-87.*

For Online Publication

1 APPENDIX I: PROOFS

Proof of Remark 2. Let $\eta \in [0, 1]$ be the probability that the bystander protests. From equation (1), any value of $\eta \in [0, 1]$ is supported by some combination of beliefs. If a bad ruler faces a good activist he represses whenever $\eta < \alpha_g$, but if a bad ruler faces a bad activist he represses whenever $\eta < \alpha_b$. Because $\alpha_g < \alpha_b$, the belief that the bad ruler is repressing the good activist violates D1. $\blacksquare$

1.1 PROOFS FOR SECTION 5: NO LIMITS ON REPRESSION

Proof of Proposition 1. We prove that $q' > q$; other claims follow from (3) by simple algebraic manipulations and differentiations and hence are omitted. From (3),

$$q' - q = q(1 - q)\frac{p(1 - \rho_g^B) + (1 - p)\rho_b^G}{q[(1 - p)\rho_b^G + p] + p(1 - q)\rho_g^B},$$

and hence $q' \geq q$. We show that the inequality is generically strict. If $q' = q$, then we must have $\rho_g^B = 1$ and $\rho_b^G = 0$. Hence, (2) requires that $\pi \in [\delta_b, \alpha_g] \subset (0, 1)$. Hence, (1) requires that $\beta_g(1 - q') - \beta_b q' = -\beta p'$. Substituting $\rho_g^B = 1$ and $\rho_b^G = 0$ into (3) gives $p' = 1$. Hence, $q' = q$ implies $\beta_g(1 - q) - \beta_b q = -\beta$, which is a knife edge case. $\blacksquare$

Proof of Proposition 2. From Lemma 1, $\rho_g^B > 0$. Hence, equation (2) implies that $\pi \leq \alpha_g$. An equilibrium with $\pi \in (\delta_b, \alpha_g)$ is generically impossible. If $\pi \in (\delta_b, \alpha_g)$, then $\rho_b^G = 0$ and $\rho_g^B = 1$. Hence, $p' = 1$ and $q' = q$. Because $\pi \in (0, 1)$, equation (1) implies that $\beta_g(1 - q) - \beta_b q = -\beta$, but this is non-generic. Similarly, an equilibrium with $\pi \in (0, \delta_b)$ is generically impossible. Hence, there are only three possibilities: $\pi = \alpha_g$, $\pi = \delta_b$, and $\pi = 0$.

Suppose that $\pi = \alpha_g$. Because $\delta_b < \alpha_g$, equation (2) implies that $\rho_b^G = 0$ and $\rho_g^B \in [0, 1]$, and hence (3) implies:

$$q' = \frac{q}{q + (1 - q)\rho_g^B} \qquad p' = 1.$$

Because $\pi \in (0, 1)$, equation (1) implies that $\beta_g(1 - q') - \beta_b q' = -\beta p'$. Substituting (p', q')

gives:

$$\beta_g(1 - \frac{q}{q + (1-q)\rho_g^B}) - \beta_b \frac{q}{q + (1-q)\rho_g^B} = -\beta \iff \rho_g^B = \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta},$$

and hence, $\rho_g^B \in [0,1] \Leftrightarrow q \leq \frac{\beta + \beta_g}{\beta_b + \beta_g}$. Hence, an equilibrium with $\pi = \alpha_g$, $\rho_b^G = 0$, and $\rho_g^B = \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}$ exists if and only if $q \leq q_2$.

Suppose that $\pi = \delta_b$. Because $\delta_b < \alpha_g$, equation (2) implies that $\rho_b^G \in [0,1]$ and $\rho_g^B = 1$. Hence, (3) implies:

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

Because $\pi \in (0,1)$, equation (1) implies that $\beta_g(1 - q') - \beta_b q' = -\beta p'$. Substituting (p', q') gives:

$$\beta_g(1 - \frac{q[(1-p)\rho_b^G + p]}{q[(1-p)\rho_b^G + p] + p(1-q)}) - \beta_b \frac{q[(1-p)\rho_b^G + p]}{q[(1-p)\rho_b^G + p] + p(1-q)} = -\beta \frac{p}{p + (1-p)q\rho_b^G},$$

which implies:

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

Hence, $\rho_g^B \in [0,1] \Leftrightarrow \frac{(\beta + \beta_g)p}{\beta_b + \beta_g p} \leq q \leq \frac{\beta + \beta_g}{\beta_b + \beta_g}$, where we recognize that $\frac{(\beta + \beta_g)p}{\beta_b + \beta_g p} < \frac{\beta + \beta_g}{\beta_b + \beta_g}$ for $p < 1$. Hence, an equilibrium with $\pi = \delta_b$, $\rho_b^G = \frac{p}{1-p}\frac{(\beta + \beta_g) - (\beta_b + \beta_g)q}{\beta_b q}$, and $\rho_g^B = 1$ exists if and only if $q_1(p) \leq q \leq q_2$.

Suppose that $\pi = 0$. Equation (2) implies that $\rho_b^G = \rho_g^B = 1$. Hence, (3) implies:

$$q' = \frac{q}{q + p(1-q)} \qquad p' = \frac{p}{p + (1-p)q}.$$

Because $\pi = 0$, equation (1) implies that $\beta_g(1 - q') - \beta_b q' \leq -\beta p'$. Substituting (p', q') gives:

$$\beta_g(1 - \frac{q}{q + p(1-q)}) - \beta_b \frac{q}{q + p(1-q)} \leq -\beta \frac{p}{p + (1-p)q} \iff q \geq \frac{(\beta + \beta_g)p}{\beta_b + \beta_g p}.$$

Hence, an equilibrium with $\pi = 0$, $\rho_b^G = \rho_g^B = 1$ exists if and only if $q \geq q_1(p)$. ∎

1.2 PROOFS FOR SECTION 6: EXOGENOUS LIMITS ON REPRESSION

Proof of Proposition 3. Let $g(p) \equiv 1 - C\frac{\beta_b - p\beta}{\beta_b - p\beta + (1-C)p(\beta + \beta_g)}$. Note that $g(p) \in (0,1)$ for $C \in (0,1)$, and that $g'(p) = C(1-C)\frac{\beta_b(\beta + \beta_g)}{(\beta_b - p\beta + (1-C)p(\beta + \beta_g))^2} > 0$.

First, observe that the best responses of the bystander and the ruler are still described by equations (1) and (2). Hence, Lemma 1 applies. Hence, equation (2) implies that $\pi \leq \alpha_g$. An equilibrium with $\pi \in (\delta_b, \alpha_g)$ is generically impossible. If $\pi \in (\delta_b, \alpha_g)$, then $\rho_b^G = 0$ and $\rho_g^B = 1$. Hence, $p' = 1$ and $q' = q/(q + (1-q)(1-C))$. Because $\pi \in (0,1)$, equation (1) implies that $\beta_g(1-q') - \beta_b q' = -\beta$, but this is non-generic. Similarly, an equilibrium with $\pi \in (0, \delta_b)$ is generically impossible. Hence, there are only three possibilities: $\pi = \alpha_g$, $\pi = \delta_b$, and $\pi = 0$.

Suppose that $\pi = \alpha_g$. Because $\delta_b < \alpha_g$, equation (2) implies that $\rho_b^G = 0$ and $\rho_g^B \in [0,1]$, and hence (4) implies:

$$q' = \frac{q}{q + (1-q)(1-C)\rho_g^B} \qquad p' = 1.$$

Because $\pi \in (0,1)$, equation (1) implies that $\beta_g(1-q') - \beta_b q' = -\beta p'$. Substituting (p', q') gives:

$$\beta_g\left(1 - \frac{q}{q + (1-q)(1-C)\rho_g^B}\right) - \beta_b\frac{q}{q + (1-q)(1-C)\rho_g^B} = -\beta \iff \rho_g^B = \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}\frac{1}{1-C},$$

and hence, $\rho_g^B \in [0,1] \iff q \leq \frac{\beta + \beta_g}{\beta_b + \beta_g}\left(1 - C\frac{\beta_b - \beta}{\beta_b - \beta + (1-C)(\beta + \beta_g)}\right) = q_2 \times g(1)$. Hence, an equilibrium with $\pi = \alpha_g$, $\rho_b^G = 0$, and $\rho_g^B = \frac{q}{1-q}\frac{\beta_b - \beta}{\beta_g + \beta}\frac{1}{1-C}$ exists if and only if $q \leq g(1)q_2$.

Suppose that $\pi = \delta_b$. Because $\delta_b < \alpha_g$, equation (2) implies that $\rho_b^G \in [0,1]$ and $\rho_g^B = 1$. Hence, (4) implies:

$$q' = \frac{q[(1-p)\rho_b^G + p]}{q[(1-p)\rho_b^G + p] + p(1-q)(1-C)} \qquad p' = \frac{p[q + (1-q)(1-C)]}{p[q + (1-q)(1-C)] + (1-p)q\rho_b^G}.$$

Because $\pi \in (0,1)$, equation (1) implies that $\beta_g(1-q') - \beta_b q' = -\beta p'$. Substituting (p', q')

gives:

$$
(5) \qquad \beta_g \left(1 - \frac{q[(1-p)\rho_b^G + p]}{q[(1-p)\rho_b^G + p] + p(1-q)(1-C)} \right) - \beta_b \frac{q[(1-p)\rho_b^G + p]}{q[(1-p)\rho_b^G + p] + p(1-q)(1-C)} =
$$
$$
-\beta \frac{p[q + (1-q)(1-C)]}{p[q + (1-q)(1-C)] + (1-p)q\rho_b^G},
$$

which implies:

$$
\rho_b^G = \frac{p}{1-p} \left(\frac{1-q}{q} \frac{\beta + \beta_g}{\beta_b}(1-C) - \frac{\beta_b - \beta}{\beta_b} \right).
$$

Hence,

$$
\rho_g^B \in [0,1] \Leftrightarrow
$$
$$
\frac{(\beta + \beta_g)p}{\beta_b + \beta_g p} \left(1 - \frac{C(\beta_b - p\beta)}{\beta_b - p\beta + (1-C)p(\beta + \beta_g)} \right) \leq q \leq \frac{\beta + \beta_g}{\beta_b + \beta_g} \left(1 - \frac{C(\beta_b - \beta)}{\beta_b - \beta + (1-C)(\beta + \beta_g)} \right),
$$

or equivalently, $q_1(p)g(p) \leq q \leq q_2 g(1)$, where we recognize that $q_1(p)g(p) < q_2 g(1)$ for $p < 1$. Hence, an equilibrium with $\pi = \delta_b$, $\rho_b^G = \frac{p}{1-p} \left(\frac{1-q}{q} \frac{\beta + \beta_g}{\beta_b}(1-C) - \frac{\beta_b - \beta}{\beta_b} \right)$, and $\rho_g^B = 1$ exists if and only if $g(p)\, q_1(p) \leq q \leq g(1)\, q_2$.

Suppose that $\pi = 0$. Equation (2) implies that $\rho_b^G = 1$ and $\rho_g^B = 1$. Hence, (4) implies:

$$
q' = \frac{q}{q + p(1-q)(1-C)} \qquad p' = \frac{p[q + (1-q)(1-C)]}{p[q + (1-q)(1-C)] + (1-p)q}.
$$

Because $\pi = 0$, equation (1) implies that $\beta_g(1-q') - \beta_b q' \leq -\beta p'$. Substituting (p', q') gives:

$$
\beta_g(1 - \frac{q}{q + p(1-q)(1-C)}) - \beta_b \frac{q}{q + p(1-q)(1-C)} \leq -\beta \frac{p[q + (1-q)(1-C)]}{p[q + (1-q)(1-C)] + (1-p)q},
$$

which holds if and only if $q \geq q_1(p)g(p)$. Hence, an equilibrium with $\pi = 0$, $\rho_b^G = 1$ and $\rho_g^B = 1$ exists if and only if $q \geq q_1(p)g(p)$. $\blacksquare$

Proof of Corollary 2. In the intermediate and high repression equilibria with international pressure, from Proposition 3, the likelihood that the activist is repressed is $p(1-q)(1-C)$. Absent international pressure, from Proposition 2, in the low repression equilibrium, the

corresponding likelihood is $p(1-q)\rho_g^B = p(1-q)\frac{\beta_b-\beta}{\beta_g+\beta}\frac{q}{1-q}$. Moreover,

$$\frac{\beta_b-\beta}{\beta_g+\beta}\frac{q}{1-q} < 1-C \text{ if and only if } q < \frac{(1-C)(\beta_g+\beta)}{(1-C)\beta_g+\beta_b-C\beta} = g(1)q_2.$$

Further, the likelihood that the bystander protests is 0 in the high repression equilibrium with international pressure, δ_b in the intermediate repression equilibrium with international pressure, and $\alpha_g > \delta_b$ in the low repression equilibrium (without pressure). The result follows. ∎

Proof of Corollary 3. Observe that, without international pressure, the low repression equilibrium is unique in region I. In this equilibrium, the good reform is blocked with probability $p(1-q)\rho_g^B(1-\alpha_g) = p(1-q)\frac{\beta_b-\beta}{\beta_g+\beta}\frac{q}{1-q}(1-\alpha_g)$. Moreover, the high repression equilibrium with international pressure is also unique in region I. In this equilibrium the good reform is blocked with probability $p(1-q)(1-C)$. Further, the boundaries of region I, $g(1)q_2$ and $q_1(p)$, do not depend on α_g. Thus, in region I, the likelihood that good reforms are blocked is higher with international pressure if and only if:

$$(6) \qquad \frac{\beta_b-\beta}{\beta_g+\beta}\frac{q}{1-q}(1-\alpha_g) < 1-C.$$

From the proof of Corollary 2, for $\alpha_g = 0$, $\frac{\beta_b-\beta}{\beta_g+\beta}\frac{q}{1-q} > 1-C$ if and only if $q > g(1)q_2$, which includes region I. Further, the left hand side of (6) is strictly decreasing in α_g, with $\rho_g^B(1-C) < 1-C$ at $\alpha_g = C$. Thus, there exists $\hat{\alpha}_g \in (0,C)$ such that $\frac{\beta_b-\beta}{\beta_g+\beta}\frac{q}{1-q}(1-\alpha_g) < 1-C \Leftrightarrow \alpha_g > \hat{\alpha}_g$. ∎

Proof of Proposition 4: The results follow from the following Lemma together with our discussion in the text.

Lemma 2 *Consider a society with parameters (p,q) such that $q < q_1(p)$. There exists $0 < \underline{C}(p,q) < \overline{C}(p,q) < 1$ with $p(1-q)(1-\underline{C}) > R_g(0)$ and $p(1-q)(1-\overline{C}) = R_g(0)$ such that:*

- *If either intermediate or high repression equilibrium with international pressure is se-*

5

lected in region II, then

$$R_g(C) = \begin{cases} R_g(0) & \text{if } C < \underline{C} \\ p(1-q)(1-C) & \text{if } C > \underline{C}. \end{cases}$$

- *If the low repression equilibrium with international pressure is selected in region II, then*

$$R_g(C) = \begin{cases} R_g(0) & \text{if } C \leq \overline{C} \\ p(1-q)(1-C) & \text{if } C \geq \overline{C}. \end{cases}$$

Proof of Lemma 2. From the proof of Proposition 3, $g(p,C) = 1 - C\frac{\beta_b - p\beta}{\beta_b - p\beta + (1-C)p(\beta + \beta_g)}$, and hence $\frac{\partial g(p,C)}{\partial C} = -\frac{(\beta_b - p\beta)(\beta_b + p\beta_g)}{(\beta_b + p\beta_g - Cp(\beta + \beta_g))^2} < 0$. That is, $g(p,C)$ is strictly decreasing in C on $C \in [0,1]$ and for all $p \in [0,1]$, with $g(p,0) = 1$ and $g(p,1) = 0$. Thus, for a given p, if $q < q_1(p)$, (i) there exists a unique $\underline{C}$ that satisfies $g(p,\underline{C})q_1(p) = q$, and (ii) there exists a unique $\overline{C}$ that satisfies $g(1,\overline{C})q_2 = g(1,\overline{C})q_1(1) = q$, and (iii) $0 < \underline{C} < \overline{C} < 1$. Thus, given a p we have: $q < g(p,C)q_1(p)$ for $C < \underline{C}$; $q = g(p,\underline{C})q_1(p)$; $g(p,C)q_1(p) < q < g(1,C)q_1(1)$ for $\underline{C} < C < \overline{C}$; $q = g(1,\overline{C})q_1(1)$; and $q > g(1,C)q_1(1)$ for $C > \overline{C}$. Thus, a (p,q) with $q < q_1(p)$ is below region II when $C \in (0,\underline{C})$, in region II when $C \in (\underline{C},\overline{C})$, and in region I when $C \in (\overline{C},1)$.

From Proposition 3, in the intermediate and high repression equilibrium with international pressure, the likelihood that the good activist is repressed $R_g(C)$ is $p(1-q) \times (1-C)$; and in the low repression equilibrium with international pressure, that likelihood is $p(1-q) \times \frac{\beta_b - \beta}{\beta_g + \beta}\frac{q}{1-q}\frac{1}{1-C} \times (1-C) = p(1-q) \times \frac{\beta_b - \beta}{\beta_g + \beta}\frac{q}{1-q}$, which is the same as the likelihood that the good activist is repressed in the low repression equilibrium without international pressure $R_g(0)$. Moreover, from the proof of Proposition 3, $\frac{\beta_b - \beta}{\beta_g + \beta}\frac{q}{1-q}\frac{1}{1-C}$ is strictly increasing in q for $q < g(1,C)q_2$, and becomes 1 at $q = g(1,C)q_2$. Hence, $R_g(0) < p(1-q)(1-C)$ for $q < g(1,C)q_2$, and $R_g(0) = p(1-q)(1-C)$ at $q = g(1,C)q_2$, i.e., as C increases $p(1-q)(1-C)$ crosses $R_g(0)$ from above at $\overline{C}$. ∎

2 APPENDIX II: ESTIMATION OF GAMES WITH MULTIPLE EQUILIBRIA

Recent developments in the estimation of incomplete information games with private information provide a fruitful and practical approach to tackle empirical challenges in studying the determinants of repression, which arise due to multiple equilibria—see de Paula (2013) and Borkovsky et al. (2015) for reviews.[26] To illustrate, we briefly discuss estimation approaches used in two different settings. Multiple equilibira arise in Sweeting's (2009) model of radio stations' decisions of when to air commercials, where stations have private information and are concerned about the consumers' decision to switch station. To estimate the game, he uses a *finite mixture model*, "where the components [of the likelihood function] are outcome distributions conditional on a realized equilibrium and the equilibrium selection mechanism characterizes the mixture weights" (Henry et al. 2014, p. 124). Sweeting (2009) chooses constant weights, and estimates the weights along with other parameters of the model to maximize the likelihood function, showing a significant degree of strategic complementarities between the stations' equilibrium strategies.[27]

Bisin, Moro, and Topa (2011) study high school students' decisions to smoke in the presence of pressure for conformity. Similar to our analysis, strategic complementarities between the students' decisions generate three equilibria, which are ranked according to the level of smoking in school into high, intermediate, and low equilibira, with opposing comparative statics. They develop a two-step estimation method and combine it with a random coefficient specification to estimate the model.[28] They show that almost all schools in their sample are in the low or intermediate smoking equilibrium, with a majority being in the intermediate one. Critically, their empirical finding shows that "depending on whether a given school is

---

[26]More broadly, empirical analyses of models with multiple equilibria are used to study a wide range of economic topics, including macroeconomic fluctuations (Dagsvik and Jovanovic 1994), entry games in oligopolistic settings (Bresnahan and Reiss 1991; Berry and Tamer 2007), and wage discrimination in labor markets (Moro 2003).

[27]See Grieco (2014) for a finite mixture model with more flexible weights, and its application to firm entry in local grocery markets, where firms have private information about local demands and their costs.

[28]Their computationally heavy *direct estimation method* consists of first computing all the equilibria for a set of parameters (to be estimated) and their associated likelihood values, and then choosing the equilibrium that maximizes the likelihood correspondence.

in the low or the intermediate equilibrium, the same reduction in social interactions may increase or decrease the equilibrium level of smoking (assuming the school stays in the same equilibrium). This is in contrast to simple repression analyses that did not consider the existence of multiple equilibria" (p. 39). For example, "if the adoption of rules for tobacco use by school staff reduces the strength of school-wide social interactions, such a policy may have the unintended consequence—in some schools—of actually increasing smoking prevalence in the school" (p. 40).

A less demanding approach is to use *random coefficient models* to allow different countries to have different coefficients (with potentially opposite signs) for covariates such as income or inequality. The distribution of slopes can then be estimated by a variety of techniques (e.g., Fox et al. 2011). In this approach, one assumes that the equilibrium played between the government and the public typically remains the same within a country or a region, with potential changes in equilibrium allowed only after major event such as regime changes or wars. As de Paula (2013) argues in his review of econometric analysis of games, "if an equilibrium is established as a mode of behavior by past play, custom, or culture, this equilibrium becomes a focal point for those involved. When observed games are drawn from a population that is culturally or geographically close, sharing similar norms and conventions, one would expect this assumption to be adequate" (p. 120). Random coefficient models are common in economics and can be implemented using standard software (Cameron and Trivedi 2005, Ch. 22).[29] Finally, an ad hoc approach would be to "infer" the equilibrium by observing how repression changes with income or inequality in a particular country in the near past; for example, if increases in income have increased the likelihood of repression in the recent past, then this implies that the country is in the low repression equilibrium. It is worth mentioning that Besley and Persson (2011) emphasize the use of fixed-effects in their ordered logit estimation of the causes of violence to circumvent bias due to country-specific unobservables. The approaches presented here are more demanding as they address identification and estimations issues that arise due to multiple equilibria, which typically cannot be remedied by standard fixed-effect approaches.

---

[29]Recent applications in political economy include the study of vote swings in elections (Gelman et al. 2016).

2.1 REFERENCES

Bisin, Alberto, Andrea Moro, and Giorgio Topa. 2011. "Empirical Content of Models with Multiple Equilibria in Economies with Social Interactions." NBER Working Paper 17196.

Berry, Steve, and Elie Tamer. 2007. "Identification in Models of Oligopoly Entry." In *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, vol. II,* ed. by R. Blundell, W. Newey, and T. Persson. Ch. 2, 46-85. Cambridge: Cambridge University Press.

Besley, Timothy, and Torsten Persson. 2011. "The Logic of Political Violence." *Quarterly Journal of Economics* 126: 1411-45.

Borkovsky, Ron N., Paul B. Ellickson, Brett R. Gordon, Victor Aguirregabiria, Pedro Gardete, Paul Grieco, Todd Gureckis, Teck-Hua Ho, Laurent Mathevet, Andrew Sweeting. 2015. "Multiplicity of Equilibria and Information Structures in Empirical Games: Challenges and Prospects." *Marketing Letters* 26: 115-25.

Bresnahan, Timothy F., and Peter C. Reiss. 1991. "Empirical Models of Discrete Games." *Journal of Econometrics* 64: 57-81.

Cameron, A. Colin, and Pravin K. Trivedi 2009. *Microeconometrics: Methods and Applications.* New York: Cambridge University Press.

Dagsvik, John, and Boyan Jovanovic. 1994. "Was the Great Depression a Low-Level Equilibrium?" *European Economic Review* 38: 1711-29.

de Paula, Áureo. 2013. "Econometric Analysis of Games with Multiple Equilibria." *Annual Review of Economics* 5: 107-31.

Fox, Jeremy T., Stephen P. Ryan, and Patrick Bajari. 2011. "A Simple Estimator for the Distribution of Random Coefficients." *Quantitative Economics* 2: 381-418.

Gelman, Andrew, Sharad Goel, Douglas Rivers, and David Rothschild. 2016. "The Mythical Swing Voter." *Quarterly Journal of Political Science* 11: 103-30.

Grieco, Paul L. 2014. "Discrete Games with Flexible Information Structures: An Application to Local Grocery Markets." *RAND Journal of Economics* 45: 303-40.

Henry, Marc, Yuichi Kitamura, and Bernard Salanié. 2014. "Partial Identification of Finite

Mixtures in Econometric Models." *Quantitative Economics* 5: 123-44.

Moro, Andrea. 2003. "The Effect of Statistical Discrimination on Black-White Wage Inequality: Estimating a Model with Multiple Equilibria" *International Economic Review* 44: 457-500.

Sweeting, Andrew. 2009. "The Strategic Timing of Radio Commercials: An Empirical Analysis Using Multiple Equilibria." *RAND Journal of Economic* 40: 710-42.