Repression Works
(just not in moderation)

Yuri M. Zhukov

University of Michigan

September 2022

Abstract

Why does government violence deter political challengers in one context, but inflame them in the next? This paper argues that repression increases opposition activity at low and moderate levels, but decreases it in the extreme. There is a threshold level of violence, where the opposition becomes unable to recruit new members, and the rebellion unravels — even if the government kills more innocents. We find empirical support for this proposition in disaggregated data from Chechnya and a meta-analysis of sub-national conflict dynamics in 71 countries. The data suggest that a threshold exists, but the level of violence needed to reach it varies. Many governments, thankfully, are unable or unwilling to go that far. We explore conditions under which this threshold may be higher or lower, and highlight a fundamental trade-off between reducing government violence and preserving civil liberties.

Repression is the use of violence and intimidation to maintain political power.[1] When confronting behavioral challenges to their authority, governments often respond by threatening and punishing suspected dissidents and rebels. The coercive purpose of these actions is to compel challengers to stop their fight, and to deter others from joining it. The intensity of repression can vary greatly. To reestablish control in Chechnya after 1999, for example, the Russian government used methods ranging from house demolitions and targeted killings to shelling and indiscriminate sweeps. Rebels' responses ranged from quiet acquiescence in one village to violent escalation in the next.

Why does government violence sometimes deter political challengers, but other times inflame them?[2] The dominant view in political science is that violent efforts to maintain power can create grievances that embolden the regime's opponents.[3] Others disagree, noting that repression can deter rebellion by making it unacceptably costly.[4]

These perspectives are not necessarily at odds. We argue that what rebels do depends on how much violence the government uses: repression inflames opposition activity at low and moderate levels, but deters it in the extreme. There is a threshold level of violence, at which repression outpaces the opposition's ability to recover losses. If the government can escalate violence past this point, civilians will believe that supporting the opposition is costlier than supporting the government, and will generally not rebel — even if the government is more responsible for civilian suffering. If the level of repression falls short, repression will invite new and more aggressive behavioral challenges.

We find empirical evidence of this threshold at the sub-national and cross-national levels. We begin with a disaggregated analysis of violence in Russia's Chechnya region. We examine Chechnya due to its prominence in recent literature on political violence (Lyall, 2009, 2010; Toft and Zhukov, 2015), and its geopolitical significance as a "test case," whose lessons other governments have sought to learn (e.g. Ukraine, Kazarin 2014) and emulate (e.g. Syria, Hill 2013). To ensure that the threshold is not unique to Chechnya, we evaluate the generalizability of these results with a meta-analysis of four

---

[1] While definitions of repression can include broader transgression of civil liberties and discrimination (Goldstein 2001, p. xxviii, Davenport 2007, p. 2), our focus is on the use or threat of physical sanctions.

[2] We use the terms "repression" and "government violence" interchangeably below.

[3] This perspective is particularly dominant among social scientists studying civil war and terrorism (Gurr and Lichbach, 1986; Mason and Krane, 1989; Francisco, 1995; Mason, 1996; Lichbach, 1987; Heath et al., 2000; Carr, 2002; Findley and Young, 2007; Saxton and Benson, 2008). Skepticism of repression — especially indiscriminate repression — is also a central theme of the "population-centric" school of counterinsurgency policy research (Galula, 1964; Thompson, 1966; Kitson, 1971; Nagl, 2002; Smith, 2007) and Kilcullen 2009, and is embedded in U.S. counterinsurgency doctrine (FM 3-24).

[4] Examples include Langer (1969); Hibbs (1973); Tilly (1978); Trinquier (1961); Opp and Roehl (1990); Rasler (1996); Lyall (2009); Downes and Cochran (2010); Beissinger (2007); Weyland (2010). Some have argued that coercive effectiveness depends on social network structure (Siegel, 2011), and on whether repression is overt vs. covert (Davenport, 2015), or selective vs. indiscriminate (Lyall, 2010).

widely-used conflict event datasets, covering 71 countries.

The idea of a threshold effect is not new. A rich literature on contentious politics and social movements has hypothesized an "inverted-U" relationship between repression and dissent (Bwy, 1968; Gurr, 1970; Feierabend et al., 1972; Muller and Seligson, 1987; Olivier, 1991; Khawaja, 1993; DeNardo, 2014). The scope of this research, however, has been mainly on protests and forms of resistance short of armed conflict. Most empirical tests, moreover, have relied on macro-level data, and indirect measures of coercion. This article's contribution is to unpack the theoretical logic behind the threshold effect, and to conduct the most comprehensive empirical test yet fielded in the literature. We demonstrate that the threshold effect holds at the sub-national level, and is robust across multiple estimation strategies, countries and datasets.

Beyond showing that the threshold effect exists, our paper contributes to recent research on why states don't always resort to such extreme measures (Conrad and Ritter, 2013; Tyson, 2018). In many cases, mass repression does not occur because it is infeasible. A government may simply lack the resources to do it: the intensity of violence needed to reach the threshold exceeds what the state is capable and willing to produce. In such instances, repression is strictly inflammatory, and never achieves a deterrent effect.

In other cases, mass repression may not occur because it is unnecessary. If the government has highly-accurate information on rebels' identities and whereabouts, it shouldn't need to resort to overwhelming force. If the government can isolate rebels from sources of external support, the rebels become more sensitive to coercion. If the government can restrict freedom of expression, the opposition will have more difficulty making positive appeals to supporters. Under each of these scenarios, the government can reach its threshold at a lower level of violence. Yet these "solutions" all come at a price: coercion becomes less lethal, but the population becomes less free, and the government less publicly accountable. This builds on work by Young (2009), Ritter (2014) and others who have explored how expectations of political survival impact the use of repression.

Governments who defeat their challengers through repression are likely to govern by these same means. A central implication of our argument is that, for a government to maintain its monopoly on the use of force, the policies used to achieve this monopoly must remain in place. Any effort to dismantle the police state — or to tie the regime's hands by, for example, ratifying human rights treaties — risk upsetting this fragile equilibrium, should an opportunistic challenger arrive. This result explains why violence in Chechnya re-emerged in the 1990s despite two centuries of Soviet and Russian efforts to suppress it, from forcible disarmament to mass deportation.

The logic of two-sided coercion pushes governments to repress massively or not at all.

We present tentative evidence that this dilemma — between inaction and mass murder — creates a need to limit the amount of violence needed for governments to stay in power. The resulting institutions (e.g. mass surveillance, travel restrictions, censorship) curtail civil liberties and may explain the emergence of autocracy after civil war.

1 A THEORY OF COERCION IN CIVIL CONFLICT

The scope of this inquiry is on the dynamics of an armed civil conflict, defined as the sustained use of organized violence, by at least two groups of actors within the same state, toward the pursuit (or maintenance) of political power.[5] We do not seek to explain the original causes of rebellion or repression (ala Hegre et al. 2001; Fearon and Laitin 2003; Collier and Hoeffler 2004). Rather, our interest is in the subsequent violent interaction between political actors, and their competitive efforts to build a base of support. The narrative begins after government forces and rebels fail to reach a bargain that both prefer to war (Fearon, 1995; Powell, 2006). The narrative ends when one of the two sides re-establishes a monopoly on the use of force, either through the other party's cessation of violence, or through neutralization of their ability to generate it (Tilly, 1978, p. 7:5).

To establish a theoretical benchmark, we make several simplifying assumptions, each of which we will subsequently loosen to allow for a more realistic conflict environment (Section 3). First, we assume that the quality of the intelligence each side uses to identify and punish opponents is fixed, and that any improvement or deterioration in intelligence occurs too slowly to immediately affect the selectivity of violence. Second, we assume that the government and opposition receive all of their support from the local population, and none from external sources.[6] Our core claims do not depend on these assumptions.

1.1 SUMMARY OF THE ARGUMENT

Consider a stylized conflict zone populated by two combatants — government forces and rebels — and a group of neutral civilians. The combatants both seek to establish a monopoly on the use of force — locally, regionally or country-wide. They pursue this goal by extracting the resources needed to continue fighting and establish a viable state — principally taxes, intelligence, supplies and labor — while denying these same

---

[5] Rather than using a casualty-based inclusion criterion (e.g. 25+ or 1000+ battle deaths), our definition requires only that the violence be two-sided and repeated over multiple interactions.

[6] Although too restrictive to convey the complexities of real-world combat (e.g. most rebel groups since 1975 have received weapons, training, logistical or financial support from external sponsors, Högbladh and Themnér 2011), these assumptions are common in the literature (Kalyvas, 2006; Balcells, 2010).

resources to their opponent. The civilians — whose cooperation both sides need to collect these resources — are interested in security above all else, and will cooperate with one of the sides or remain neutral — whichever is least costly.

To deter civilians from supporting the opponent, each combatant needs to make collaboration as costly as possible, including by killing and capturing more of the opponent's supporters than the opponent can of one's own. The opponent, in turn, has strong incentives to reciprocate. Absent any constraints on the use of force, equilibrium behavior becomes one of mutual escalation, as each side attempts to "outbid" the opponent's use of coercion. Yet mass violence requires significant resources, whose mobilization is constrained by norms, rules of engagement, ammunition stocks, and legal obligations.

To establish a monopoly on the use force, each combatant needs to escalate to the point where the opponent's response would require more resources than it is able to extract. This dynamic implies the existence of a threshold of violence, at which one side is unable to replace its losses with new recruits, and can no longer sustain the fight.

1.2 THE LOGIC OF COERCIVE OUTBIDDING

Let's walk through the main components of this theoretical story, beginning with what each group of actors is trying to achieve. For the combatants, what's at stake is sovereignty: supreme, independent authority over a body politic in a geographic area. The existence of a civil conflict implies that sovereignty is divided, with two or more centers of authority (e.g. government, rebels) making competing, exclusive claims to power (Tilly, 1978). These claims rest, in part, on the ability to extract resources from the population — which is essential not only to the short-term military effort, but also to the establishment of a durable political regime. Because "sovereign power is conferred by the consent of the people" (Hobbes, 2010, p. 162), both combatants seek to maximize the relative size of their political coalitions. A government monopoly emerges when the share of the population actively supporting rebels falls to zero and the government becomes the only actor capable of wielding sustained coercive force. A rebel monopoly is the opposite.

For the civilian population, the stakes are physical survival. Living in a war zone is dangerous, and in deciding whom to support or whether to stay neutral, civilians must first consider the implications of this choice for their safety.

Combatants maximize their shares of popular support by increasing the costs of cooperation with their opponents. As the relative intensity of violence inflicted against a group increases, cooperation with that group becomes more costly. Violence, of course, is not the only tool in the combatants' arsenal, which includes a mix of positive and

negative incentives. Yet if we assume that civilians are risk averse, and prefer a course of action whose worst possible outcome is the least harmful of the alternatives, then negative incentives should ultimately prove more powerful than positive ones. Confronted with the possibility of being harmed by violence, civilians will want to minimize the maximum potential damage they might suffer.

If a combatant has complete information about the decisions civilians make, neutral members of the population should receive no punishment. But combatants cannot inflict costs on opponents with perfect accuracy. A combatant's coercive force will vary in its selectivity, with some proportion of punishment being correctly inflicted against a group's opponents, while another share erroneously befalls neutral civilians. Where selectivity is high, punishment relies on individual criteria (e.g. "target is a known rebel"). Where selectivity is low, punishment uses collective criteria (e.g. "targets live where rebels may be active"), ensnaring otherwise passive individuals. The availability of information depends on various barriers to intelligence collection, like ethno-linguistic differences, rough terrain, and the public's hesitancy to provide tips. The recruitment and cultivation of an informant network, however, is a costly and time-consuming undertaking. While combatants can control the level of violence they use, they typically have far less control — especially in the short term — over how selective that violence will be.

Coercion with low selectivity not only inflicts fewer costs on opponents; it also harms non-combatants. Yet by doing so, indiscriminate violence helps solve the combatants' collective action problem, rendering "free-riding" (i.e. neutrality) potentially more costly than cooperation (Kalyvas and Kocher, 2007). Because civilians absorb damage from both government and rebel violence, staying neutral can become costlier than cooperating with one of the combatants — each of whom only absorbs damage from one side.[7]

Higher levels of indiscriminate violence make it more costly for civilians to stay neutral. Yet these formerly-neutral civilians will not necessarily align themselves against the more indiscriminate side. Assuming this choice is primarily security-driven, civilians should cooperate with the combatants in proportion to rates of survival in each group. If the government can inflict more selective violence against the rebels' supporters than the rebels can against government supporters, then civilians will choose the government. This can still be the case if the government is quite indiscriminate in its use of force. If this violence is so overwhelming that a sufficient number of rebel supporters are punished — even if purely by chance — then cooperation with the ruthless government may still be less costly than cooperation with the better-informed and restrained rebels.

---

[7] A fourth option may be to flee the conflict zone. Yet as long as those fleeing are neutral, their exit does not drive the outcome — it only reduces the size of the "pie" over which the combatants are competing.

Without constraints, equilibrium behavior becomes one of mutual escalation, where each combatants seeks to "outbid" the opponent's use of coercion. But the two sides do not escalate equally. Coercive leverage rests on a demonstrated ability to distinguishing compliant from non-compliant behavior, and to punish non-compliance in a manner appropriate in severity and timeliness. To ensure that cooperation with the opponent remains unacceptably high, the less selective combatant faces a greater incentive to escalate the use of force. The arithmetic is straightforward: if three of my supporters are hit for every four bullets my opponent fires, but I can only hit one of their supporters with every four bullets I fire, then I will need to fire at least three times as many bullets as my opponent just to keep the body count event. The flip side of this calculation is that for every one of my bullets that correctly reaches its intended target, the rest may strike innocent civilians. But I must escalate my violence precisely because it is so indiscriminate. Inefficient force can only maintain its coercive effect if it is so overwhelming that the guilty are punished along with the innocent. Violence is a substitute for information.

To achieve and maintain a monopoly on the use of force, each combatant must escalate violence past a minimum *stalemate threshold*. The system reaches a stalemate when each side matches the other's intensity of violence, adjusted by relative selectivity. Where the government has an intelligence disadvantage — as is common in civil conflicts — it will need to employ a relatively higher level of force to break even. The stalemate is broken when one side escalates violence beyond this threshold, but the other cannot keep up.

The threshold exists because escalation cannot continue indefinitely. Constraints on the use of force can come in the form of societal norms, human rights treaties, restrictive rules of engagement, or even a lack of ammunition. Some of these constraints may be more decisive than others, and their relative salience may vary over time. Collectively, these constraints impose an upper bound on the level of violence each side can use. If the amount of violence the government must produce to reach the threshold exceeds the maximum level feasible under these constraints, then the government will be unable to outbid the rebels and public support will flow overwhelmingly to the opposition. If the threshold is below the government's upper bound, then it should be possible to push the rebels past their breaking point, and restore a government monopoly.

The upper bound varies across combatants. Its value is private information that the other side can try to estimate, but cannot immediately observe. Due to uncertainty over their opponent's upper bound, combatants also cannot precisely know the value of the threshold they must surpass. If this information was common knowledge, the deadly race up the escalatory ladder could in theory be avoided, in favor of a bargain that reflects the known balance of coercive capabilities. Instead, this information is typically

revealed only through the other side's continued ability to withstand coercive pressure.

1.3 Observable implications

Figure 1 plots the expected relationship between government and rebel violence. The solid black curve shows levels of rebel violence associated with each hypothetical level of government violence. The figure assumes that both sides punish at stalemate level, subject to some exogenous constraint, or upper bound. The dashed, diagonal line shows the rebels' response curve without this constraint.

Figure 1: Stalemate threshold. Dashed, diagonal line shows rebels' best responses to government coercion, without upper bounds on violence. Solid black curve shows the same response curve, when violence is subject to upper bounds.
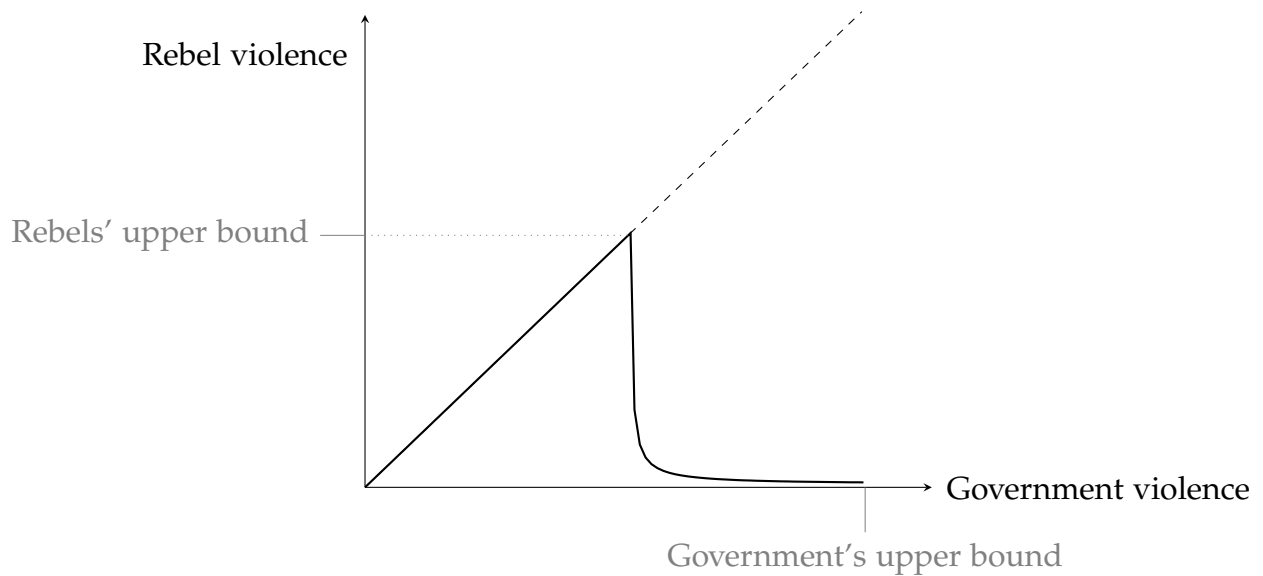
The hypothesized threshold effect is readily visible in Figure 1. Rebel violence first rises in response to increases in government violence, and then drops exponentially. This change occurs where the government escalates to a level that would force the rebels to produce violence at maximum capacity (rebels' upper bound). Because the rebels cannot escalate violence beyond this upper bound, they can no longer maintain a stalemate if the government pushes further. When the government does so, rebels start taking disproportionately high losses, civilians stop cooperating with them, and the opposition becomes unable to replace losses and, eventually, unable to produce violence.

This pattern is starkly different from what would happen in the absence of constraints on violence (dashed line in Figure 1). Here, the two sides continue their coercive outbid-

7

ding, with no breakthrough. The contrast between the two lines highlights an inherent tension between incentives to escalate, and practical limits on military power. As Clausewitz (1832, Book 1, Ch. 1) wrote, "both [combatants] proceed to extremities, to which the only limitations are those imposed by the amount of counteracting force on each side."

2 EMPIRICAL TEST

Our central theoretical claim is that, to suppress a rebellion, a government must escalate violence beyond a stalemate threshold, inflicting more costs on the rebels than the rebels can against the government. We test this proposition with disaggregated data on government and opposition violence in Russia's Chechnya region (2000-2012).[8] We examine the generalizability of these results with disaggregated data on dozens of civil conflicts since 1989, from multiple event datasets covering Africa, Asia, the Americas and Europe.

To analyze the conflict in Chechnya, we use Toft and Zhukov (2015)'s dataset on violence in the North Caucasus, which relies on incident reports from Memorial, a Russian human rights NGO. Within the territory of the Chechen Republic, this dataset includes 35,130 incidents of government violence and 9451 incidents of rebel violence. We aggregated these events to the district-month level, yielding an overall sample size of 2016 (14 districts × 144 months between May 2000 and March 2012). For robustness, we replicated all analyses with regular grid cells (.5×.5 degree PRIO-grid) as spatial units.

The outcome variable, rebel violence (Y_{it}), is a local monthly event count, ranging from 0 to 64 (mean of 5).[9] This measure includes any act of violence (e.g. firefight, ambush, hit-and-run attack, terrorist attack, hostage-taking, bombing) by armed opposition groups like the Chechen Republic of Ichkeria or the Caucasus Emirate. The treatment variable, government violence (T_{it}), is similarly distributed, from 0 to 346 (mean of 17), and includes any violence (e.g. arrests, assassinations, sweeps, search and destroy missions, artillery shelling, air strikes) by Russian security forces and their local allies.[10] Figure 2 shows the temporal and geographic distribution of these data.

2.1 ESTIMATION STRATEGY

To assess whether the expected threshold effect in Figure 1 aligns with real-world events in Chechnya, we estimate a dose-response function (DRF). The DRF represents the conditional expectation of the outcome (i.e. intensity of rebel violence), given each level of

---

[8] These dates exclude the conventional phase of the Second Chechen War in August 1999 - April 2000.

[9] In grid-cell level data, the range is 0 to 149, with a mean of 6.

[10] In grid-cell level data, the range is 0 to 714, with a mean of 21.

Figure 2: Geographic and temporal distribution of Chechnya data. Height of bars in time plot and shading in maps represent the intensity of violence per (a) district or (b) grid cell, over 2000-2012. T = government violence, Y = rebel violence.

(a) Districts   (b) PRIO grid cells
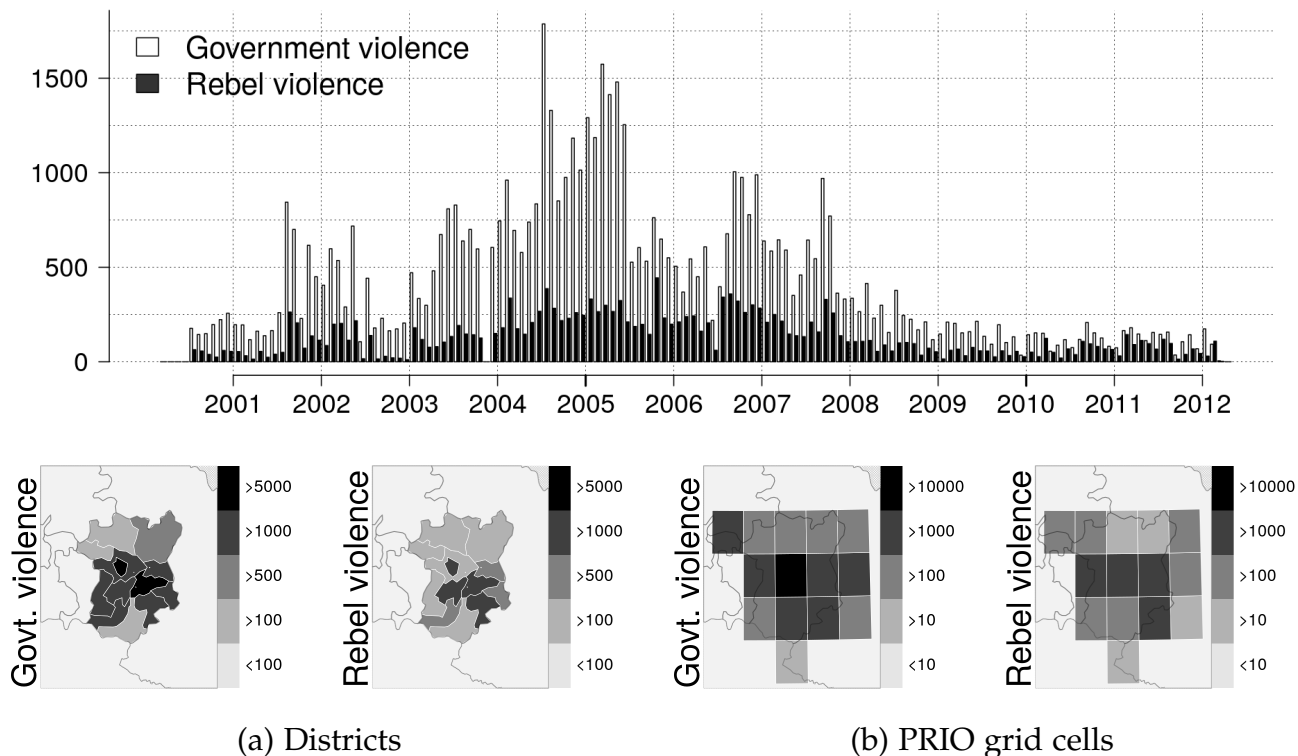
treatment (i.e. recent exposure to government violence):

$$\psi(t) = E[Y_{it}(\tau)] \tag{1}$$

where $Y_{it}(\tau)$ is the potential outcome at $T = \tau$.

A challenge in estimating the DRF is that governments do not repress at random. For example, the theoretical discussion predicts that incentives for escalation are greatest where selectivity is low — due to poor information about one's enemies — and constraints on the use of force are limited. The resulting bias makes it difficult to assess whether variation in rebellion is the outcome of government efforts to suppress it, or of other confounding factors that may have preceded the government's actions.

To adjust for this covariate imbalance — while accounting for the continuous nature of the treatment — we employ inverse generalized propensity score (GPS) weights of the form $w_{it} = f(T_{it})/f(T_{it}|X)$, where $f(T_{it}|X)$ is the conditional density of treatment given covariates X and $f(T_{it})$ is a stabilizing factor based on the marginal probability

9

of treatment (Robins et al., 2000). The logic of this approach is to create a re-weighted dataset, in which more common government actions receive less weight (e.g. higher repression in rugged, difficult-to-control areas), and the level of treatment is weakly unconfounded by observable pre-treatment factors (Imbens, 2000).[11]

The estimation strategy proceeds in three steps. First, we find the GPS using several estimators, including Generalized Linear Models (Hirano and Imbens, 2004; Guardabascio and Ventura, 2013), covariate-balancing generalized propensity scores (CBGPS) (Imai and Ratkovic, 2014), and their non-parametric variants (npGBPS) (Fong et al., 2017).[12]

Second, we use each GPS estimate to calculate inverse probability weights $w_{it} = f(T_{it})/f(T_{it}|X)$, and model the conditional expectation of rebel violence. We begin with a second-order polynomial approximation:

$$\ln(y_{it+1}) = \beta_1 T_{it} + \beta_2 T_{it}^2 + \gamma X_{it-1} + \alpha_i + v_t + u_{it} \tag{2}$$

where y_{it+1} is the local intensity of rebel violence at $t+1$, T_{it} is government violence at t, and X_{it-1} is a matrix of time-variant pre-treatment covariates, including temperature and rainfall. To account for time-invariant local factors, we include district-level (or grid cell-level) fixed effects α_i. Time fixed effects v_t account for common temporal shocks.

To ensure that the quadratic functional form in (2) is not driving the results, we also model the outcome with a threshold regression,

$$\ln(y_{it+1}) = \beta_1 T_{it}\, 1\{q_{it} \leq \tau^*\} + \beta_2 T_{it}\, 1\{q_{it} > \tau^*\} + \gamma X_{it-1} + \alpha_i + v_t + u_{it} \tag{3}$$

where $q_{it} = q(T_{it})$ is a threshold variable, and τ^* is the threshold value, estimated by $\widehat{\tau^*} = \arg\min_{\tau^* \in [\tau_0^*, \tau_1^*]} SSE(\tau^*)$. While the polynomial regression assumes that $E[Y_{it}(\tau)]$ is a continuous function of T, the threshold regression allows the DRF to be discontinuous.

In the last step, we use the results of (2) and (3) to obtain estimates of the full dose-response function, $\widehat{E[Y(\tau)]}$, by estimating the average potential outcome at each level of treatment. As a benchmark, we also report results with unweighted data.

Figure 3 visualizes the covariate balance achieved by each GPS method. The plot reports the distribution of absolute Pearson correlation coefficients between the treatment and each covariate, before and after weighting, with districts (a) and grid cells (b) as

---

[11] Weak unconfoundedness implies pairwise independence of treatment with each potential outcome, $D(\tau) \perp Y(\tau)|X$, where $D(\tau)$ is an indicator of receiving treatment τ.

[12] The GLM estimator uses a Negative Binomial conditional distribution for the treatment, $T_{it}|X_{it-1} \sim NB(\theta X_{it-1} + \iota_i + \zeta_t, k)$, where X denotes pre-treatment covariates, ι_i are local fixed effects, ζ_t are time fixed effects, and k is a dispersion parameter. Following Fong et al. (2017), we use a Box-Cox transformation on the treatment for CBGPS. The npCBGPS makes no distributional assumptions about T.

geographic units. Because CBGPS weighting improves covariate balance the most, we discuss mainly those results below; other estimators yield similar findings.

Figure 3: COVARIATE IMBALANCE. Box plots represent the distribution of absolute Pearson correlation between treatment and each covariate after weighting. Whiskers indicate maxima and minima, boxes indicate upper and lower quartiles, thick lines are medians.

(a) District-month   (b) PRIO grid cell-month
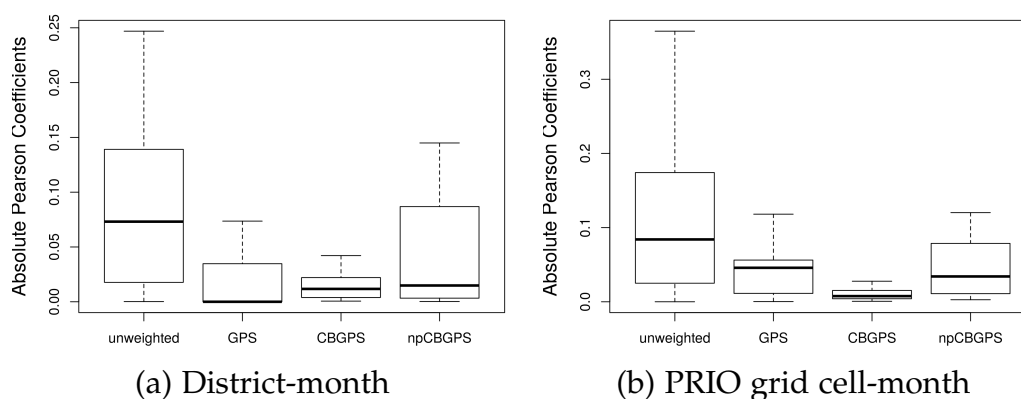
2.2 EVIDENCE FROM RUSSIA'S OPERATIONS IN CHECHNYA

The dynamics of violence in Chechnya provide strong evidence of a threshold effect. Figure 4 reports the average DRF for government and rebel violence in Chechnya at the district-month level, estimated with the polynomial model (4a) and threshold regression (4b).[13] Similarly to the predicted threshold effect in Figure 1, the DRF shows that intermediate levels of repression increase rebel violence, but higher levels decrease it.

The polynomial model (Figure 4a) finds an "upside-down U" relationship between government and rebel violence. In an average locality, fewer than one rebel attack occurred in months following no use of government violence, and 18 attacks if the government escalated to 100 operations per month. This number dropped to less than one attack per month where the government was more extreme (250 operations per month).

Similar patterns emerge from the threshold regression (Figure 4b). Here, the relationship is initially strongly inflammatory: less than one rebel attack following months with no government violence, and 26 attacks after 100 government operations. Once repression exceeds 187 operations per month, however, the model predicts a sharp drop-off from 297 to under 20 rebel attacks per month. After that point, the relationship is flat.

The relative fit of models (2) and (3) depends on units of analysis. Polynomial regression fits district-month data slightly better than threshold regression, with Akaike

---

[13] Results for grid cell-month data are in Appendix A1

Information Criteria (AIC) of 4758.1 versus 4857.1, respectively. In grid cell-month data, threshold regression performed better, with AICs of 5666.0 versus 5460.5.

Figure 4: DOSE-RESPONSE FUNCTION, VIOLENCE IN CHECHNYA (district-month). Dark line is conditional expectation of rebel violence (vertical axis) in the month following each level of government repression (horizontal axis). Shaded area is 95% confidence interval. Short dashes indicate empirical distribution of treatment variable. Inverse probability of treatment weights estimated with CBGPS. Vertical axis on logarithmic scale.

(a) Polynomial regression
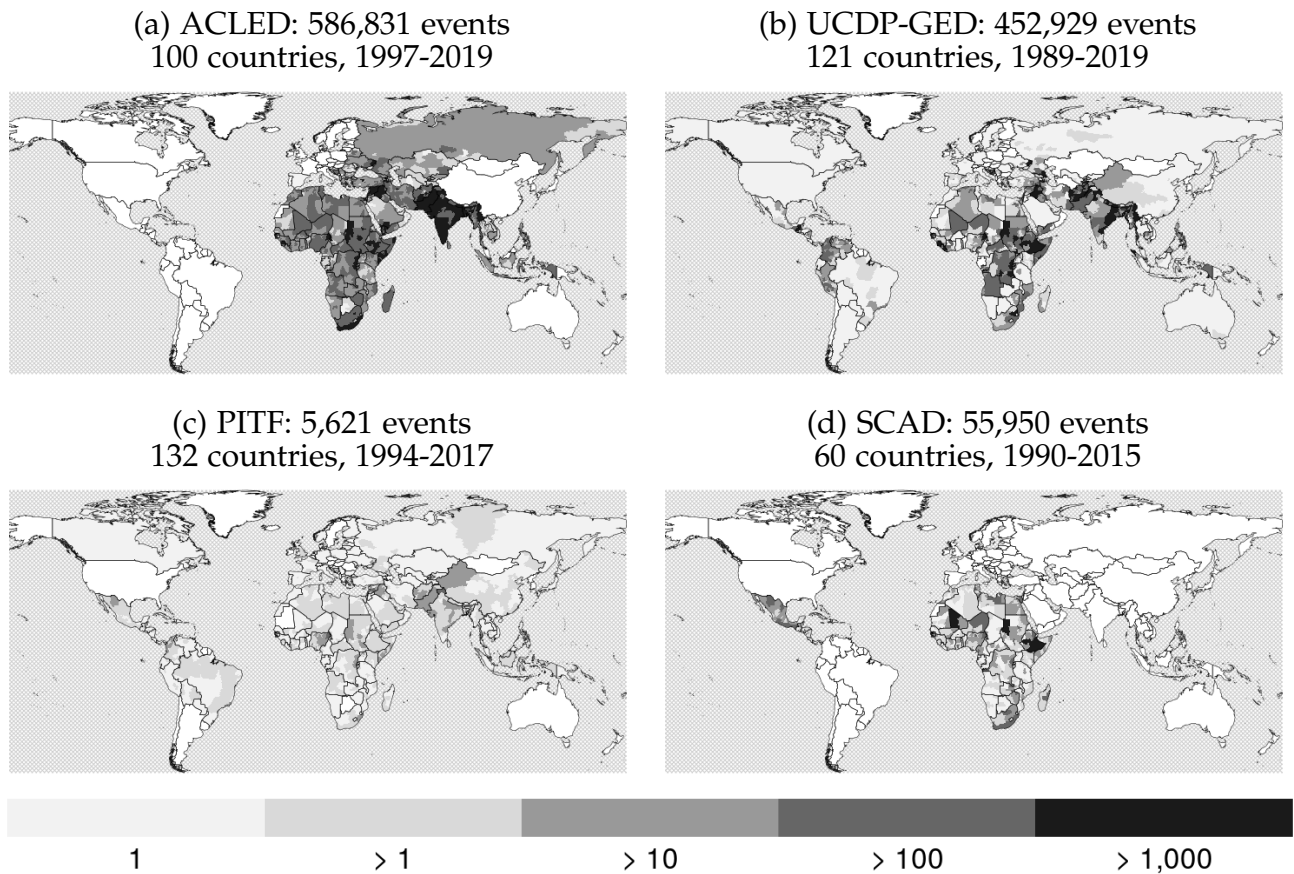
(b) Threshold regression
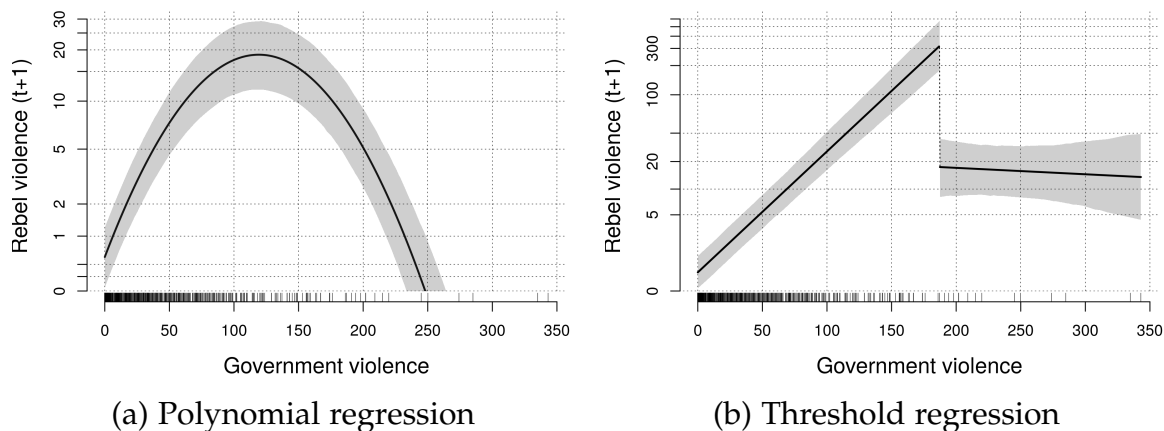
2.3 EVIDENCE FROM CONFLICTS IN 71 OTHER COUNTRIES

Is the threshold effect unique to Chechnya, or part of a broader trend? While the empirical patterns uncovered in this case align closely with theoretical expectations, one may worry that Chechnya is an idiosyncratic outlier, where relatively isolated rebels have confronted an unusually powerful government, one with few material or normative constraints on the use of force. To evaluate the generalizability of the Chechen case, we conducted a meta-analysis of sub-national conflict trends around the globe.

The meta-analysis seeks to replicate the Chechen results with armed conflict data on dozens of countries, from four established multi-national event datasets: Armed Conflict Location and Event Data Project (ACLED) (Raleigh et al., 2010), the UCDP Georeferenced Event Dataset (UCDP-GED) (Sundberg and Melander, 2013), Political Instability Task Force (PITF) Worldwide Atrocities Dataset (Schrodt and Ulfelder, 2016), and Social Conflict Analysis Database (SCAD) (Salehyan et al., 2012). These data sources track individual incidents of political violence, specifying the location and date of an event, the actors involved, and the tactics they used (i.e. "who did what to whom, when and where"). These data vary in geographic and temporal scope, but each covers violence on multiple continents, from the 1990s to the 2010s (Figure 5).

12

To facilitate comparisons across countries and sources, we assembled these events into consistent categories and spatio-temporal units of analysis.[14] We began by organizing events by actor and tactic. Government actions involved state security forces, pro-government militias, and other third parties acting on the incumbent's behalf (e.g. foreign troops, mercenaries); opposition actions involved rebels, dissidents, anti-government militias, rioters and protesters, third parties acting on rebels' behalf, and other groups directly challenging the regime. Figure 5 summarizes the aggregate datasets. To facilitate estimation (see below), we dropped countries with insufficient variation in treatment, reducing the combined sample size to 71 countries across the four datasets.[15]

Figure 5: GEOGRAPHIC SCOPE OF DATA USED IN META-ANALYSIS. Colors denote incidents of government and rebel violence per province, over full period of observation.

(a) ACLED: 586,831 events
100 countries, 1997-2019

(b) UCDP-GED: 452,929 events
121 countries, 1989-2019

(c) PITF: 5,621 events
132 countries, 1994-2017

(d) SCAD: 55,950 events
60 countries, 1990-2015

| 1 | > 1 | > 10 | > 100 | > 1,000 |

To examine variation in the DRF across countries and within them, we estimate a

---

[14] We used the xSub R package (Zhukov et al., 2019) to pre-process the event data. We use "province" as a shorthand for level-1 administrative units, like U.S. states or Russian oblasts.

[15] Because data sparsity varies across data sources, our inclusion criteria required at least 10 unique treatment values per country for ACLED and UCDP-GED, at least 5 for SCAD, and at least 3 for PITF.

series of varying slope and intercept models, first with a polynomial term

$$\ln(y_{ikt+1}) = \beta_{k1} T_{ikt} + \beta_{k2} T_{ikt}^2 + \kappa_k + \gamma X_{ikt-1} + v_t + u_{ikt} \tag{4}$$

and then with a threshold variable

$$\ln(y_{ikt+1}) = \beta_{k1} T_{ikt} \, 1\{q_{ikt} \leq \tau_k^*\} + \beta_{k2} T_{ikt} \, 1\{q_{ikt} > \tau_k^*\} + \kappa_k + \gamma X_{ikt-1} + v_t + u_{ikt} \tag{5}$$

where y_{ikt+1} is the number of rebel attacks in province i, county k during month $t+1$, T_{ikt} is the number of government operations in the same province at t, and X_{ikt-1} are pre-treatment covariates. Each country k has a unique baseline level of violence (κ_k) and a uniquely-shaped relationship between repression and rebellion (β_k). In (5) we use country-specific threshold values τ_k^*.[16] We ran these models separately for each of the four datasets, and used the $\hat{\kappa}_k, \hat{\beta}_k$ parameters to estimate country-specific DRFs, $E[Y_{ikt}(\tau)]$. Although less rigorous with respect to covariate imbalance than our estimates for Chechnya, these models allow us to efficiently assess heterogeneity across conflicts.
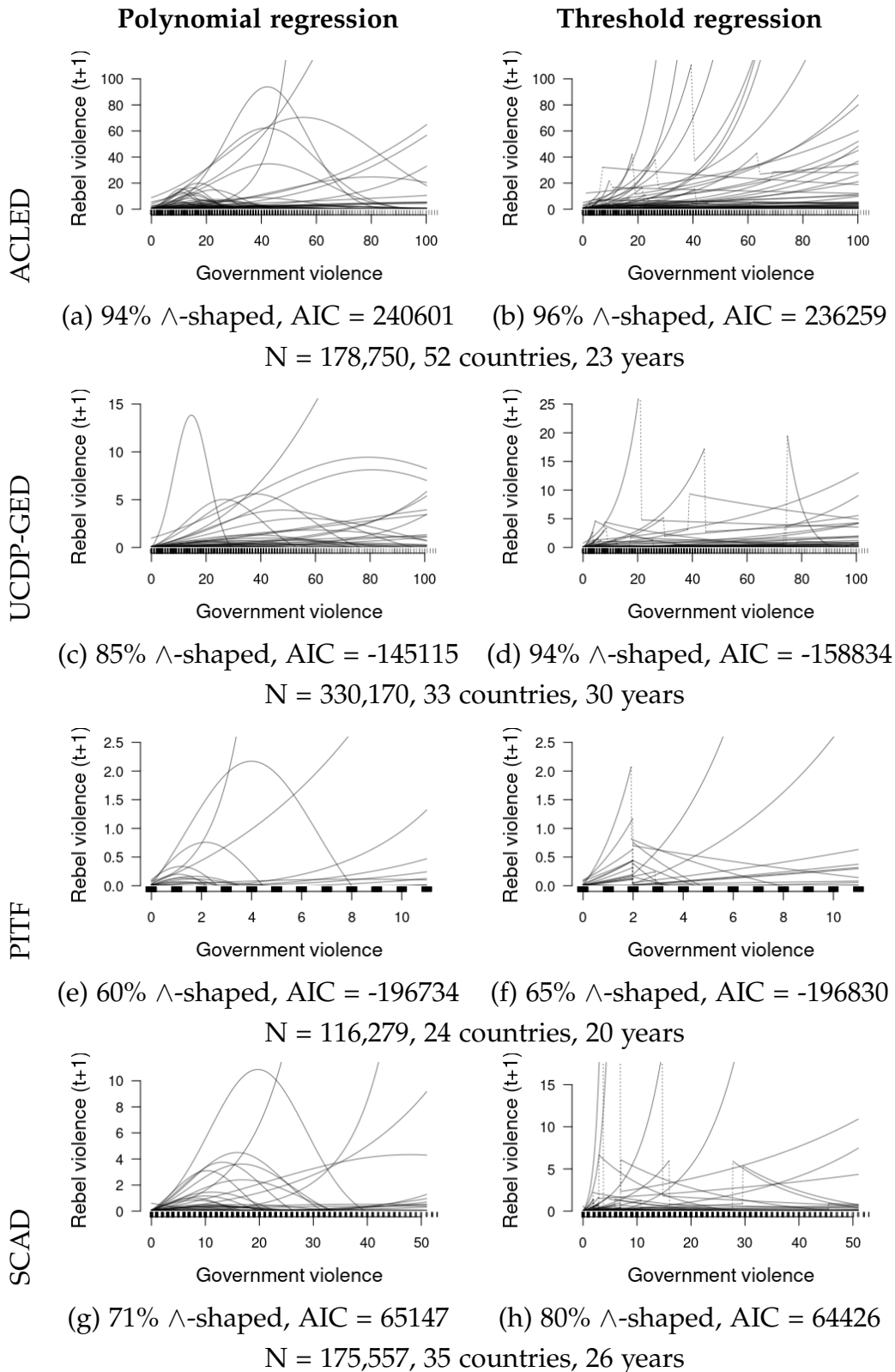
The global meta-analysis largely corroborates evidence from Chechnya. Figure 6 reports the results, with each line representing the estimated DRF for a single country. The left column ("Polynomial regression") reports DRF estimates for model (4), the right column ("Threshold regression") does the same for model (5). For most countries in each dataset, the DRF is $\wedge$-shaped ("upside-down U"). For the polynomial regressions, this concave shape appears in 94% of conflicts in ACLED, 85% is UCDP-GED, 60% in PITF and 71% in SCAD. In threshold regressions, a flip in sign from positive to negative appears in 96% of the conflicts in ACLED, 94% in UCDP-GED, 65% in PITF and 80% in SCAD. The scale of violence, however, varies greatly from country to country, as does the level of repression needed to reach the threshold.

Also visible in Figure 6 are cases where repression is completely inflammatory. In polynomial regressions, the DRF is strictly positive in 4% of conflicts in ACLED, 12% in UCDP-GED, 5% in PITF, and 9% in SCAD. In threshold regressions, it is strictly positive in 2% (ACLED), 6% (UCDP-GED), 5% (PITF), and 3% (SCAD) of countries.

Finally, these results suggest that the threshold specification in model (5) fits the data slightly better that the polynomial specification in model (4). AICs are consistently lower for the threshold regressions, across all four data sources.

---

[16] We estimated $\widehat{\tau_k^*}$ by fitting a restricted form of model (5, dropping κ_k) for each country, and selecting τ values that satisfy $\widehat{\tau_k^*} = \text{argmin}_{\tau_k^* \in [\tau_{k0}^*, \tau_{k1}^*]} SSE(\tau_k^*)$. This procedure assumes independence across countries.

Figure 6: COUNTRY-SPECIFIC DOSE-RESPONSE FUNCTIONS. Each line represents the estimated DRF for a single country in the dataset.

(a) 94% ∧-shaped, AIC = 240601   (b) 96% ∧-shaped, AIC = 236259

N = 178,750, 52 countries, 23 years

(c) 85% ∧-shaped, AIC = -145115   (d) 94% ∧-shaped, AIC = -158834

N = 330,170, 33 countries, 30 years

(e) 60% ∧-shaped, AIC = -196734   (f) 65% ∧-shaped, AIC = -196830

N = 116,279, 24 countries, 20 years

(g) 71% ∧-shaped, AIC = 65147   (h) 80% ∧-shaped, AIC = 64426

N = 175,557, 35 countries, 26 years

3 How high the threshold?

While generally supportive of the theory, our empirical analysis raises two important questions. First, why does the repression-rebellion relationship sometimes resemble a threshold (or "upside-down U") but in other times is strictly inflammatory? Second, why is the threshold high in some cases, and low in others? The second question offers a potential answer to the first: the higher the threshold, the harder it is to reach it.

As is clear even in Chechnya (Figure 4), the empirical distribution of treatment (repression) is highly skewed, with most observations falling on the left tail of the DRF. Much of what we observe in practice, therefore, may be cases where the government uses a level of violence insufficient to "break" the opposition. This censoring may explain why repression often appears counterproductive: escalation to the extreme is empirically rare.

If true, the theoretical implications are quite severe: backlash happens not because governments use repression, but because they sometimes do not repress enough. How much repression is "enough" varies from case to case. The most efficient coercion is one that hardly requires any violence at all. In this sense, a government that can deter with one arrest is more "efficient" than one who can only do so after arresting a thousand.

Why do some governments reach the threshold at a low level of violence, while others escalate to the extreme? One possibility is that some governments enter into these conflicts with institutional features that directly affect how high their threshold is. We now return to the theoretical discussion, and relax some of its more restrictive assumptions. In doing so, we examine several ways in which the government might regulate the level of violence needed to reach the threshold, like expanding surveillance, cutting off opponents' external support, and restricting free speech. We also consider the role of international legal constraints on violence, in the form of of human rights treaties.

3.1 Information

Sustainable deterrence requires that the government not only achieves a threshold effect, but sustains it indefinitely. The alternative to doing so through a massive (and endless) campaign of violence is to develop institutions that make mass violence less essential to regime survival. This applies not only to how governments militarily confront their challengers, but also to how they sustain their monopoly after the military phase ends.

One of the most direct ways to lower the threshold is to collect more information about the opponent's supporters. If the government has a better grasp of who these supporters are and where they hide, it can more selectively use violence against them. Returning to

an analogy from earlier in this paper, a well-informed government can hit more targets with fewer bullets, enabling it to heighten the expected costs of cooperation with rebels at a lower level of effort. This improved efficiency of government violence will push the rebels to use a higher level of violence to maintain a stalemate — increasing the slope of the response curve in Figure 1, and causing the rebels to hit their upper bound sooner.

How can governments improve their information, and make violence more selective? Option one is to enhance human intelligence with a network of local informants. Recruitment of these informants, however, is subject to the same challenges as recruitment of supporters more generally: few will cooperate if it is not safe for them to do so.

Option two is to invest in intelligence capabilities that are less dependent on local support, like surveillance. Solutions here can range from low-tech (e.g. taking a census, clearing forests to improve visibility) to high-tech (e.g. electronic intercepts, CCTV cameras, facial recognition). Surveillance is not a perfect substitute for human intelligence. The information it reveals is more plentiful, but also noisier. Yet by allowing the government to passively monitor the population's movements, contacts and activities — at least those which are most readily visible — surveillance can improve selectivity on the margins. This new flow of information can enable the government to target rebels with higher precision, including where it is very costly for informants to provide tips.

To evaluate the relationship between surveillance and violent repression, we considered whether governments with more secret police agencies behave differently from those with fewer. Secret police agencies, like the Soviet People's Commissariat for Internal Affairs (NKVD) or Uganda's State Research Bureau, are institutions that covertly track, arrest and sometimes kill the incumbent's suspected political opponents. Their means of surveillance can include both electronic eavesdropping and human informants. If secret police agencies improve the government's ability to monitor and selectively punish rebels, we should expect their existence to correlate with a lower coercive threshold.

For each country's estimated DRF in Figure 6, we identified treatment values associated with its mode (polynomial regression) or point of discontinuity (threshold regression). We also recorded whether the curve was strictly increasing, as an indicator that the government — due to some constraint — could not exceed its threshold.

The top two rows of Table 1a report the proportion of governments — across all datasets and models in Figure 6 — whose estimated DRFs were monotonically increasing ($\hat{\pi}$).[17] $\widehat{\pi_{\text{hi sp}}}$ is the proportion for countries with a high number of secret police agencies at the start of the conflict (90th percentile); $\widehat{\pi_{\text{lo sp}}}$ is the proportion for those with fewer.

As Table 1a shows, strictly inflammatory relationships between repression and rebel-

---

[17]Appendix A2 breaks these results down by data source.

lion are less common among governments with more secret police agencies ($\widehat{\pi_{\text{hi sp}}} < \widehat{\pi_{\text{lo sp}}}$). According to polynomial (threshold) regressions, 0% (9%) of states with a high number of secret police agencies had positive DRFs, compared to 4% (33%) for the rest.

For governments whose dose-response curves were ∧-shaped, the remaining rows of Table 1a report median estimated thresholds: $\widehat{\tau_{\text{hi sp}}}$ for countries with more secret police agencies and $\widehat{\tau_{\text{no sp}}}$ for those with fewer. Because the scale of violence varies across countries and datasets, we report estimated threshold values as standard deviations above mean levels of repression per province-month (z-score standardization) and as proportions of maximum repression observed per province-month (Min-Max standardization).

Across all scales and models, countries with more secret police agencies have lower estimated thresholds ($\widehat{\tau_{\text{sp}}} < \widehat{\tau_{\text{no sp}}}$). According to polynomial (threshold) regressions, reaching the threshold requires governments with multiple secret police agencies to escalate violence 5.7 (1.3) standard deviations above the monthly mean. This is a high bar, to be sure, but less than half of the violence that countries with fewer police agencies must use — 11.5 (3.9) standard deviations above the mean. Kolmogorov-Smirnov test statistics are significant at the $p < 0.05$ level or better, rejecting the null hypothesis that the two samples of threshold values are drawn from the same distribution.

| | | (a) Information | | | (b) External support | | | (c) Free discussion | | | (d) Legal constraints | | |
	Model	$\widehat{\pi_{\text{hi sp}}}$	$\widehat{\pi_{\text{lo sp}}}$	K.S.	$\widehat{\pi_{\text{es}}}$	$\widehat{\pi_{\text{no es}}}$	K.S.	$\widehat{\pi_{\text{fd}}}$	$\widehat{\pi_{\text{no fd}}}$	K.S.	$\widehat{\pi_{\text{hr}}}$	$\widehat{\pi_{\text{no hr}}}$	K.S.
	Polynomial	0.00	0.04	0	0.04	0.01	0	0.05	0.03	0	0.04	0.03	0.005
	Threshold	0.09	0.33	0.2*	0.29	0.37	0	0.45	0.20	0.3***	0.40	0.02	0.4***
Scale of $\widehat{\tau}$	Model	$\widehat{\tau_{\text{hi sp}}}$	$\widehat{\tau_{\text{lo sp}}}$	K.S.	$\widehat{\tau_{\text{es}}}$	$\widehat{\tau_{\text{no es}}}$	K.S.	$\widehat{\tau_{\text{fd}}}$	$\widehat{\tau_{\text{no fd}}}$	K.S.	$\widehat{\tau_{\text{hr}}}$	$\widehat{\tau_{\text{no hr}}}$	K.S.
z-score	Polynomial	5.70	11.49	0.4***	10.03	10.03	0.1	16.45	10.03	0.4***	10.03	13.87	0.2
	Threshold	1.29	3.93	0.3**	1.98	1.98	0.2'	1.98	4.64	0.1	3.93	1.98	0.2
Min-Max	Polynomial	0.49	0.51	0.3*	0.52	0.50	0.2*	0.53	0.50	0.3***	0.51	0.52	0.2'
	Threshold	0.16	0.24	0.3**	0.23	0.17	0.2*	0.36	0.19	0.2*	0.23	0.17	0.2

Table 1: VARIATION IN THRESHOLD LEVELS. π is the proportion of countries for which the DRF is /-shaped (monotonically increasing). $\widehat{\tau}$ is the median estimated threshold for countries with a ∧-shaped DRF. "hi sp" ("lo sp") denotes governments with (without) a high number of secret police agencies prior to the conflict (90th percentile). "es" ("no es") denotes governments whose challengers did (not) receive external support. "fd" ("no fd") denotes countries where freedom of discussion was unrestricted (restricted) when conflict began. "hr" ("no hr") denotes countries that had (not) ratified the ICCPR before the conflict began. "K.S." is Kolmogorov-Smirnov test statistic and significance level (***$p < 0.001$, ***$p < 0.01$, *$p < 0.05$, '$p < 0.1$). z-score standardization expresses threshold values as standard deviations above mean levels of repression per province-month. Min-Max standardization rescales threshold levels to the inverval of 0 and 1.

The establishment of secret police agencies, even those that predate a given conflict,

can hardly be considered exogenous to the overall political threat environment facing an incumbent regime. To the extent that governments create these institutions when facing more formidable opponents, however, we would expect the direction of bias to be in the opposite direction of the difference we observe — with higher estimated thresholds in states with more agencies. While a more rigorous assessment of state surveillance lies outside the scope of this study, the patterns reported here are consistent with the view that better-informed governments are less reliant on mass violence.

One of the implications of this discussion is that a demand for surveillance should endure after governments militarily defeat their challengers. A close monitoring of subjects helps the government more selectively punish its challengers, and sends a signal to potential rebels that detection and arrest are almost certain. As more challengers are deterred from acting, the need for a violent government response abates. We should therefore expect post-conflict governments to be especially likely to establish secret police agencies, since doing so helps maintain a monopoly at a lower level of violence.

To test this possibility, we estimate the following generalized linear model:

$$y_{it} = g^{-1}(s(\tau) + \beta_1 x_{it} + \alpha_i + \kappa_t + \epsilon_{it}) \tag{6}$$

where y_{it} is the number of secret police agencies in country i in year t, $g^{-1}()$ is an inverse quasipoisson link, $s(\tau)$ is a cubic polynomial of years since i's most recent intrastate conflict, X_{it} are time-variant covariates, and α_i, κ_t are country and year fixed effects.[18]

As the fitted values in Figure 7a show, states emerging from civil conflict tend to have substantially more secret police agencies than states emerging from protracted periods of peace. This institutional legacy can persist for a generation: the average number of agencies begins to decline from its peak only 40 years after conflict's end.

Do these surveillance capabilities actually help governments stay in power? To answer this question, we estimate a Cox Proportional Hazards model

$$h(t|x_i) = h_0(t)e^{\beta x_i + u_i} \tag{7}$$

where $h(t|x_i)$ is the conditional probability of regime change at t, given survival up to time t, $h_0(t)$ is the baseline hazard, x_i is the establishment of a secret police agency by i between t_0 and t, and u_i are country-clustered standard errors.[19]

---

[18] Full list of secret police agencies in Appendix A3. Intrastate conflict dates are from Correlates of War 4.1 (Singer and Small, 1994).

[19] We define regime change as either a change in excess of one standard deviation in Polity2 scores (Marshall and Jaggers, 2002), or the beginning of an interruption, interregnum or transition (-77,-88,-99).

Figure 7: Institutional legacies of repression, 1816-2016. Quantities represent (a) $E[y_{it}|\tau]$ estimates from Equation 6, or the expected value of each dependent variable, at τ years after civil conflict (fixed effects i = Russia, t = 1991), and (b) Kaplan-Meier estimates from Equation 7, or the estimated fraction of regimes of each type surviving to time t. Shaded region is bootstrapped 95% confidence interval.
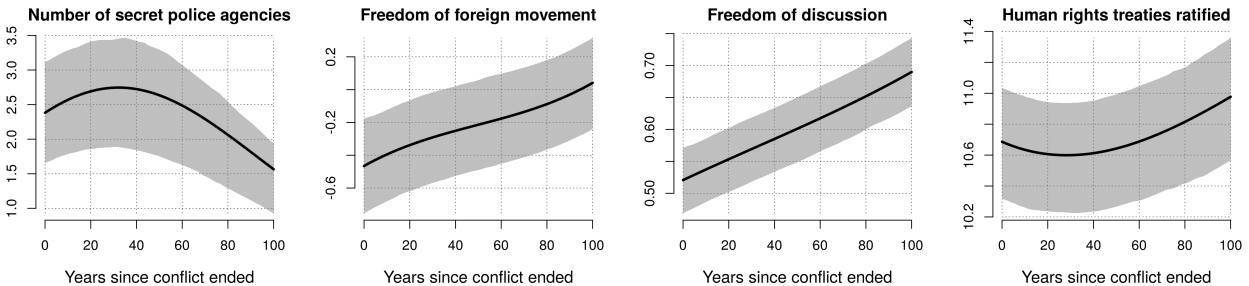
(a) Civil liberties in post-conflict states
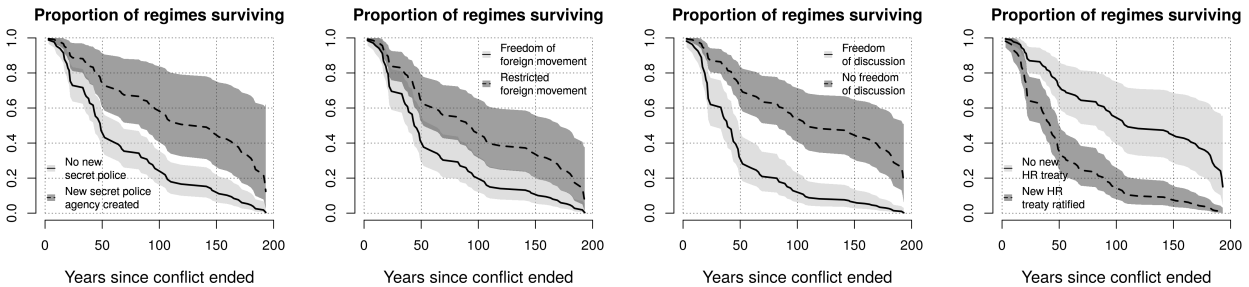
(b) Civil liberties and post-conflict regime survival

Figure 7b reports Kaplan-Meier estimates, representing the fraction of regimes surviving up to year t. Regimes that create new secret police agencies (dashed line) survive longer than post-conflict regimes who do not (solid line). 50 years out, over 75% of governments with new secret police agencies are still standing, versus 47% of those without.

3.2 External support

Another way to lower the threshold is to isolate one's opponents, cutting them off from outside aid. While our initial theoretical discussion assumed that combatants rely exclusively on resources from the local population, external support can complicate the armed struggle in non-trivial ways. Unlike local support, which requires interaction with the local population, external support does not depend on civilian cooperation. By diversifying their support base, and reducing their dependence on local resources, rebels become better equipped to sustain themselves even where it is too costly for locals to support them. This makes it more difficult for the government to reach the threshold

— local repression deters only locals from supporting the rebels, and does not interdict resources flowing in from outside. The government's incentives are to escalate violence, to the point where this external lifeline is insufficient to offset rebels' losses.

To assess whether external support for the opposition elevates the conditions needed for government success, we compared the distributions of $\hat{\pi}$ and $\hat{\tau}$ estimates for subsets of conflicts where rebels did versus did not (initially) receive military support from external sponsors.[20] The results of these analyses are in Table 1b.

While the probability of a strictly increasing DRF does not strongly depend on external support, median estimated thresholds are generally higher in conflicts where rebels had some external backing. The differences, however, are smaller and more uncertain than those we observed when analyzing the role of secret police. This may in part reflect lower statistical power due to missing data.[21] There are also broader inferential challenges in studying the role of external support in this context. First, outside sponsors tend to be strategic about whom they support. For example, they may be more willing to provide support when they believe the rebels are likely to win, or when the government is believed to be committing atrocities. Second, external support can vary over time, and looking at its availability on the conflict's first year — as we did here to ensure that this variable is temporally prior to observed repression — may obscure its full effect.

If external support for rebels does create incentives for more government violence, then cutting off this support should have the opposite effect. By closing borders, setting up roadblocks and otherwise restricting population mobility, the government can reduce the flow of outside goods and personnel. This isolation should make potential rebels more reliant on local support, and hence more vulnerable to government pressure.

Consistent with this claim, cross-national evidence suggests that governments emerging from civil conflict give their citizens less freedom of movement — domestically (Appendix A4) and abroad — than governments without recent armed challenges. Figure 7a shows that citizens in post-conflict environments are less able to travel freely overseas and face higher administrative barriers to emigration, compared to their counterparts in countries without a recent history of conflict.[22] As with mass surveillance, regimes that impose these movement restrictions survive longer than those who do not (Figure 7b).

---

[20] We coded rebels as receiving external support if at least one non-state actor received militarily relevant aid from outside sponsors in the conflict's first year (data from UCDP-ESD, Meier et al. 2022).

[21] While most of the countries in ACLED (90%), UCDP-GED (100%), PITF (100%) and SCAD (86%) are represented in UCDP-ESD, smaller shares had matching UCDP-ESD observations on the year of the first reported event (ACLED 44%, UCDP-GED 35%, PITF 37%, SCAD 24%).

[22] We used the specification in (6), with Gaussian link function $g^{-1}()$ and V-Dem's freedom of foreign movement index as the dependent variable (Pemstein et al. 2018).

3.3 Freedom of discussion

An alternative to coercive violence is the recruitment of supporters through positive inducements, including not only material incentives (Berman et al., 2011; Weinstein, 2007), but also ideological appeals and charismatic leadership. If we loosen the assumption that civilians are solely security-driven, then rebels offering a competitive set of inducements could attract supporters despite severe coercive pressure from the state. This, in turn, incentivizes the government to escalate violence. One way to reduce the opposition's appeal is to control the narrative: restrict freedom of expression, and limit opportunities to publicly criticize the government. By narrowing the scope of permissible speech, such measures hamper the opposition's outreach to sympathetic audiences, and leave the government's own message — and characterization of opponents — unchallenged.

Table 1c suggests that strictly increasing DRFs are more common in countries with unrestricted freedom of discussion ($\widehat{\pi_{\text{fd}}} > \widehat{\pi_{\text{no fd}}}$).[23] According to polynomial (threshold) regressions, among states where citizens were able to talk openly about politics — privately or publicly, without fear of prosecution — 5% (45%) never reached their thresholds. Among states without free discussion, the number was 3% (20%).

Table 1c further shows that countries where freedom of discussion is respected have higher estimated thresholds ($\widehat{\tau_{\text{fd}}} > \widehat{\tau_{\text{no fd}}}$). In 3 of 4 cases, differences in medians are positive, with Kolmogorov-Smirnov test statistics significant at the $p < 0.01$ level.

Incentives to silence the opposition persist after the violence ends. As Figure 7 shows, post-conflict regimes tend to have less freedom of discussion.[24] Such states also tend to have fewer critical print and broadcast media outlets, and higher levels of censorship overall (Appendix A4). From the standpoint of regime survival (Figure 7b), these measures seem to pay: post-conflict governments without free expression or critical media last longer, on average, than post-conflict governments that respect freedom of speech.

3.4 Normative and legal constraints

If there were no limits on the amount of violence a government can produce, there would be little need to lower the threshold by transforming the political environment. Yet these limits do exist in various forms, including the need to comply with legal obligations and international norms. For example, human rights treaties may constrain states by creating costs for noncompliant behaviors (see reviews in Simmons, 2010; Matanock, 2020).

---

[23] We coded country-years as "free" if they received above average V-Dem freedom of discussion scores on the first year of the conflict.

[24] The specification here is the same as in (6), with V-Dem's freedom of discussion index as the outcome.

In the context of our theory, treaty ratification may relate to government violence in two ways. First, compliance with human rights treaties may lower states' upper bound on violence, making it harder for them to reach the threshold. The observable implication would be a higher incidence of monotonically increasing DRFs among ratifying states. Alternatively, governments that ratify these treaties may face less severe challenges, and may already have some of the institutions in place to avoid the need for escalation. The empirical implication here would be fewer strictly increasing DRFs and lower estimated thresholds among ratifying states. Such a pattern would be consistent with the view that observed differences between ratifiers and non-ratifiers are the result of selection rather than behavioral change. The first pattern would instead suggest that treaties can elicit compliance beyond what could be explained by selection alone.

We find traces of both patterns in the data, but — at least during the active phase of a conflict — evidence of compliance is stronger than evidence of selection. We compared the DRFs of countries that had ratified the 1966 International Covenant on Civil and Political Rights (ICCPR) prior to the conflict to those that had not. Past research has used ratification of the ICCPR as a measure of international human rights obligations because it covers a broad range of rights — prohibiting torture, cruel or degrading punishment, arbitrary detention, and infringement of free speech (Lupu, 2015).

Table 1d shows that monotonically increasing DRFs are significantly more likely among ratifying states than non-ratifying states ($\widehat{\pi_{hr}} > \widehat{\pi_{no\,hr}}$). According to the polynomial (threshold) model, 4% (40%) of states that had ratified the ICCPR failed to reach their thresholds, compared to 3% (2%) for states that had not ratified it. Among states that do reach their thresholds, however, the relationship between ratification and violence is relatively weak. Polynomial regressions suggest the median threshold is lower for ratifying states ($\widehat{\tau_{hr}} < \widehat{\tau_{no\,hr}}$) — consistent with selection — but threshold regressions suggest the opposite ($\widehat{\tau_{hr}} > \widehat{\tau_{no\,hr}}$). Kolmogrov-Smirnov statistics are insignificant for three sets of estimates, and marginally significant ($p < 0.1$) for the fourth. This evidence is too weak to support the selection interpretation, but also insufficient to definitively rule it out.

To the extent that human rights treaties constrain government violence during active confrontations with rebels — as Table 1d suggests they might — we could expect governments emerging from civil conflict to be more reluctant to ratify new treaties of this kind. Using the specification in (6), we regressed the total number of human rights treaties a country had ratified on the number of years since its last civil conflict.[25] As

---

[25] The full list of treaties includes the 1965 International Convention on the Elimination of All Forms of Racial Discrimination, the 1966 ICCPR (and its optional protocols), the 1966 International Covenant on Economic, Social and Cultural Rights (and optional protocols), the 1979 Convention on the Elimination of All Forms of Discrimination against Women (and optional protocols), the 1984 Convention against

Figure 7a shows, there is a decline in ratification in the two decades after a conflict's end, and a gradual increase thereafter. This pattern may be suggestive of post-conflict states selecting out of human rights treaties, but the predictions are highly uncertain.

One area where the evidence is slightly stronger is on the relationship between treaty ratification and the longevity of post-conflict regimes. The rightmost pane in Figure 7b reports Kaplan-Meier curves for governments that ratified at least one new treaty post-conflict (dashed line) and governments that ratified no new treaties (solid line). Of these two, governments that had not ratified any new treaties survive significantly longer, on average: 50 years out, 72% of non-ratifiers remain standing, compared to 35% of ratifying regimes. The same legal mechanisms that constrain violent repression during conflict may also make it harder for regimes to survive peacetime challenges.

3.5 WHY POLICE STATES EMERGE

Almost all of the steps one might take to lower the threshold — and thereby reduce the violence needed for a government to stay in power — are ones that make the population less free. This pattern points to a deep institutional legacy of political violence. The idea that "war makes states" is deeply embedded in our understanding of political-economic development (Tilly, 1985). Research on interstate wars has found that the peaceful settlement of disputes increases the prevalence of democracy (Gibler and Tir, 2010; Rasler and Thompson, 2004), and that the positive relationship between war termination and democracy does not depend on whether a country wins or loses (Reiter, 2001).

Questions of causality aside (Kim and Rousseau, 2013; Hegre, 2014), the apparently positive correlation between war termination and democracy does not seem to carry over to contemporary civil conflicts. As Gurr (1988) observed, regimes that successfully use coercion to survive armed challenges are likely to rely on coercion in response to future challenges. The dilemma is how to make this coercion less overt, but no less credible. Governments navigate this dilemma by building institutions that make mass violence less essential to regime survival — improving surveillance, cutting ties to the outside world, restricting free speech. The result is a polity that is less violent, but also less free.

Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (and optional protocols), the 1989 Convention on the Rights of the Child (and optional protocols), the 1990 International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families, the 2006 International Convention for the Protection of All Persons from Enforced Disappearance, and the 2006 Convention on the Rights of Persons with Disabilities (and its optional protocols).

4 CONCLUSION

The central finding of this article is grim: repression works, just not in moderation. Government violence can suppress rebellion, but only if that violence is sufficiently high to convince civilians that supporting the rebels is more costly than supporting the government. If the government is unable or unwilling to escalate to this point, it will only provoke reciprocal escalation by the rebels. This non-monotonic relationship holds in dozens of modern civil conflicts, across multiple datasets and units of analysis, and is robust to multiple estimation strategies.

These findings help illuminate the wartime origins of autocracy: efforts that governments take to reduce the amount of violence needed to stay in power (e.g. surveillance, travel restrictions, censorship) are also ones that make its citizens less free. A citizen under constant monitoring, with no freedom of movement or expression, is a citizen with few opportunities to rebel. Such a citizen may feel a strong motivation to oppose her government, and many of her compatriots may agree. Yet if she cannot organize and maintain an armed struggle, her rebellion will not succeed.

Our paper offers several lessons for future research. First, it illustrates the value of meta-analysis in establishing the extent to which empirical results are transportable to other geographic contexts, time periods, treatments, and units of analysis. This approach can complement more bespoke single-country sub-national work, which prioritizes internal validity over generalizability. Integrating event data into a set of consistent definitions, measures, and units can remove key barriers to the accumulation of knowledge.

Second, we urge future researchers to further explore estimation methods for continuous and multi-valued treatments (e.g. intensity of government violence, not just its presence or absence). Doing so requires confronting thorny selection problems, since situations where the government represses at a high level are likely to be systematically different from situations where it does not. But this approach can help reveal important insights about the shape of the repression-dissent relationship, and why it varies.

The purpose of such research, needless to say, is not to advise dictators on how to repress their own people. They do not need such advice. As we have shown, political actors of many stripes already act in a manner consistent with our theory's predictions. If we are to understand why these acts of unspeakable cruelty happen, it is necessary to examine the incentives their perpetrators face, and how their targets are likely to react.

References

Balcells, L. (2010). Rivalry and revenge: Violence against civilians in conventional civil wars1. *International Studies Quarterly*, 54(2):291–313.

Beissinger, M. R. (2007). Structure and example in modular political phenomena: The diffusion of Bulldozer/Rose/Orange/Tulip Revolutions. *Perspectives on Politics*, 5(2):259–276.

Berman, E., Callen, M., Felter, J. H., and Shapiro, J. N. (2011). Do working men rebel? insurgency and unemployment in Afghanistan, Iraq, and the Philippines. *Journal of Conflict Resolution*, 55(4):496–528.

Bwy, D. P. (1968). Political instability in latin America: the cross-cultural test of a causal model. *Latin American Research Review*, 3(2):17.

Carr, C. (2002). *The Lessons of Terror*. Random House, New York.

Clausewitz, C. v. (1832). *On War*. Princeton University Press, Princeton, NJ.

Collier, P. and Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers New Series*, 56(563-595).

Conrad, C. R. and Ritter, E. H. (2013). Treaties, tenure, and torture: The conflicting domestic effects of international law. *The Journal of Politics*, 75(2):397–409.

Davenport, C. (2007). State repression and political order. *Annual Review of Political Science*, 10:1–23.

Davenport, C. (2015). *How Social Movements Die*. Cambridge University Press.

DeNardo, J. (2014). *Power in numbers: The political strategy of protest and rebellion*. Princeton University Press.

Downes, A. and Cochran, K. M. (2010). Targeting civilians to win? assessing the military effectiveness of civilian victimization in interstate war. In Lawrence, A. and Chenoweth, E., editors, *Rethinking Violence: States and Non-State Actors in Conflict*. MIT Press, Cambridge, MA.

Fearon, J. (1995). Rationalist explanations for war. *International Organization*, 49(3):379–414.

Fearon, J. and Laitin, D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1):75–90.

Feierabend, I. K., Feierabend, R. L., and Gurr, T. R. (1972). *Anger, violence, and politics: theories and research*. Prentice Hall.

Findley, M. G. and Young, J. K. (2007). Fighting fire with fire? how (not) to neutralize an insurgency. *Civil Wars*, 9(4):378–401.

Fong, C., Hazlett, C., and Imai, K. (2017). Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements.

Francisco, R. A. (1995). The relationship between coercion and protest an empirical evaluation in three coercive states. *Journal of Conflict Resolution*, 39(2):263–282.

Galula, D. (1964). *Counterinsurgency Warfare*. Praeger, New York.

Gibler, D. M. and Tir, J. (2010). Settled borders and regime type: Democratic transitions as consequences of peaceful territorial transfers. *American Journal of Political Science*, 54(4):951–968.

Goldstein, R. J. (2001). *Political repression in modern America from 1870 to 1976*. University of Illinois Press.

Guardabascio, B. and Ventura, M. (2013). Estimating the dose-response function through the glm approach.

Gurr, T. R. (1970). *Why Men Rebel*. Princeton University Press, Princeton, NJ.

Gurr, T. R. (1988). War, revolution, and the growth of the coercive state. *Comparative Political Studies*, 21(1):45–65.

Gurr, T. R. and Lichbach, M. I. (1986). Forecasting internal conflict: A competitive evaluation of empirical theories. *Comparative Political Studies*, 19(1):3–38.

Headquarters, Department of the Army (2014). *Insurgencies and Countering Insurgencies, Field Manual No. 3-24*. Headquarters, Department of the Army, Washington, DC.

Heath, J. A., Mason, T. D., Smith, W. T., and Weingarten, J. P. (2000). The calculus of fear: Revolution, repression, and the rational peasant. *Social Science Quarterly*, 81(2):622–633.

Hegre, H. (2014). Democracy and armed conflict. *Journal of Peace Research*, 51(2):159–172.

Hegre, H., Ellingsen, T., Gates, S., and Gleditsch, N. P. (2001). Toward a democratic civil peace? democracy, political change, and civil war, 1816-1992. *American Political Science Review*, 95(01):33–48.

Hibbs, D. A. (1973). *Mass Political Violence*, volume 253. John Wiley.

Hill, F. (2013). The real reason putin supports assad: Mistaking Syria for chechnya. *Foreign Affairs*.

Hirano, K. and Imbens, G. W. (2004). The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, 226164:73–84.

Hobbes, T. (1651/2010). *Leviathan, revised edition*. Broadview Press.

Högbladh, Stina, T. P. and Themnér, L. (2011). External support in armed conflict 1975–2009. presenting new data. Paper presented at the 52nd Annual International Studies Association Convention, Montreal, Canada, 16-19 March, 2011.

Imai, K. and Ratkovic, M. (2014). Robust estimation of inverse probability weights for marginal structural models. *Journal of the American Statistical Association*, 110(just-accepted):00–00.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, 87(3):706–710.

Kalyvas, S. N. (2006). *The Logic of Violence in Civil War*. Cambridge.

Kalyvas, S. N. and Kocher, M. A. (2007). How 'free' is free-riding in civil war? violence, insurgency, and the collective action problem. *World Politics*, 59(2):177–216.

Kazarin, P. (2014). Chechenskie uroki dlya ukrainy [chechen lessons for Ukraine]. *LB.ua*.

Khawaja, M. (1993). Repression and popular collective action: Evidence from the west bank. In *Sociological Forum*, volume 8, pages 47–71. Springer.

Kilcullen, D. (2009). *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One*. Oxford University Press, USA, New York.

Kim, H. M. and Rousseau, D. L. (2013). The reciprocal relationship between military conflict and democracy. *Defence and Peace Economics*, 24(1):47–72.

Kitson, F. (1971). *Low Intensity Operations: Subversion, Insurgency and Peacekeeping*. Faber & Faber, Limited.

Langer, W. L. (1969). *Political and social upheaval, 1832-1852*. Number v. 14 in Rise of modern Europe. Harper & Row.

Lichbach, M. I. (1987). Deterrence or escalation? : The puzzle of aggregate studies of repression and dissent. *Journal of Conflict Resolution*, 31.

Lupu, Y. (2015). Legislative veto players and the effects of international human rights agreements. *American Journal of Political Science*, 59(3):578–594.

Lyall, J. (2009). Does indiscriminate violence incite insurgent attacks? evidence from chechnya. *Journal of Conflict Resolution*, 53(3):331–362.

Lyall, J. (2010). Are coethnics more effective counterinsurgents? evidence from the second chechen war. *American Political Science Review*, 104(1).

Marshall, M. G. and Jaggers, K. (2002). Polity iv project: Political regime characteristics and transitions, 1800-2002.

Mason, T. D. (1996). Insurgency, counterinsurgency, and the rational peasant. *Pubic Choice*, 86(1/2):63–83.

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

Matanock, A. M. (2020). How international actors help enforce domestic deals. *Annual Review of Political Science*, 23:357–383.

Meier, V., Karlén, N., Pettersson, T., and Croicu, M. (2022). External support in armed conflicts: Introducing the ucdp external support dataset (esd), 1975–2017. *Journal of Peace Research*, page 00223433221079864.

Muller, E. N. and Seligson, M. A. (1987). Inequality and insurgency. *American political science Review*, 81(2):425–451.

Nagl, J. A. (2002). Learning to eat soup with a knife: Counterinsurgency lessons from malaya and Vietnam.

Olivier, J. L. (1991). State repression and collective action in South Africa, 1970–84. *South African Journal of Sociology*, 22(4):109–117.

Opp, K.-D. and Roehl, W. (1990). Repression, micromobilization, and political protest. *Social Forces*, 69(2):521–547.

Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y.-t., Krusell, J., and Miri, F. (2018). The v-dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data. *University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21, 3d edition.*

Powell, R. (2006). War as a commitment problem. *International Organization*, 60(1):169–203.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: An armed conflict location and event dataset special data feature. *Journal of Peace Research*, 47(5):651–660.

Rasler, K. (1996). Concessions, repression, and political protest in the Iranian revolution. *American Sociological Review*, 61(1):132–152.

Rasler, K. and Thompson, W. R. (2004). The democratic peace and a sequential, reciprocal, causal arrow hypothesis. *Comparative Political Studies*, 37(8):879–908.

Reiter, D. (2001). Does peace nurture democracy? *The Journal of Politics*, 63(3):935–948.

Ritter, E. H. (2014). Policy disputes, political survival, and the onset and severity of state repression. 58(1):143–168.

Robins, J. M., Hernan, M. A., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560.

Salehyan, I., Hendrix, C. S., Hamner, J., Case, C., Linebarger, C., Stull, E., and Williams, J. (2012). Social conflict in Africa: A new database. *International Interactions*, 38(4):503–511.

Saxton, G. and Benson, M. (2008). Means, motives and opportunities in ethno-nationalist mobilization. *International Interactions*, 34(1):53–83.

Schrodt, P. A. and Ulfelder, J. (2016). Political instability task force atrocities event data collection, version 1.1b1.

Siegel, D. A. (2011). When does repression work? collective action in social networks. *The Journal of Politics*, 73(4):993–1010.

Simmons, B. (2010). Treaty compliance and violation. *Annual review of political science*, 13:273–296.

Singer, J. D. and Small, M. (1994). Correlates of war project: International and civil war data, 1816-1992 (ICpsr 9905). *Ann Arbor, MI: Inter-University Consortium for Political and Social Research.*

Smith, R. (2007). *The Utility of Force: The Art of War in the Modern World.* Knopf Doubleday Publishing Group.

Sundberg, R. and Melander, E. (2013). Introducing the ucdp georeferenced event dataset. *Journal of Peace Research*, 50(4):523–532.

Thompson, R. S. (1966). *Defeating Communist Insurgency*. Praeger, New York.

Tilly, C. (1978). *From Mobilization to Revolution*. Addison-Wesley.

Tilly, C. (1985). War making and state making as organized crime. In Evans, P. and D. Rueschemeyer, T. S., editors, *Bringing the State Back In*, pages 169–191. Cambridge.

Toft, M. D. and Zhukov, Y. M. (2015). Islamists and nationalists: Rebel motivation and counterinsurgency in Russia's north caucasus. *American Political Science Review*, 109(2):222–238.

Trinquier, R. (1961). *La Guerre moderne*. La Table ronde, Paris.

Tyson, S. A. (2018). The agency problem underlying repression. *Journal of Politics*, 80(4):1297–1310.

Weinstein, J. (2007). *Inside Rebellion: The Politics of Insurgent Violence*. Cambridge University Press, Cambridge.

Weyland, K. (2010). The diffusion of regime contention in European democratization, 1830-1940. *Comparative Political Studies*, 43(8/9):1148–1176.

Young, J. K. (2009). State capacity, democracy, and the violation of personal integrity rights. *Journal of Human Rights*, 8(4):283–300.

Zhukov, Y. M., Davenport, C., and Kostyuk, N. (2019). Introducing xsub: A new portal for cross-national data on sub-national violence. *Journal of Peace Research*, 56(4):604–614.