The Punisher's Dilemma*

Bruce Bueno de Mesquita, David P. Myatt,
Alastair Smith, and Scott A. Tyson

April 2023

Abstract

We develop a theory of resistance and repression where citizens face a coordination problem and participation in collective resistance must be large enough for a government to concede rather than repress citizens who participate. Repression is costly for both citizens (the punished) and the government (the punisher). Harsher punishments can sometimes raise participation via a strategic-feedback channel that follows from the credibility of implementing repression: a higher cost to the punisher makes a government less willing to repress larger resistance, and this can encourage citizens' coordinated participation. We identify three factors that determine whether this channel dominates: the shape of the response of the government's costs to the intensity and scale of repression; the presence of selective incentives; and the presence or absence of focal optimism. We examine a government's desired choice of repression technology and the response of the social cost of political instability to harsher punishments.

*We thank Karen Albert, Jessica Sun, Randy Stone, and participants of the Virtual Formal Theory Workshop in spring 2021 and the internal political science seminar at the University of Rochester.

In April 2018, US Attorney General Jeff Sessions announced a "zero tolerance" immigration policy on the southwest border of the United States. This policy change included housing migrants in deplorable conditions and separating children from their parents. It was a repressive policy that was intentionally designed to increase the costs migrants bear when coming to the US through the southern border with the goal of reducing the number of asylum seekers. However, by April 2019, the US Customs and Border Protection Agency reported that "the US Border Patrol has seen a more than 370% increase in the number of family units apprehended" year-on-year.[1] Were the consequences of zero tolerance insufficiently severe? Or is the standard deterrence logic implicit in conventional theories of repression incomplete? History is replete with examples where citizens participate in resistance knowing that doing so can mean arrest, serious injury, or even death. But in many cases the threat of harsh punishment seems to have stimulated rather than stifled resistance movements, which is inconsistent with the standard deterrence logic. Is this because repression changes preferences in the short run, via some kind of "backlash," or is there a more strategic explanation?

We present a two-stage game between a large mass of citizens who face a coordination problem and a government, whose threats of repression may lack credibility. In our model citizens first decide whether to mobilize, and then the government observes the size of mobilized resistance and decides whether to repress it. Mobilized citizens are successful if and only if the government does not repress. Key to our contribution is that an exogenous parameter describing the intensity of repression determines both (i) the cost a citizen pays when repressed and (ii) the cost the government pays for repressing, which reflects the need to compensate members of the repressive apparatus for implementing repression (Tyson 2018). Harsher punishments, i.e., more expansive and severe repressive measures, are more costly to endure but are also more costly to administer since the cost the government incurs for repressing is increasing in the size of resistance.

[1] See `https://www.justice.gov/opa/pr/attorney-general-announces-zero-tolerance-policy-criminal-illegal-entry` and `https://www.cbp.gov/newsroom/national-media-release/cbp-releases-march-statistics-southwest-border-migration`. On border apprehensions see `https://www.cbp.gov/newsroom/stats/sw-border-migration`.

We demonstrate how increasing repression's (anticipated) intensity has cross-cutting effects on civilian participation in resistance, and hence, the likelihood that resistance succeeds. First, the threat of more intense repression deters civilians from mobilizing, depressing participation. Second, and novel to our model, more intense repression deters the government from repressing a group of a fixed size—because it is too costly. Although a government may threaten severe repression, if they cannot commit to implementing such repression, citizens anticipate that if enough of them participate, the threat of being repressed lacks *credibility*. Consequently, citizens become more likely to participate. This commitment problem faced by a repressive regime is the *punisher's dilemma*.

The punisher's dilemma has been starkly illustrated in journalistic accounts that document debates within former Soviet regimes during the 1989 revolutions (Garton Ash 1990; Nalepa 2010), or during the Tiananmen Square protests of 1989 in China, where the conservative wing led by Premier Li Peng pushed for harsh repression, and the progressive camp led by Zhao Ziyang argued to concede by making the party more transparent (Liang, Nathan and Link 2001). These cases illustrate the tension among government officials over whether to implement harsh repression or grant concessions. Indeed, in East Germany, over 70,000 protesters showed up on October 9, 1989, despite rumors that the Communists planned to massacre protesters. Ultimately, no massacre took place—repressive agents refused to shoot protesters—and within weeks all members of government had resigned. These examples illustrate policymakers grappling with the underlying strategic logic identified by the punisher's dilemma. In some cases the threat of repression materialized, while in other cases it did not—our theory explores the factors that determine this difference.

In our theory repression generates two kinds of exogenous punishment costs: (i) those that are exclusive to citizens participating in collective resistance; and (ii) costs that are *shared* by participating citizens and the government who implements repression. We refer to the punishment cost borne by citizens as repression's *intensity*, which can include being jailed, beaten, or subjected to human and civil rights violations. The government's cost

(as punisher) depends on repression's intensity as well as its *scale*; that is, the number of citizens who must be repressed. This cost reflects private repressive measures, like bullets/jail cells, and can include compensating the repressive apparatus, via material rewards (to agents) for compliance (Egorov and Sonin 2011; Tyson 2018).

We first derive the conventional direct effect of increased intensity (of punishment) and show that it dissuades citizen participation, resulting from punishment costs borne solely by citizens (e.g., Lichbach 1987; Moore 1998; Gates 2002; Pierskalla 2010). We then shift to our primary focus— the indirect strategic channel—where increased punishment intensity encourages participation. Isolating what gives rise to this strategic channel is important for empirical studies of repression because depending on how the indirect effect varies in magnitude across cases, ignoring it can lead to wrong counterfactuals and misleading empirical findings (Hill and Jones 2014).

We identify three ways that an increase in the intensity of repressive punishments can encourage citizens' participation in political resistance, and in so doing, we help to isolate ways to control for such indirect strategic channels, or suggest research designs to avoid it, in future empirical studies. The first channel we isolate is the *technology of repression*, which determines how the punisher's cost responds to the intensity and scale of punishment, where these responses are conveniently measured by the elasticities of the government's repression cost function.[2] For example, if the *intensity elasticity* exceeds one, then doubling the punishment suffered by a repressed citizen (the intensity) more than doubles the cost borne by the government. Similarly, if the *scale elasticity* is below one, then doubling repression's scale increases the government's cost by less than double. If the intensity elasticity exceeds the scale elasticity, then the cost the punisher bears from repressing responds more strongly to repression intensity than to its scale. As a result, increasing repression's intensity decreases the government's willingness to repress fast enough that citizens become more aggressive in equilibrium. We find that it is in exactly this kind of situation (given some conditions) that an increase in punishment

[2]An elasticity measures the percentage change in the cost borne by the government in response to a percentage change in one of its inputs.

intensity results in greater participation.

Secondly, increased repression intensity leads to increased participation when there is a *selective incentive* for citizens to participate. Selective incentives are private benefits associated with participation in collective resistance that are conditional on its success. Such incentives are often provided by resistance leaders and political entrepreneurs and can include access to services or protection (Popkin 1988).[3] Increasing selective benefits is equivalent (for citizen behavior) to increasing the overall cost of punishment from participation in failed resistance (through the loss of the selective benefit). Consequently, to offset a doubling of participation in resistance, when selective benefits are present, the intensity of punishment must more than double. This has the same effect as increasing government's repression cost by more than double, leading an increase in repression intensity to result in greater participation.

Thirdly, we identify the role of *focal point optimism*. This is when common information (the prior or a public signal) suggests that participation will be sufficiently beneficial that it serves to coordinate citizens' expectations (Schelling 1960; Hardin 1995; Myerson 2004).[4] We increase the "focalness" of common information without altering overall citizen informedness and find that increasing the intensity of repressive punishment causes the credible threat of repression to decrease faster than citizen participation. In equilibrium, this raises participation when citizens are already optimistic about the government granting concessions (and not repressing).

Taken together, our results outline precisely when the standard deterrence logic holds, namely, when either (i) the intensity of punishment is experienced solely by citizens; (ii) if the scale of the government's repressive costs is large relative to its intensity; (iii) selective incentives are large; or (iv) there is a sufficiently pessimistic "focal point" commonly suggesting that the benefit of collective success is small. Our results arise when punishment costs between citizens and the government are shared (i.e., not entirely private), and hence the public good/private cost characterization of political resistance (Tullock

[3]See Wood (2010) for detailed examples.

[4]Examples of focal points include salient places/dates that have historical significance (Truex 2019).

1971) can be misleading. Indeed, this characterization is largely based on treatments where the government is a "passive" party. We depart from that perspective, and highlight an important, and empirically relevant, strategic interplay between citizens looking for political change and their government whose repressive threats may not be credible.

Finally, we use the link between citizens' participation and collective success in resistance movements to explore two additional sets of results. First, we develop an extension where the government chooses repression intensity ex ante, i.e., before it sees the size of collective resistance, and show that it will choose either the maximum or minimum level of repression intensity available. Second, when increasing the intensity of punishment leads to higher collective success, does reducing repression's intensity actually benefit citizens? Although decreasing the shared cost of repression implies that citizens endure less severe repression when they participate, such decreases in the shared cost of repression mean the government concedes less, leading to a higher incidence of repression. There is a third channel arising from our novel results, where decreasing the shared repression cost can increase or decrease participation, depending on the mechanisms we highlighted above. Citizen welfare (in expectation) thus depends on how these different channels manifest from case to case.

Related Literature

The factors determining when a government represses its citizens—rather than granting concessions—is at the heart of the "repression-dissent nexus" and is often studied in a principal-agent setting (Davenport 2007; Svolik 2013; Gehlbach, Sonin and Svolik 2016; Ritter and Conrad 2016). Pierskalla (2010) studies the direct deterrence effect in a game-theoretic setting where third parties introduce perverse signaling incentives. Danneman and Ritter (2014) show that increased repression responds to rebellion in neighboring countries. Shadmehr (2014) shows a U-shaped relationship between income inequality and protest, where repression is the mediating factor. Di Lonardo, Sun and Tyson (2020)

show that variation in repression can result from variation in foreign threats.

A prominent view regarding the association between government repression and dissent refers to a "backlash effect," where increased government repression leads to increased dissent (Lichbach 1987; Mason and Krane 1989; Gibilisco 2020). It is important in backlash arguments that increased dissent occurs *in response* to increased repression, and it is often attributed to being the result of changing preferences. Empirical studies suggest that there may not be a backlash effect (Tilly 1978; Lyall 2009), or instead that there is an "inverted-U" relationship between repression and dissent (DeNardo 2014; Zhukov 2022). We study an entirely different phenomenon that is linked to the credibility of repression: increased dissent by dissidents who anticipate that repression will not materialize—a mechanism commonly associated with deterrence (Di Lonardo and Tyson 2022).

Political resistance is often thought of as a collective action problem (Olson 1965; Lichbach 1995; Tullock 1971). In such contexts, opposition can depend on an informational cascade, where observing political resistance leads to increased opposition through learning about the government (Lohmann 1994). An eruption of political opposition could be the result of hidden preferences that quickly unravel to reveal large-scale opposition (Kuran 1989). Empirical studies have considered the role of different technologies that improve coordinating capability, like cell phones and social media (Zeitzoff 2017). Specifically, Pierskalla and Hollenbach (2013) show that wider cell phone coverage led to more political violence in Africa, and Shapiro and Weidmann (2015) present similar findings, showing that wider cell phone use reduces insurgent violence. Shadmehr and Bernhardt (2011) focus on dilemmas that arise from the combination of coordination with informational concerns, deriving a positive relationship between punishment intensity and its incidence resulting from a tradeoff between the frequency of punishment and successful coordination.

There is a large literature that uses a global game approach to capture coordination problems in political settings, including studies of strategic voting (Myatt 2007, 2017), party leadership (Dewan and Myatt 2007), authoritarian power-sharing (Boix and Svolik

2013), propaganda and censorship (Edmond 2013), and international conflict (Chassang and Padró i Miquel 2010). The closest model to ours is Tyson and Smith (2018) who study the relationship between seemingly independent coordination problems, one for citizens and one for government members, which become strategically linked because of political conflict.[5] Egorov and Sonin (2021) develop a model that focuses on a dictator's incentives to hold elections. They show that increased "cruelty" of repressive policies reduces protest size directly as well as indirectly by increased pessimism in protest success. Our results suggest that this intuition's scope is limited when the cruelty of repressive policies can undermine the credibility of implementing repression.

A Model of Collective Action and Repression

Our model has two-stages: (1) a unit mass of citizens (she), indexed by $i \in [0, 1]$, independently and simultaneously decide whether to participate in collective resistance; (2) after seeing the proportion of citizens who participate, ρ, the government (it) either concedes (granting concessions) or represses (maintaining a status quo).

Citizens' Payoffs. A citizen's direct benefit (or cost if negative) from participation (relative to abstention) is θ. If the government concedes then every citizen enjoys a public-good benefit $w \geq 0$. Additionally, those who participated in the (successful) resistance enjoy a selective benefit $b \geq 0$. A participating citizen suffers a cost $c_E + c_S$ if she is repressed. The *exclusive cost* c_E is only experienced by citizens, whereas the *shared repression cost* c_S also influences the cost of implementing repression for the government. Bringing these elements together, citizen payoffs are summarized by:

	Repress	Concede
Participate	$\theta - c_E - c_S$	$\theta + b + w$
Abstain	0	w

[5] See also Casper and Tyson (2014).

The Government's Payoff. The government enjoys a benefit $v > 0$ from maintaining the status quo and a payoff normalized to 0 from granting concessions. The cost of repressing a resistance of size ρ is $k(c_S, \rho)$. This is a smoothly and strictly increasing function of the *intensity* (of repression), captured by the shared repression cost c_S, and the *scale* (of repression), measured by ρ. If the government represses then its payoff (relative to concession) is $v - k(c_S, \rho)$. We restrict to $k(c_S, 0) < v < k(c_S, 1)$, to avoid trivial cases where the government always/never represses.

Information. Our specification generates multiple equilibria under complete information. Specifically, if $c_E + c_S > \theta > -b$ then there is simultaneously an equilibrium where all citizens abstain and the government represses any citizen who participates, as well as an equilibrium where all citizens participate and the government concedes.

Noting this equilibrium-selection problem, we take a "global game" approach to coordination frictions (Morris and Shin 2003) by specifying uncertainty over θ and equipping citizens with private signals of it. Citizens share a diffuse (flat or improper) prior over the participation payoff θ.[6] Citizen i sees a signal $x_i = \theta + \varepsilon_i$ where the idiosyncratic noise terms (ε_i) are independently drawn from a continuous distribution $F(\cdot)$ with mean zero and full support on $\mathbb{R}$. The posterior expectation of citizen i is $\mathrm{E}[\theta \mid x_i] = x_i$, meaning a citizen's signal is her expected benefit from participation in collective resistance.

Equilibrium Characterization. The timing is: (1) nature determines citizens' private signals; (2) citizens independently decide whether to participate; and then (3) the government observes the resistance and decides whether to repress it. We find a (unique) perfect Bayesian equilibrium with thresholds $\rho^\star$ and $x^\star$ in which a citizen participates if and only if $x_i > x^\star$, and then the government represses if and only if $\rho \leq \rho^\star$.[7]

Comments on the model. Firstly, c_S indexes a repression technology with a joint impact on both the punisher (government) and the punished (citizens). Equating this to the cost experienced by a repressed citizen is a normalization. Specifically, one could have a repression-technology parameter γ_S which generates costs $C(\gamma_S)$ to a citizen and $K(\rho, \gamma_S)$

[6] We study focal point optimism by specifying a proper prior over θ in a subsequent section.

[7] Without loss of generality, an indifferent citizen abstains and an indifferent government represses.

to the government. We normalize by setting $c_S = C(\gamma_S)$ and $k(\rho, c_S) = K(\rho, C^{-1}(c_S))$. The (repressive) government's cost $k(\rho, c_S)$ might arise from tangible factors such as jailing protesters or dismantling rebel groups, but also from the lost labor of jailed citizens or the psychological costs suffered by those who carry out repressive punishments. This last example illustrates that the repressive agent tasked with implementing repression, i.e., the player who bears $k(c_S, \rho)$, could be a different player from the leadership of the regime. In this case, an increase in c_S represents an increase in the difficulty (material or psychological) of repressing citizens. For example, a move from plastic bullets (baton rounds) to fully lethal ammunition, corresponds to an increase in c_S because it raises citizens' participation cost. Additionally, it may also imply that the government's leaders need to compensate members of the repressive apparatus more for implementing harsher repression—a point made by Tyson (2018). This latter feature highlights how factors outside of our model, such as the relationship between the regime and repressive apparatus, partially determine the severity of repression, c_S.

Secondly, after observing the citizens' behavior the government makes a binary "repress or not" choice. Of course, the government might vary the intensity of its repression at this stage. However, it would choose the least costly way of stopping the collective resistance, and this is what we mean by the cost $k(c_S, \rho)$. The government might also choose the number of citizens to repress. In an extension (Supplement Appendix D) we examine a variation of our model where the government can vary the number of repressed citizens which in turn influences the probability that their collective action is successfully subdued. We find conditions under which the government either represses everyone or nobody, leading back to the (simpler) binary choice of our core model.

Thirdly, the technology of repression is (for our core model) an exogenous component, and we (as investigators) change c_S as a comparative-static exercise. The shared cost, c_S, might, however, be an endogenous choice. In a later extension we add a prior stage in which the government chooses its repression technology (through calibrating c_S). By choosing a higher value of c_S a government makes it more costly for those who resist and

are repressed, but also makes it more costly (and less credible) for its agents to carry out that repression, thus potentially undermining their own credibility in the future.

Fourthly, the public benefit from collective success, w, does not influence behavior because each individual citizen perceives her impact on the aggregate outcome as negligible (they are members of a continuum); we retain this parameter solely for completeness. In our model, participation generates individual and selective benefits via the parameters θ and b—these can be thought of as club-good components.[8] A citizen's net benefit from participation, θ, can be positive if participation is enjoyable, or negative, if a citizen incurs an opportunity cost from participation. That citizens lack common knowledge about θ generates a coordination problem in our model (we have specified uncertainty exclusively over a common θ for simplicity). An equivalent formulation (developed in Supplemental Appendix B) is one where citizens have different participation costs but do not know the average participation cost among citizens.

Finally, we focus on factors that reverse the standard deterrence logic from conventional accounts of repression, and so some real-world features are intentionally left lean or omitted. Such omissions are essentially equivalent to holding factors fixed in an empirical study (Paine and Tyson 2020), which allows us to interpret our comparative-static results as counterfactual comparisons (Ashworth, Berry and Bueno de Mesquita 2021). For instance, one could imagine that c_S increases in ρ because the government resorts to more intense repressive measures as ρ increases, and that c_E decreases in ρ because the probability of being targeted decreases as ρ increases. These competing channels, although interesting, are not about the relationship we study, but the empirical relationship between the size of resistance and the cost from repression an individual experiences, and we leave their examination for future work.

[8] If there were a finite number of citizens then they would take into account the chance of being pivotal (similar to a voting model) and w would become relevant (see, e.g., Shadmehr 2019).

How Repression Shapes Collective Action

We now analyze citizen participation in collective resistance, when citizens anticipate the government's decision whether to maintain a status quo through repression.

Proposition 1 (Equilibrium) *There is a unique equilibrium, characterized by the pair* $(\rho^\star, x^\star)$ *that simultaneously satisfy (i)* $v = k(c_S, \rho^\star)$ *and (ii)* $x^\star = \rho^\star(c_E + c_S) - (1 - \rho^\star)b$.

The equality (i) corresponds to the government's behavior: it represses if $v \geq k(c_S, \rho)$, which holds if the resistance falls below the threshold $\rho^\star$ satisfying (i). By standard arguments, $\rho^\star$, which satisfies $0 < \rho^\star < 1$, is increasing in the government's desire to retain power, v, and decreasing in the shared repression cost, c_S.

Fixing $\rho^\star$, the citizens play a binary-action coordination game. As we have noted, with complete information there are two fully-coordinated equilibria (all citizens participate, or all abstain). With strategic uncertainty (see Supplemental appendix A) a citizen participates if and only if her signal x_i (of the parameter θ) exceeds a threshold $x^\star$, which satisfies (ii) in Proposition 1.

To find $x^\star$, notice that a citizen (optimally) participates if and only if

$$\underbrace{\mathrm{E}[\theta \,|\, x_i] + b \Pr[\text{Concession} \,|\, x_i]}_{\text{expected benefits}} \geq \overbrace{(c_E + c_S)[1 - \Pr[\text{Concession} \,|\, x_i]]}^{\text{expected costs}}. \tag{1}$$

The posterior expectation $\mathrm{E}[\theta \,|\, x_i] = x_i$, and the probability concessions are granted, $\Pr[\text{Concession} \,|\, x_i]$, are both increasing in x_i, and so citizen i participates if her signal exceeds the threshold $x^\star$. To pin down the equilibrium value of this threshold we observe that (1) holds as an equality when $x_i = x^\star$, and re-arranging, this condition is:

$$x^\star + (b + c_E + c_S) \Pr[\text{Concession} \,|\, x_i = x^\star] = c_E + c_S. \tag{2}$$

The government concedes if the size of resistance, ρ, exceeds the threshold $\rho^\star$ from (i), and so $\Pr[\text{Concession} \,|\, x_i = x^\star] = Pr[\rho > \rho^\star | \, x_i = x^\star]$. For a citizen with a threshold-

equalling signal ($x_i = x^\star$) the question "what is the probability $\rho^\star$ (or more) other citizens will participate?" is equivalent to "what is the probability that a fraction $\rho^\star$ (or more) of other citizens have signals greater than mine?" This event holds if and only if her own signal is in the bottom fraction $1 - \rho^\star$ of signals. She therefore asks: "what is the probability I am in the $1 - \rho^\star$ most pessimistic signals?" The answer is $1 - \rho^\star$, hence $\Pr[\rho > \rho^\star | \, x_i = x^\star] = 1 - \rho^\star$. It follows that condition (2) is

$$x^\star + (b + c_E + c_S)(1 - \rho^\star) = c_E + c_S,$$

which pins down the threshold $x^\star$, giving the expression reported in Proposition 1.[9]

The Standard Logic. To establish a benchmark, we first consider the influence of a change in the exclusive punishment cost, c_E, which only impacts citizens. Note that participation is higher whenever the threshold $x^\star$ is lower.

Remark 1 (Standard Deterrence Logic) *Citizens' participation in collective resistance is decreasing in the exclusive cost of participation and increasing in the selective benefits from participation: the threshold $x^\star$ is increasing in c_E and decreasing in b.*

Remark 1 follows by direct differentiation of the second equilibrium condition (ii):

$$\frac{dx^\star}{dc_E} = \rho^* \quad \text{and} \quad \frac{dx^\star}{db} = -(1 - \rho^*). \tag{3}$$

An increase in c_E harms a participating citizen when collective resistance is repressed. For an indifferent citizen ($x_i = x^\star$) the perceived chance of this is $\rho^\star$, which is the probability that a fraction $1 - \rho^\star$ or more have lower signals. To offset this, the threshold $x^\star$ adjusts upward, thus reducing the level of participation by citizens. From the second part of (3), a change in the selective benefit b yields a qualitatively similar effect. This direct effect arises since the exclusive cost, c_E, and the selective benefit, b, do not influence the government, and so do not change $\rho^\star$.

[9] Uniqueness can be straightforwardly established and is relegated to the online supplement.

Shared Repression Costs and Participation. The shared repression cost, c_S, feeds directly, via the standard deterrence channel, into the threshold $x^\star$ in the same way as c_E, but it also influences participation indirectly by altering the government's cutoff rule $\rho^\star$, a channel that is novel to our framework. Explicitly:

$$\frac{dx^\star}{dc_S} = \rho^\star + (c_E + c_S + b)\frac{d\rho^\star}{dc_S}. \tag{4}$$

The first term is the direct effect of the shared cost of repression, which by comparison with (3), shows the deterrence channel from increasing citizens' costs. The second term in (4) represents the indirect channel via the change in the government's threshold $\rho^\star$. This captures the government's commitment problem: although the government would like to commit to repress any resistance, i.e., $\rho = 1$, doing so is not sequentially rational.

To work out how $\rho^\star$ changes, we can (implicitly) differentiate equality (i) from Proposition 1:

$$\frac{d\rho^\star}{dc_S} = -\frac{\partial k(c_S, \rho^\star)/\partial c_S}{\partial k(c_S, \rho^\star)/\partial \rho}, \tag{5}$$

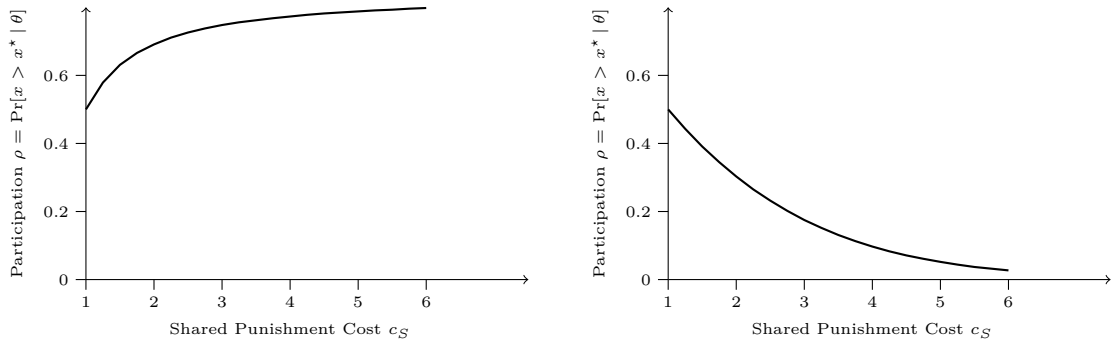
which generates one of our main results.

Proposition 2 (Total Effect of Shared Repression Costs) *A higher shared repression cost c_S strictly increases citizen participation in collective resistance if and only if*

$$\underbrace{\frac{\partial \log k(c_S, \rho^\star)}{\partial \log \rho}}_{\textit{scale elasticity}} < \underbrace{\left(1 + \frac{c_E + b}{c_S}\right)}_{\textit{impact of selective incentives}} \times \underbrace{\frac{\partial \log k(c_S, \rho^\star)}{\partial \log c_S}}_{\textit{intensity elasticity}}. \tag{6}$$

An increase in c_S strictly decreases citizen participation if and only if the opposite inequality holds, and it has no effect on citizen participation if (6) *holds as an equality.*

The *scale elasticity* on the left-hand side of (6) measures how the size of resistance affects a government who represses. If it exceeds one then there are decreasing returns to scale, meaning that an $m\%$ increase in the size of resistance imposes a more than $m\%$ increase in the government's costs of repression. The scale elasticity captures how the cost



(a) More responsive to intensity ($\beta = 2.0$) (b) Less responsive to intensity ($\beta = 0.4$)

We illustrate the effect of c_S on citizen participation ρ. We use $k(c_S, \rho) = c_S^{\beta} \cdot \rho^{\gamma}$, $v = 1$, $\gamma = 1$, $c_E = b = 0$, $\theta = 1$. Signals are normal with unit variance.

Figure 1: The Effect of Shared Punishment Costs on Participation

of implementing repression can increase, and even become excessively high, as the size of collective resistance, ρ, changes. On the right-hand side of (6) is the *intensity elasticity*, which measures how the government's costs respond to changes in repression intensity, measured by c_S. If the intensity elasticity is more than one, then doubling the intensity of punishment that citizens endure (doubling c_S) more than doubles the government's cost, $k(c_S, \rho)$. Finally, the middle term in (6) captures the relative importance of selective incentives: the exclusive cost c_E, and the selective benefit b from participation in a successful resistance. If these are absent then this term simplifies to one.

A pair of corollaries serve to highlight the impact of each factor in isolation.

Corollary 1 (Relative Response to Scale and Intensity) *If there are no selective incentives, $b = c_E = 0$, then a higher shared repression cost c_S increases participation if and only if the government's cost is more responsive to intensity than to scale.*

Figure 1 illustrates how the elasticities of $k(c_S, \rho)$ affect the relationship between collective resistance and the cost of repression. To build intuition, begin with the case where there are constant returns to both scale and intensity (so that both elasticities are equal to one). A doubling of c_S doubles the government's per-citizen cost of repression,

and so halves the resistance that the government is willing to repress. The threshold citizen $x_i = x^\star$ sees a punishment cost that is twice as large, but also perceives half the risk of failure; these considerations offset exactly. Now suppose that the government's cost responds more to intensity, i.e., the intensity elasticity exceeds 1. Doubling c_S now lowers the resistance the government is willing to repress by more than half, which less than halves the expected cost incurred from participation in collective resistance. This extra movement down in $\rho^\star$ increases participation from citizens.

To isolate the role of selective incentives, we balance the effects of intensity and scale by equating their elasticities (for example, by setting $k(c_S, \rho) \propto c_S \cdot \rho$).

Corollary 2 (Selective Incentives) *Suppose that the government's cost of repression is equally responsive to scale and intensity, so that the scale elasticity of $k(c_S, \rho)$ is equal to the intensity elasticity. An increase in the shared punishment cost c_S increases citizens' participation if and only if there are selective incentives: $b + c_E > 0$.*

To explain this result, suppose that there are no selective benefits ($b = 0$) but that there are exclusive costs from repression, $c_E > 0$. The total cost of punishment felt by a citizen is $c_E + c_S$. Doubling this requires more than doubling the shared-cost parameter c_S, which more than doubles the government's cost. This means that within the relationship between costs experienced by citizens and the government there are, in effect, decreasing returns to the intensity of repression. The selective-benefit parameter b plays an identical role (for decision-making) to c_E, and so the same kind of argument applies when $b > 0$.

The Selma Marches. An insight of our model is illustrated by the March (or really Marches) from Selma to Montgomery, Alabama in 1965 as part of the Civil Rights Movement. Demonstrators in Selma sought the right to vote, but faced resistance in the Jim Crow South. On January 18, 1965, about 300 people marched to the courthouse in Selma intending to register to vote, but were prevented from doing so by local police. Their effort was repeated by 110 black teachers on January 22. The number of protesters continued to swell as violence against them increased. On February 5, Dr. Martin Luther

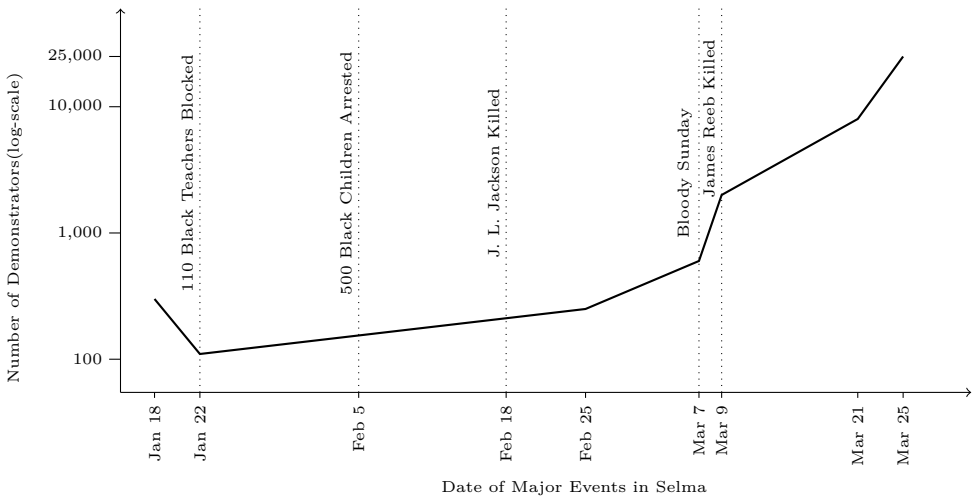


Figure 2: Participation in the Selma March

King Jr. led 500 black school children on a march in Selma, all of whom were arrested. On March 7, 600+ demonstrators walked across the Edmund Pettus Bridge where they were confronted by police who tear gassed and billy clubbed them, and two days later, James Reeb was murdered during a protest march of about 2,000 people. By March 21, an estimated 3,200–8,000 people participated in a march from Selma to Montgomery to assert their voting rights, and by late March, 25,000 marchers had arrived in Montgomery. Figure 2 shows some key events, with the number of demonstrators.

What motivated protesters to to show up *after* they witnessed Alabama's increasingly severe response? The federal government's immediate response to the events of Bloody Sunday (March 7) was to consider prosecuting local and state officials involved in the event and to accelerate passage of the Civil Rights Act of 1965 (March 15). In addition, President Johnson nationalized the Alabama National Guard to protect demonstrators, forcing Alabama to back down and participate in protecting the rights of the very people they had previously repressed. These events show the relationship our main results highlight at work: as the costs from resisting the Alabama state government increased, the ranks of those putting themselves in harms way grew rather than shrank. As the numbers became large, the sustained and growing resistance became too costly for the federal government to endure, ultimately forcing the state government of Alabama to

concede.

The Impact of the Technology of Repression

Proposition 2 shows how a weakened government response (a reduction in $\rho^\star$) can offset (if (6) holds) the direct deterrent effect of increased repression, leading to increased participation in collective resistance. We now evaluate the success of resistance and extend our model to characterize a government's choice of repressive technology.

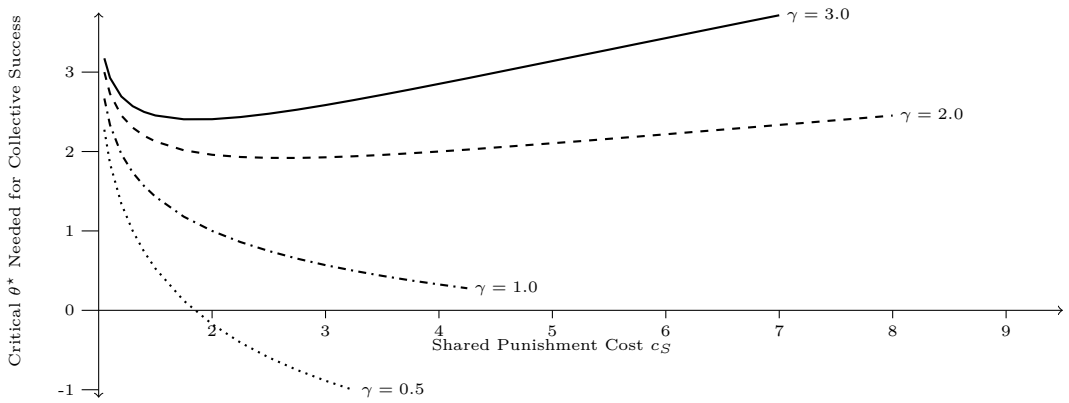
The Success of Collective Resistance. Recall that citizen i's signal is $x_i = \theta + \varepsilon_i$ where ε_i is drawn from a distribution $F(\cdot)$, and so the proportion of citizens who participate is $\rho = 1 - F(x^\star - \theta)$. It needs to exceed the (endogenous) threshold $\rho^\star$ for the resistance to succeed.[10] This happens if and only if θ exceeds the critical value

$$\theta^\star = x^\star - F^{-1}(1 - \rho^\star).$$

This depends on c_S (the shared cost of repression) via two channels. First, there is the government's behavior: $\theta^\star$ is increasing in $\rho^\star$ (the second term of $\theta^\star$) and so decreasing in c_S. The government is less willing to repress when costs are high, which expands the circumstances when it concedes. Second there is citizens' behavior: if $x^\star$ is decreasing in c_S then the first term also favors the success of collective resistance following an increase in c_S. If not, then there are competing effects from the channels and the overall effect of an increase in c_S will vary from case to case.

Evaluating the success of collective resistance simplifies when the variation in citizens' signals is small, so that $F(\cdot)$ is tightly clustered around zero, thereby reducing coordination frictions. This means that citizens have near-perfect information about θ, implying that $F^{-1}(1 - \rho^\star) \approx 0$. Then, because $\theta^\star \approx x^\star$, if $\theta > \theta^\star$, all citizens participate and the government concedes; whereas if $\theta < \theta^\star$, no citizens participate and the government incurs no cost; there is perfect coordination and citizens are never repressed.

[10] We are able to "skip" this part of the analysis in Proposition 1, by exploiting the flat prior over θ.



Notes. The figure illustrates the effect of c_S on the critical state $\theta^\star$. We use $k(c_S, \rho) = c_S \cdot \rho^\gamma$ (so that $\beta = 1$), $v = 1$, and $c_E = b = 0$. The cases correspond to different values for the γ parameter, and so vary the extent of scale elasticity in the cost of punishment. Signals are: $x_i \,|\, \theta \sim N(\theta, \xi^2)$ with $\xi^2 = 1$.

Figure 3: The Threshold Needed for Collective Success

Proposition 3 (Collective Success) *If an increase in repression intensity (an increase in c_S) raises participation, so that* (6) *from Proposition 2 holds, then it also decreases $\theta^\star$ and so it expands the range of circumstances where collective resistance succeeds.*

If variation in citizens' signals is sufficiently small then Proposition 3 becomes "if and only if" in the sense that heightened repression intensity expands the circumstances in which collective resistance succeeds if and only if (6) from Proposition 2 holds.

Proposition 3 is illustrated in Figure 3. The horizontal axis is the shared cost of repression, and the vertical axis is the critical state determining when concessions are achieved. We have eliminated any selective incentives (by setting $c_E = b = 0$) and we have chosen a specification for the government's cost function which exhibits constant returns in the repression intensity (so that it is linear in c_S). Over the four cases illustrated, we vary the returns to scale via a scale elasticity parameter.

The lowest (dotted) line in Figure 3 (labeled as $\gamma = 0.5$) specifies a scale elasticity less than one. This is a case where (6) from Proposition 2 holds, and so an increase in the shared cost of repression helps the success of the collective resistance. Similarly,

the second-lowest (dashed and dotted) line (labeled as $\gamma = 1.0$) corresponds to a scale elasticity of exactly one. For this case, the threshold $x^\star$ is invariant to c_S. However, the direct effect of $\rho^\star$ in the solution for $\theta^\star$ means that an increase in repression costs continues to increase resistance (given that citizens' signals are not perfectly accurate).

The remaining two cases (the solid and dashed lines) in Figure 3 are cases where the scale elasticity exceeds one and (6) no longer holds: an increase in c_S results in an increase in $x^\star$, thus reducing citizen participation—increased repression can work against the success of collective resistance. In these two cases, the relationship between the cost of repression and the collective success of the citizens is inverse U-shaped: success is highest for intermediate values of the shared-cost parameter, c_S. Equivalently, the government maintains the status quo when repression intensity is very low or very high.

Repression Costs with Constant Elasticity. We now examine a special case where the intensity and scale elasticities are constant, with $k(c_S, \rho) = c_S^\beta \cdot \rho^\gamma$.[11] The constants β and γ are the intensity elasticity and scale elasticity respectively. Under this specification, inequality (6) is

$$\frac{\gamma}{\beta} < 1 + \frac{b + c_E}{c_S}. \tag{7}$$

The left-hand side is the relative importance (to the government) of scale versus intensity. The right-hand side is the relative importance of selective incentives (b and c_E) compared to the shared element of the cost of repression (c_S).

If $\gamma < \beta$ then since $b + c_E \geq 0$, (7) is always satisfied, which also means that the conclusion of Proposition 3 applies. The more interesting case is when $\gamma > \beta$ (the scale elasticity exceeds the intensity elasticity) and selective incentives are strictly positive ($b + c_E > 0$). Under these conditions, inequality (7) holds if c_S is small but fails if c_S is large. This implies that there is a unique value of c_S for which (7) holds as an equality.

Proposition 4 (Inverse U-Shaped Response of Collective Success) *Suppose that* $k(c_S, \rho) = c_S^\beta \cdot \rho^\gamma$. *If* $\gamma < \beta$ *then participation in collective resistance, and the collective*

[11] Everything that we say holds if $k(c_S, \rho)$ is any smoothly increasing function of $c_S^\beta \cdot \rho^\gamma$.

success of that resistance, are both increasing in the shared repression cost c_S. If $\gamma > \beta$ (the scale elasticity is larger), and if selective incentives are strictly positive, then $x^\star$ is "U-shaped" in the shared repression cost c_S, and participation in collective resistance is "inverse U-shaped" in c_S. Moreover, there exists a unique value

$$c_S^\dagger = \frac{\beta \cdot (b + c_E)}{\gamma - \beta} \tag{8}$$

of the shared repression cost that maximizes citizen participation. If the variation in signals vanishes then it also maximizes the circumstances in which the resistance succeeds.

To illustrate this result consider when c_S is large, then the shared cost of repression is much more important than the exclusive cost, c_E, borne solely by citizens. This means that for large c_S, the properties of the government's cost function $k(c_S, \rho)$ are of most importance. The condition $\gamma > \beta$ corresponds to when the cost of implementing repression is more responsive to scale than intensity. This means that (other things equal) the government is relatively helped by an increase in c_S because it shifts emphasis to smaller levels of participation when the decreasing-returns-to-scale property of the government's cost function bites less. For that reason, $\theta^\star$ is increasing in c_S, hence the government is helped by an increase in the shared-cost parameter when c_S is large.

Now consider when c_S is very low. This means that the exclusive cost c_E is relatively high, and so c_S represents a small proportion of the cost faced by a repressed citizen. This means that to double the intensity felt by citizens requires much more than a doubling of c_S which means that (in effect) there are decreasing returns to intensity. Equivalently, the scale elasticity is of relatively lower importance. This means $\theta^\star$ is decreasing in c_S (the government is hindered by an increase in the shared-cost parameter) when c_S is small.

The Government's Choice of Repression Technology. In our model the shared-cost parameter, c_S, represents the technology of repression and captures things such as the lethality of the weapons and tactics used by repressive agents. This technology, of course, might be something that has been strategically chosen by the government, who

may (at least partially) anticipate the effect of c_S on resistance. Extending our main model slightly, we introduce an initial stage where the government selects the intensity of repression $c_S \in [c_{S\min}, c_{S\max}]$, after which the game proceeds as above.

To evaluate the government's choice, we need to specify its payoffs and prior beliefs, and to keep the analysis simple, we suppose that this early-stage government is the same player as the final stage government, so that the government shares the same payoffs as its repressive agent. Specifying a prior belief over the parameter θ, we would obtain

$$\text{Government payoff} = \Pr[\theta < \theta^\star]\left(v - \mathrm{E}\left[k(c_S, 1 - F(x^\star - \theta)) \,|\, \theta < \theta^\star\right]\right)$$

This incorporates the expected benefit, equal to $v \Pr[\theta < \theta^\star]$, from maintaining the desired status quo, as well as the expected cost of repression. This simplifies appreciably when there is little noise in citizens' signals since coordination frictions disappear and all citizens participate if and only if $\theta > \theta^\star$, and the government concedes. This means that the government's payoff reduces to $v \Pr[\theta < \theta^\star]$.

A second approach is to suppose that the early-stage government cares only about maintaining the status quo. In this case, the government does not care about the costs incurred by its agents, but recognizes that an increase in c_S reduces the credibility of a threat to repress. Such a government maximizes $\Pr[\theta < \theta^\star]$, and once again, an early stage government will seek to minimize the range of collective success (maximizes $\theta^\star$). We thus focus our analysis on a government whose objective is simply to maximize θ^* and remain agnostic as to the exact reason why.

Proposition 5 (The Choice of Repression Technology) *Suppose that the intensity and scale elasticities are both constant, and that a government seeks to minimize the range of circumstances that lead to collective success. Further suppose that the noise in citizens' signals is negligible. Such a government optimally chooses either the most lenient ($c_S = c_{S\min}$) or the most punitive ($c_S = c_{S\max}$) repression technology.*

This gives a "bang-bang" result on the severity of repression when it is chosen endoge-

nously by the government, which allows them to avoid, at least partially, the punisher's dilemma. Specifically, the government, when choosing the severity of repression, will choose only the minimum value they can choose or the maximum value that is available to them. One possibility is that the lower-limit for the shared repression cost parameter is $c_{S\,\min} = 0$, in which case the government eliminates all shared costs (the exclusive cost, c_E, remains); we explore this in detail in Appendix D.

Proposition 6 *For the conditions stated in Proposition 5: (i) the government's optimal repression technology is independent of the government's value of the status quo, v; (ii) an increase in selective incentives via an increase in either b or c_E favors the use of the most lenient repression technology; (iii) an increase in either of the upper or lower limits $c_{S\,\min}$ and $c_{S\,\max}$ favors the use of the most punitive repression technology.*

This result identifies a number of important empirical implications. First, the value of the status quo to the government, v, does not influence their choice of repressive technology. Second, selective incentives (c_E or b) push a government toward more lenient repressive technologies. Third, increasing the upper constraint, or increasing the lower constraint, on repression severity favors more punitive repressive technologies.

Focal Points and Prior Optimism or Pessimism

We have thus far specified a flat prior over the (net) participation payoff θ. This has allowed us to focus on technological features of repression and selective incentives, but prevents an investigation of the role of a common prior belief (or a public signal) where prior optimism (or pessimism) about the success of collective resistance can influence behavior. Here we specify a proper prior and to focus fully on the information environment, we remove selective incentives: $c_E = b = 0$.

Information. Citizens commonly believe that θ is drawn from a normal distribution with mean μ and variance ζ^2.[12] This can be interpreted as a common prior, or as the

[12]Traditional models of protest that focus on "participation paradoxes" essentially assume that $\mu =$

updated beliefs of citizens following the observation of a common public signal of θ.[13] Citizen i then sees a private signal $\tilde{x}_i = \theta + \varepsilon_i$ where ε_i is an independent draw from a normal distribution with mean 0 and variance ξ^2.

This specification generates posterior beliefs about θ that are normally distributed:

$$\theta \,|\, \tilde{x}_i \sim N(x_i, \sigma^2) \text{ where } x_i \equiv \mathrm{E}[\theta \,|\, \tilde{x}_i] = \frac{\xi^2\mu + \zeta^2\tilde{x}_i}{\xi^2 + \zeta^2} \text{ and where } \sigma^2 \equiv \mathrm{var}[\theta \,|\, \tilde{x}_i] = \frac{\xi^2\zeta^2}{\xi^2 + \zeta^2}.$$

We use "$\tilde{x}_i$" rather than "x_i" for a citizen's private signal to maintain clear comparability with our earlier results, where x_i stood for a citizen's posterior expectation of θ (with a flat prior, a citizen's posterior expectation is equal to her signal).

To measure the relative importance of common information (or prior beliefs), recall that the precision of a random variable is the reciprocal of its variance, and define

$$\lambda \equiv \frac{\xi^2}{\xi^2 + \zeta^2} = \frac{1/\zeta^2}{(1/\xi^2) + (1/\zeta^2)} = \frac{1/\zeta^2}{1/\sigma^2} = \frac{\text{Prior Precision}}{\text{Posterior Precision}}.$$

Here λ is the precision of the prior as a proportion of the sum of the total precision of information citizens receive, or equivalently, the ratio of prior to posterior precision. This measures the proportion of information that is common among citizens. When λ is large, citizens get a majority of their information from common sources, whereas when λ is small, citizens base decisions on private information, which makes coordination more difficult since there is less common knowledge.

The informational environment is specified above in terms of the prior and signal variances (ζ^2 and ξ^2). However, it is equivalent (and useful) to use σ^2 and λ, allowing us to vary the importance of the prior while fixing the precision of citizens' beliefs. In terms

$\zeta^2 = 0$.

[13] Suppose, for example, that citizens begin with a flat prior over θ, as they do in our main model. If citizens commonly observe a public signal from a normal distribution with variance ζ^2 and realization μ, then they update to a common posterior of θ that is normally distributed with mean μ and variance ζ^2.

of σ^2 and λ, the (conditional) distribution of citizens' posterior expectations is

$$x_i \,|\, \theta \sim N(\lambda\mu + (1-\lambda)\theta, (1-\lambda)\sigma^2).$$

Fixing the posterior precision of their beliefs, citizens' expectations become tighter and move toward the prior (and away from the true θ) as λ increases.[14]

Proposition 7 (Participation and Repression with a Proper Prior) *There exists an equilibrium, characterized by $(\rho^\star, x^\star)$, where the government concedes only if collective resistance exceeds $\rho^\star$ satisfying $v = k(c_S, \rho^\star)$, and a citizen participates in collective resistance if and only if her expected benefit from participation exceeds $x^\star$, where $x^\star$ solves*

$$x^\star = c_S \Phi\left(\frac{\lambda(x^\star - \mu)}{\sigma(1-\lambda)} + \frac{1}{\sqrt{1-\lambda}}\Phi^{-1}(\rho^\star)\right). \tag{9}$$

This equilibrium is unique if $\lambda/(1-\lambda) < (\sigma/c_S)\sqrt{2\pi}$ (so that the prior is not too precise).

The government's decision rule is unaffected: it concedes if $\rho \geq \rho^\star$ and represses otherwise. We note that equation (9) reduces to $x^\star = c_S\rho^\star$ (this is the solution reported in Proposition 1 for $c_E = b = 0$) as $\lambda \to 0$ (so that the prior becomes diffuse).

Fixing a threshold $x^\star$ used by citizens, we characterize how this translates into the success of collective resistance. The conditional distribution of posterior expectations (of θ) is normal with mean $\lambda\mu + (1-\lambda)\theta$ and variance $(1-\lambda)\sigma^2$ and so, for cutoff $x^\star$, the proportion of citizens who participate in the resistance is

$$\rho(\theta \mid x^\star) = \Pr[x_i > x^\star \,|\, \theta] = \Phi\left(\frac{\lambda\mu + (1-\lambda)\theta - x^\star}{\sigma\sqrt{1-\lambda}}\right).$$

Given $\rho^\star$, we can calculate the critical state $\theta^\star$ above which the government chooses to

[14] The primitive parameters can be readily recovered: $\zeta^2 = \sigma^2/\lambda$ and $\xi^2 = \sigma^2/(1-\lambda)$.

concede, using the expression $\Pr[x_i > x^\star \,|\, \theta^\star] = \rho(\theta^\star \mid x^\star) = \rho^\star$. After rearranging,

$$\theta^\star = \frac{1}{1-\lambda}\left(x^\star - \lambda\mu + \sigma\sqrt{1-\lambda}\Phi^{-1}(\rho^\star)\right).$$

We now use the optimality condition of a citizen with the threshold belief $x^\star$. She must be indifferent, and so her expected benefit from participation (this is $x^\star$) equals her expectation of the punishment costs:

$$x^\star = c_S \Pr[\text{Repression} \,|\, x^\star] = c_S \Pr[\theta < \theta^\star \,|\, x^\star].$$

Since the indifferent citizen's beliefs about θ are normal with mean $x^\star$ and variance σ^2,

$$x^\star = c_S \Phi\left(\frac{\theta^\star - x^\star}{\sigma}\right) = c_S \Phi\left(\frac{\lambda(x^\star - \mu)}{\sigma(1-\lambda)} + \frac{1}{\sqrt{1-\lambda}}\Phi^{-1}\left(\rho^\star\right)\right),$$

where for the second step we have substituted in the expression for $\theta^\star$ in terms of $x^\star$, and re-arranged. This is the equilibrium condition reported in Proposition 7.[15]

Focal Points. A (proper) prior means that there are commonly held beliefs about the benefits from participation, which implies that its mean, μ, and importance, via λ, can act as coordinating devices. For this analysis it is important that the importance of the prior is changed without altering the quality of citizens' information, so we fix the parameter σ^2, and change λ, because this holds fixed the precision of posterior beliefs.

Proposition 8 *Fix the precision of posterior beliefs σ^2. Then, (i) increasing μ reduces $x^\star$, increasing citizens' participation in collective resistance; (ii) an increase in the importance of the prior, λ, reduces $x^\star$ and increases citizens' participation if and only if*

$$\mu > x^\star + \frac{\sigma\sqrt{1-\lambda}}{2}\Phi^{-1}\left(\rho^\star\right); \tag{10}$$

[15] A sufficient condition for uniqueness is that the right-hand side of (9) does not increase too quickly with $x^\star$. If the derivative of the right-hand side with respect to $x^\star$ is strictly less than one then the two sides of the equation cross only once. Imposing this generates the stated condition; details are in the online supplement.

(iii) if shared costs are such that $\rho^\star > \frac{1}{2}$, then participation is first decreasing and then increasing ("U-shaped") in λ. If shared costs are larger, so that $\rho^\star < \frac{1}{2}$, then participation is first increasing and then decreasing ("inverse U-shaped") in λ.

This shows how common information prompts a "clustering" of citizens' choices, which is consistent with focal point accounts, and captures distinct ways it manifests. The first part follows by inspection: the right-hand side of condition (9) is decreasing in the prior mean, μ, and so higher μ induces greater citizen participation; thus mobilizing people by coordinating more citizens around participation.

The second part of Proposition 8 focuses on increasing the importance of common information and shows that participation increases in the importance of the prior, λ, provided μ is large enough. This is not the result of increased information suggesting that participation will be beneficial since we are holding fixed the variance of posterior beliefs, σ^2. Rather, it depends on the *source* of that information—from private to public. This is akin to increasing the "focalness" of information, and the result shows that making information more common increases coordination among citizens; causing increased participation when the prior mean satisfies (10) and decreased participation otherwise.

The last part of Proposition 8 elucidates a dependence on whether the shared cost is relatively low (so $\rho^* > \frac{1}{2}$) or high (so $\rho^* < \frac{1}{2}$), corresponding to whether the government is willing to repress a majority or only a minority of citizens. If the government is willing to repress a majority, then increasing the importance of the prior first decreases participation, but after a point, increases participation. If the government is only willing to repress a minority then this pattern is reversed.

Shared Repression Costs and Focal Points. One of our central results is Proposition 2, which evaluates the effect of changes in the shared repression cost, c_S, on citizens' willingness to participate in collective resistance. Shared repression costs enter equilibrium condition (9) via two channels: directly by increasing the cost citizens bear from repression and indirectly from reducing the incidence of government repression.

To identify the role of focal point optimism, we shut down other potential mecha-

nisms (from Proposition 2) and specify an equal-elasticity case: $k(c_S, \rho) = c_S \cdot \rho$. This means that, absent selective incentives and with a flat prior, changing c_S has no effect on citizens' participation. This allows us to isolate how a focal point can introduce a similar consequence as Proposition 2, but through a new mechanism: focal-point optimism.

Proposition 9 (Shared Costs and Focal Points) *Citizen participation is increasing in shared repression costs, c_S, if and only if*

$$x^\star < \frac{v}{\sqrt{1-\lambda}} \exp\left(\frac{1}{2}\left[z^2 - \left(\frac{\lambda(x^\star - \mu)}{\sigma(1-\lambda)} + \frac{z}{\sqrt{1-\lambda}}\right)^2\right]\right) \quad \textit{where} \quad z \equiv \Phi^{-1}\left(\frac{v}{c_S}\right). \quad (11)$$

Participation is increasing in shared costs, c_S, if the prior mean is sufficiently optimistic and decreasing in shared costs if the prior mean is sufficiently pessimistic.

Consistent with focal-point-style arguments (Schelling 1960; Myerson 2004), Proposition 9 shows that the way c_S affects participation depends on how much optimism is shared among citizens. If they are optimistic, i.e., when (11) is satisfied, then an increase in c_S leads to higher participation, and the opposite holds if they are pessimistic. It is important to emphasize that we intentionally removed selective incentives and technological features of repression that would lead to a positive relationship between repression intensity and collective resistance to show that such a relationship can arise from the combination of optimism and coordination.

The balance in responsiveness of the government's repression costs of intensity and scale (our multiplicative functional form) implies that the government's cutoff satisfies $\rho^\star = \frac{v}{c_S}$ and that (6) in Proposition 2 holds with equality. To represent a control case in which the use of a prior has no direct effect on the equilibrium participation of citizens, we ensure that the inclusion of prior beliefs has no direct effect; meaning the prior mean is *neutral*. This is accomplished by setting the equilibrium cutoff to be the same as in the absence of a prior, which in this case is $x^\star = v$, and remains unchanged for different λ.

Proposition 10 *Using the specification $k(c_S, \rho) = c_S \cdot \rho$: (i) participation is decreasing in c_S if the cost of repressing citizens is sufficiently close to the government's valuation*

of repression, v; (ii) if the shared repression cost satisfies $c_S = 2v$, then the government concedes if and only if a majority of citizens participate. In this case, if the prior is neutral (or sufficiently close to it), then participation is increasing in shared costs, c_S.

For $c_S < v$, the government always represses and so a citizen participates if and only if her expected benefit exceeds the cost of punishment: $x^\star = c_S$, which is increasing in c_S. As c_S rises through and above v, (while remaining close to v) this result says that $x^\star$ continues increasing.[16] This essentially extends our earlier result showing that when the strategic feedback highlighted earlier is removed, the standard logic dominates.

The second part concerns an important benchmark case. If $c_S = 2v$ then the government's cutoff is $\rho^\star = \frac{1}{2}$, meaning that the government is willing to repress resistance from exactly half of the citizens, but no more. Citizens' beliefs being neutral means that the prior satisfies $\mu = v$ in this case, and the equilibrium threshold $x^\star = v$ does not vary with λ. Moving away (infinitesimally) from this benchmark solely by varying c_S, the result shows that, in this neutral benchmark case, participation is increasing in the shared repression cost, c_S.

The Rise of Christianity. We apply the focal point mechanism that our results highlight to the political context surrounding the rise of Christianity following their persecution in the mid-third to early-fourth centuries. The Plague of Cyprian was an epidemic that killed up to 5,000 Romans a day, lasting from around 249 until about 270. Christians died at a much lower rate than Roman pagans, in part because early Christian communities nursed their sick, whereas pagans did not. Consequently, there was an uptick in Christian conversions, a fact that was observed at the time (Stark 1996; Hopkins 1998).

Starting in the year 250, the Roman Emperor Decius initiated the first Empire-wide persecution of Christians, thus making "being Christian" an act of political resistance, a program continued by Emperor Valerian. But the Decian and Valerian persecutions failed to stem the tide of conversions. Moreover, there appeared to be some optimism among Christians regarding the benefits of participation, i.e., "being Christian." As Tertullian

[16]Mathematically, $z = \Phi^{-1}(v/c_S)$ diverges as $c_S \downarrow v$, and so the right-hand side of (11) vanishes.

presciently opined,"Torture us! Your...tortures are an enticement to our religion. We became more every time we are hewn down by you: the blood of Christians is seed" (Tertullian, 197). Thus, focal point optimism regarding the benefits of participation in Christianity formed in the late 3rd century (represented by (11) in Proposition 9). This, combined with the religious convictions held by Christians at the time, which in our model would be consistent to a very large θ, allowed them to sustain resistance over a half century of persecution.

The increase in Christian conversions, despite increased persecution, led to the exact commitment problem our theory indentifies. By 311, Galerius, who favored burning Christians alive, abandoned the persecution of Christians, issuing his Proclamation of Toleration, which acknowledged that the persecution was too costly. Like Galerius, Maximinus Daia Caesar had been an avid persecutor, but as noted by Stephen Mitchell, "Maximinus claimed that ... when at a fortunate moment (after being made Caesar in 305) he came to the east and perceived that the outlawing of Christians was leading to a severe drain on the number of persons able to act in the public interest ... he instructed judges to act leniently and to use persuasion not punishment to win them back to pagan beliefs" (Mitchell 1988, p.114). A few years later, Emperor Constantine chose to embrace Christianity rather than fight it. Capturing the concept that it is costly to repress, Constantine observed that, "'the very executioners ... [were] wearied out, and disgusted at the cruelties' ... the crowd ... gave its sympathies to the Christians, rather than to their own governors" (Drake 2002, p.150-1). The attempts to crush Christianity proved too much to bear.

How does our theory inform historical accounts? Constantine's decision to embrace Christianity has many explanations among historians, ranging from the claim that he was a true believer to the claim that he was purely motivated by political advantage (Drake 2002). Although the rise of Christianity during the decline of the Roman Empire was an enormously complicated political event, our theory offers one mechanism to help explain Christianity's remarkable resistance to Roman repression: it was a combination

of religious conviction and focal-point optimism. Our theory provides a novel mechanism that elucidates why Constantine ceded power, and notably, one that makes no reference to Constantine's own religious convictions.

Social Cost of Political Instability. Proposition 3 relates the shared repression cost, c_S, to the success/failure of collective resistance. Since collective resistance can be large but repressed, there is a chance that substantial costs are incurred on both sides.

Variation in signals introduces a coordination friction: if $\theta < \theta^\star$ then the government does not concede, and yet a proportion $1 - F(x^\star - \theta \mid \theta < \theta^\star)$ of citizens participate; in which case citizens and the government suffer the costs of repression. In such scenarios, increasing c_S directly raises the social cost arising from political instability. But importantly, the willingness of citizens and the government to participate and repress respectively also change (via $x^\star$ and $\rho^\star$), implying that the overall social cost of political instability thus depends on the value of θ. For example, if $x^\star$ is decreasing in c_S and if $\theta < \theta^\star$, then a small increase in c_S unambiguously raises the social cost of conflict. Denote the indicator function by $\mathbb{1}$, then the expected cost from political instability is

$$\mathrm{E}[\text{costs}] = \mathrm{E}[c_S \cdot \rho(\theta \mid x^\star)] \cdot \mathbb{1}_{\{\theta \leq \theta^\star\}} = \frac{c_S}{\xi} \int_{-\infty}^{\theta^\star} \rho(\theta \mid x^\star) \phi\left(\tfrac{\theta-\mu}{\xi}\right) d\theta, \tag{12}$$

which is the shared repression cost, c_S, conditional on the government repressing ($\theta < \theta^\star$), weighed by the proportion of citizens who bear it, averaged over θ.

How the expected social cost from political instability changes with the shared repression cost follows by differentiating (12) with respect to c_S, (scaling by ξ):

$$\xi \times \frac{\partial\, \mathrm{E}[\text{costs}]}{\partial c_S} = \overbrace{\int_{-\infty}^{\theta^\star} \rho(\theta \mid x^\star) \phi\left(\tfrac{\theta-\mu}{\xi}\right) d\theta}^{\text{Direct Effect of } c_S} + \underbrace{c_S \rho(\theta \mid x^\star) \phi\left(\tfrac{\theta^\star-\mu}{\xi}\right) \cdot \frac{d\theta^\star}{dc_S}}_{\text{Change in } \theta^\star} + \underbrace{\int_{-\infty}^{\theta^\star} c_S \cdot \frac{d\rho(\theta \mid x^\star)}{dx^\star} \cdot \frac{dx^\star}{dc_S} \phi\left(\tfrac{\theta-\mu}{\xi}\right) d\theta}_{\text{Change in Participation}}.$$

The first term is positive, and is the most straightforward: it is the increased social

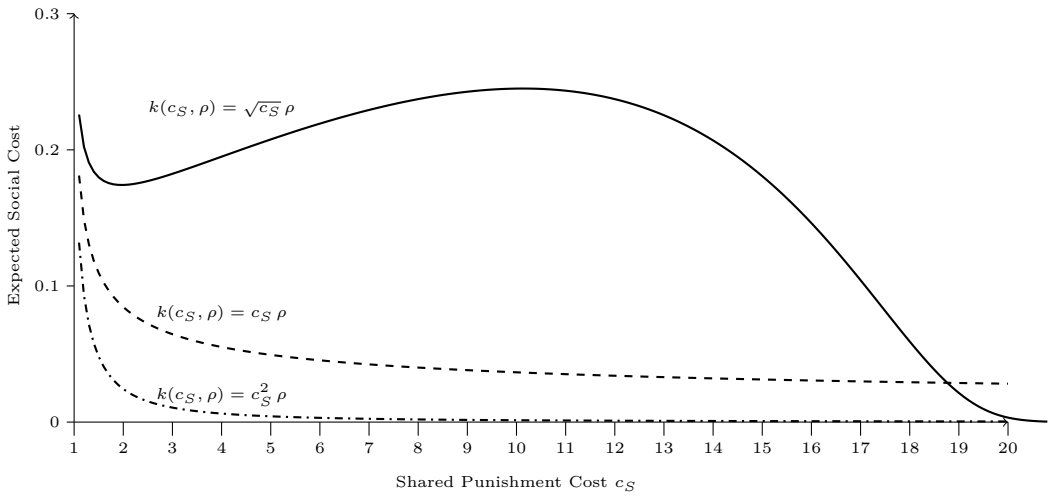


Figure 4: Expected Social Cost of Protest and Shared Punishment Cost Signals are: $x_i \,|\, \theta \sim N(\theta, 1)$; and the prior: $\theta \sim N\,(2, 10)$

cost because increasing c_S implies that repressed citizens pay a greater cost. The other terms reflect the strategic channels. The second term is negative and reflects the change in the government's decision to concede, measured by changes in the critical state $\theta^\star$. The third term is the change in participation (from changes in $x^\star$) resulting from changes in c_S. Propositions 2 and 9 show this last term may be positive or negative.

Figure 4 shows how (12), expressed on the vertical axis, changes with c_S. If the intensity of repression has a greater impact on the government, then as c_S increases, the expected social cost decreases, because the strategic effects dominate. The top solid line is for the case where a change in the scale of repression is relatively more important for the government's costs, and shows how increases in c_S lead to a non-monotonic relationship between the intensity of repression and the total costs that are endured.

Conclusion

In this article we use a model of political resistance to understand how changes in the costs of repression, among other things, influence citizen decisions to participate in collective resistance against a government. A novel feature is that the intensity of repression also

imposes costs on the government implementing repressive policies, which leads to an important commitment problem. Whereas the standard logic suggests that repressive punishments decrease citizen participation in collective resistance, we show this logic to be a special case of a more general formulation.

We identify three novel mechanisms suggesting the standard logic is not the whole story. Firstly, whether the punishment technology determining when repression's intensity, relative to its scale, has a greater effect on the costs incurred by the government from enacting repressive measures. Second, selective benefits to participation, i.e. exclusive repression costs that affect only citizens. Third, coordination incentives, a focal point kind of optimism can also lead increased punishment intensity to increase participation.

Our results illustrate the importance of a government's implementation costs applied to repression. Aside from improving the conceptual understanding of the relationship between repression and resistance, our examples illustrate the prevalence of the strategic channels we identify in substantive cases. Consequently, empirical studies that fail to account for these channels can suffer from important biases, and mispecified counterfactuals. This could be partly responsible for the seemingly inconsistent findings that plague the empirical repression literature (see Davenport (2007) or Hill and Jones (2014)).

References

Ashworth, Scott, Christopher R Berry and Ethan Bueno de Mesquita. 2021. *Theory and Credibility: Integrating Theoretical and Empirical Social Science*. PUP.

Boix, Carles and Milan W. Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-sharing in Dictatorships." *The Journal of Politics* 75(2):300–316.

Casper, Brett Allen and Scott A. Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'etat." *Journal of Politics* 76(2):548–564.

Chassang, Sylvain and Gerard Padró i Miquel. 2010. "Conflict and Deterrence Under Strategic Risk." *Quarterly Journal of Economics* 125(4):1821–1858.

Danneman, Nathan and Emily Hencken Ritter. 2014. "Contagious Rebellion and Preemptive Repression." *Journal of Conflict Resolution* 58(2):254–279.

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

DeNardo, James. 2014. *Power in numbers*. Princeton University Press.

Dewan, Torun and David P. Myatt. 2007. "Leading the Party: Coordination, Direction, and Communication." *American Political Science Review* 101(4):827–845.

Di Lonardo, Livio, Jessica S Sun and Scott A Tyson. 2020. "Autocratic stability in the shadow of foreign threats." *American Political Science Review* 114(4):1247–1265.

Di Lonardo, Livio and Scott A. Tyson. 2022. "Political Instability and the Failure of Deterrence." *The Journal of Politics* 84(1):180–193.

Drake, Harold Allen. 2002. *Constantine and the Bishops: The Politics of Intolerance*. JHU Press.

Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *Review of Economic Studies* 80(4):1422–1458.

Egorov, Georgy and Konstantin Sonin. 2011. "Dictators and their Viziers: Endogenizing the Loyalty–competence Trade-off." *Journal of the European Economic Association* 9(5):903–930.

Egorov, Georgy and Konstantin Sonin. 2021. "Elections in Non-democracies." *The Economic Journal* 131(636):1682–1716.

Garton Ash, Timothy. 1990. *The Magic Lantern: The Revolution of '89 Witnessed in Warsaw, Budapest, Berlin, and Prague*. Vintage Books.

Gates, Scott. 2002. "Recruitment and Allegiance: The Microfoundations of Rebellion." *Journal of Conflict resolution* 46(1):111–130.

Gehlbach, Scott, Konstantin Sonin and Milan W. Svolik. 2016. "Formal Models of Nondemocratic Politics." *Annual Review of Political Science* 19:565–584.

Gibilisco, Michael. 2020. "Decentralization, Repression, and the Gamble for Unity." *Journal of Politics* forthcoming.

Hardin, Russell. 1995. *One for All: The Logic of Group Conflict*. PUP.

Hill, Daniel W. and Zachary M. Jones. 2014. "An Empirical Evaluation of Explanations for State Repression." *American Political Science Review* 108(3):661–687.

Hopkins, Keith. 1998. "Christian Number and its Implications." *Journal of Early Christian Studies* 6(2):185–226.

Kuran, Timur. 1989. "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution." *Public Choice* 61:41–74.

Liang, Zhang, Andrew J. Nathan and Perry Link. 2001. "The Tiananmen Papers: The Chinese Leadership's Decision to Use Force against Their Own People.".

Lichbach, Mark. 1987. "Deterrence or Escalation? The Puzzle of Aggregate Studies of Repression and Dissent." *Journal of Conflict Resolution* 31(2):266–297.

Lichbach, Mark. 1995. *The Rebel's Dilemma*. University of Michigan Press.

Lohmann, Susanne. 1994. "The Dynamics of Information Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-1991." *World Politics* 47(1):42–101.

Lyall, Jason. 2009. "Does indiscriminate violence incite insurgent attacks? Evidence from Chechnya." *Journal of Conflict Resolution* 53(3):331–362.

Mason, T David and Dale A Krane. 1989. "The political economy of death squads: Toward a theory of the impact of state-sanctioned terror." *International Studies Quarterly* 33(2):175–198.

Mitchell, Stephen. 1988. "Maximinus and the Christians in AD 312: A New Latin Inscription." *The Journal of Roman Studies* 78:105–124.

Moore, Will H. 1998. "Repression and Dissent: Substitution, Context, and Timing." *American Journal of Political Science* 42(3):851–873.

Morris, Stephen and Hyun Song Shin. 2003. Global Games — Theory and Applications. In *Advances in Economics and Econometrics, 8th World Congress of the Econometric Society*, ed. M. Dewatripont, L. Hansen and S. Turnovsky. Cambridge University Press pp. 56–114.

Myatt, David P. 2007. "On the Theory of Strategic Voting." *Review of Economic Studies* 74(1):255–281.

Myatt, David P. 2017. "A theory of protest voting." *The Economic Journal* 127(603):1527–1567.

Myerson, Roger. 2004. "Justice, Institutions, and Multiple Equilibria." *Chicago Journal of International Law* 5:91–107.

Nalepa, Monika. 2010. *Skeletons in the Closet: Transitional Justice in Post-communist Europe*. Cambridge University Press.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

Paine, Jack and Scott A. Tyson. 2020. "The Uses and Abuses of Formal Models in Political Science." *SAGE Handbook of Political Science: A Global Perspective* 11.

Pierskalla, Jan H. 2010. "Protest, Deterrence, and Escalation: The Strategic Calculus of Government Repression." *Journal of conflict Resolution* 54(1):117–145.

Pierskalla, Jan H. and Florian M. Hollenbach. 2013. "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa." *American Political Science Review* 107(2):207–224.

Popkin, Samuel L. 1988. Political Entrepreneurs and Peasant Movements in Vietnam. In *Rationality and Revolution*, ed. Michael Taylor. Cambridge: Cambridge University Press pp. 9–62.

Ritter, Emily Hencken and Courtenay R. Conrad. 2016. "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression." *American Political Science Review* 110(1):85–99.

Schelling, Thomas. 1960. *The Strategy of Conflict*. Harvard University Press.

Shadmehr, Mehdi. 2014. "Mobilization, Repression, and Revolution: Grievances and Opportunities in Contentious Politics." *The Journal of Politics* 76(3):621–635.

Shadmehr, Mehdi. 2019. "Tullock Paradox, Pivotal Revolutionaries and the Strength of Weak States." *Quarterly Journal of Political Science* forthcoming.

Shadmehr, Mehdi and Dan Bernhardt. 2011. "Collective Action with Uncertain Payoffs: Coordination, Public Signals, and Punishment Dilemmas." *American Political Science Review* 105(4):829–851.

Shapiro, Jacob N. and Nils B. Weidmann. 2015. "Is the Phone Mightier than the Sword? Cellphones and Insurgent Violence in Iraq." *International Organization* 69(2):247–274.

Stark, Rodney. 1996. *The Rise of Christianity: A Sociologist Reconsiders History*. Princeton University Press.

Svolik, Milan W. 2013. "Contracting on Violence: The Moral Hazard in Authoritarian Repression and Military Intervention in Politics." *Journal of Conflict Resolution* 57(5):765–794.

Tertullian. 197. "The Apology, Chapter 50." *http://www.newadvent.org/fathers* .

Tilly, Charles. 1978. "From mobilization to revolution.".

Truex, Rory. 2019. "Focal points, dissident calendars, and preemptive repression." *Journal of Conflict Resolution* 63(4):1032–1052.

Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11:89–99.

Tyson, Scott A. 2018. "The Agency Problem Underlying Repression." *The Journal of Politics* 80(4):1297–1310.

Tyson, Scott A. and Alastair Smith. 2018. "Dual-layered Coordination and Political Instability: Repression, Co-optation, and the Role of Information." *The Journal of Politics* 80(1):44–58.

Wood, Reed M. 2010. "Rebel capability and strategic violence against civilians." *Journal of Peace Research* 47(5):601–614.

Zeitzoff, Thomas. 2017. "How Social Media is Changing Conflict." *Journal of Conflict Resolution* 61(9):1970–1991.

Zhukov, Yuri M. 2022. "Repression Works." *Mimeo—University of Michigan* .