Snakes in the Backyard*

Ravi Ashok Satpute †

January 31, 2026
Job Market Paper
Latest Version: link

Abstract

Why do states empower domestic armed groups that may later challenge their power? I develop a signaling model in which a state with private information about its strength chooses between direct violence and delegation to a non-state violent proxy. Direct attacks are informative, success signals strength, and failure reveals weakness. Delegation to a proxy obfuscates this mapping but endogenously creates a political claimant who may later contest power with private information about the cost of doing so. I characterize *Perfect Bayesian Equilibria* and show that in the key equilibrium, powerful states attack directly while weak states delegate, accepting blowback risk because revealing weakness immediately through failed direct action is politically costlier than potentially creating a future political competitor. The model provides a formal explanation for why proxy delegation is systematically associated with state weakness and endogenous political instability.

JEL Classification Codes: D72, D74, D82

Keywords: Proxy Warfare, Conflict, Endogenous Blowback

*I thank

†Economics Group, Indian Institute of Management Calcutta; Email: satputerafp21@email.iimcal.ac.in

1 Introduction

Why do states deliberately empower domestic violent actors that they eventually cannot fully control? From paramilitaries in Colombia and Mexico to Sudan's Janjaweed, from Hezbollah in Lebanon to Russia's Kadyrovtsy, governments across regime types tolerate, sometimes even cultivate, and use armed non-state groups that later contested political power, extracted concessions, or became autonomous centers of coercion. The pattern is puzzling. Canonical theories of state formation emphasize the imperative to monopolize violence (Weber, 1946, Tilly, 1985, North et al., 2009, Myerson, 2008); states that tolerate rival coercive organizations may invite challenges to their authority and undermine the foundations of political order. Yet empirically, the deliberate empowerment and usage of such violent non-state actors, what I term *proxy delegation*, is widespread and systematic.

The phenomenon spans historical periods and political systems. Autocracies deploy militias to enforce political control while maintaining distance from atrocities: Syria's Shabiha, Sudan's Janjaweed, and Zimbabwe's war veterans operated with state support but outside formal command structures (Ahram, 2011, Carey et al., 2015). Even democracies have historically relied on violent groups to pursue objectives that formal security forces could not openly pursue: paramilitaries in Colombia and Northern Ireland, vigilantes in the Philippines and Mexico, operated in gray zones between state and non-state violence (Mazzei, 2009, Carey et al., 2013, 2015)[1]. Acemoglu et al. (2013) document how the Colombian state's electoral incentives and relationship with paramilitaries undermined its monopoly on violence, with armed groups later becoming autonomous political actors. In most of these settings, the proxy relationship created risks that materialized over time: groups turned against patrons, demanded political representation, or leveraged their coercive capacity for autonomous ends.

The costs of such arrangements are substantial and often predictable. Proxies accumulate organizational capacity, territorial control, and popular legitimacy through the violence they conduct. Success increases popularity and smoothens the recruitment of new fighters, but, importantly, it also breeds ambition. Armed groups that prove effective at violence frequently discover that the same capabilities translate into political leverage; leverage that may be directed against the state itself. This dynamic, which I call *blowback*, represents a fundamental tension in proxy relationships: the very effectiveness that makes proxies useful

[1] In the international context the US relied on proxies in Nicaragua and Afghanistan in the 1980s

also makes them dangerous for the state.

There are three plausible explanations across the literature from diverse fields for state reliance on violent proxies. First, proxies are *militarily efficient* in that they are cheaper, more specialized, possess local intelligence and ethnic networks, or are better suited to particular conflict environments than regular forces (Kalyvas, 2006, Hughes and Tripodi, 2009, Lyall, 2010). This explanation captures real advantages but cannot account for why states empower proxies when direct action is feasible, more likely to be effective, and especially when the proxy deployment creates long-run political costs that might outweigh short-run military gains. A second explanation, *Plausible deniability*, emphasizes that proxies allow states to pursue violent objectives while obscuring their involvement; particularly to international audiences concerned with human rights or sovereignty norms (Salehyan, 2010, Carey et al., 2015). By delegating violence, states may escape sanctions, reputational costs, or intervention (Stanton, 2015, Mitchell et al., 2014). This account has merit for explaining the use of proxies in international conflicts, but it focuses almost exclusively on external audiences. It does not address the domestic political consequences of delegation, which are often more salient for regime survival than international opprobrium. Lastly, proxy reliance can be seen as evidence of state weakness or *capacity constraint*: governments delegate because they lack the capacity to project force directly and effectively (Besley and Persson, 2010, 2011). This explanation faces two difficulties. First, it risks circularity: if state weakness is inferred from the very behavior, proxy reliance, that it seeks to explain, the account has limited predictive content. The challenge is to specify a credible, context-specific, and importantly *independent* measure of state capacity and show that it predicts delegation. Second, the explanation cannot account for why strong states also delegate, despite possessing overwhelming conventional advantages. Taken together, these explanations fail to offer a credible reconciliation of the puzzle: why would a rational state, aware of blowback risk, deliberately create a potential political rival?

I propose a new explanation rooted in domestic political information. States face a fundamental tradeoff when they use violence: violence outcome reveals information about state capacity. Successful violence signals strength; failure signals weakness. Domestic audiences, voters, elites, and potential challengers observe these outcomes and update their beliefs about the incumbent's ability to maintain order and project power. These beliefs, in turn, shape political survival.

The informational consequences of violence depend critically on its mode of attack. When the state attacks directly, outcomes are diagnostic: success reflects capacity, failure reflects incapacity and limitations. When the state delegates to a proxy, outcomes are less informative about state capacity because they depend primarily on the proxy's effectiveness rather than the state's own strength. Delegation, in other words, obscures the mapping from violence outcomes to beliefs about incumbent quality.

This creates a strategic incentive for weak states to delegate. Delegation trades off current political survival against future political competition: the state avoids exposure today but creates a claimant for power tomorrow. The model in this paper formalizes this logic. An incumbent state has private information about its strength and must choose between direct or delegated violence against a target. Direct attacks succeed with a probability that depends on state strength; proxy attacks succeed with a probability that depends on proxy capability. Domestic audiences observe violence outcomes and update beliefs about the incumbent. The proxy's decision to challenge the state is itself endogenous. Following a successful operation, the proxy privately observes its cost of political activation, the organizational and reputational investments required to contest power, and decides whether to enter the political arena. This creates a genuine uncertainty about blowback: the state cannot perfectly predict whether delegation will trigger a political challenge. The outcome of this challenge depends on perceived state strength relative to the proxy's known capability.

I characterize Perfect Bayesian Equilibria of this game and identify qualitatively distinct configurations. The most substantively important is the *separating equilibrium*, in which powerful states attack directly and weak states delegate to the proxy. Under separation, violence mode perfectly reveals type: direct action signals strength, and delegation signals weakness. Voters correctly infer that direct attackers are powerful and that delegators are weak. Consequently, a weak state that delegates faces certain electoral defeat if the proxy activates politically. Yet, weak states accept this blowback risk because the alternative is worse: direct action with a high failure probability reveals weakness with certainty. The separating equilibrium exists when the cost differential between direct and delegated violent attack cost lies in an intermediate range and the weak state's direct attack success probability is sufficiently low; that is, when weak states are militarily disadvantaged enough that the informational benefit of delegation outweighs the blowback cost. This equilibrium captures the core "snakes in the backyard" logic: a weak state rationally empowers a domestic rival to avoid exposing its own weakness. The state's short-run interest in surviving the current period politically, conflicts with its long-run interest in maintaining a monopoly on political power.

Two pooling equilibria also exist. When the cost of direct action is sufficiently high relative to delegation, and the prior probability of state strength exceeds the proxy's capability, both types delegate, and no information is revealed; blowback does not arise in this equilibrium. When the cost of direct action is sufficiently low relative to delegation, both types attack directly, and delegation never occurs. The separating equilibrium arises for intermediate cost configurations, precisely when weak states have the most to gain from informational obscurity but face a binding incentive compatibility constraint. I also show that semi-pooling equilibria, where the powerful state plays a pure strategy and the weak state mixes, exist only at knife-edge parameter values and are non-generic.

Comparative statics on separating equilibrium provide some important insights. A more generously compensated proxy has a higher opportunity cost of political activation, reducing the likelihood of proxy's political entry. Moreover, as proxy inefficiency increases, the proxy becomes both militarily weaker, which reduces the value of delegation for the state, and shrinks the parameter region supporting the separating equilibrium. A central insight is that a very weak proxy, despite posing less electoral threat, is also of little military value; the relevant constraint on blowback is not the proxy's ability to win elections but the state's incentive to delegate in the first place.

These results have several implications. First, proxy deployment should correlate with incumbent weakness, conditional on violence occurring. Strong states have less to hide and face lower costs from revealing their type through direct action. This prediction distinguishes the theory from capacity-constraint explanations, which predict a monotonic relationship between state weakness and proxy reliance. Here, weak states delegate not because they lack the capacity to act directly, but because revealing their type is politically costlier than creating a rival. Second, blowback should be systematically associated with state-sponsored delegation rather than with the mere presence of armed groups. Groups cultivated by the state are particularly likely to challenge their patrons because the information structure of the delegation game creates conditions under which blowback is both possible and, in equilibrium, anticipated. Third, the terms of proxy contracts matter for political stability: proxies that receive larger shares of captured resources have weaker incentives to contest political power, suggesting a tradeoff between short-run resource extraction and long-run political stability.

This paper contributes to several literatures. *Political agency and accountability:* A central theme in political economy concerns how voters discipline politicians through reelection incentives and how politicians respond strategically to electoral pressure (Besley, 2006, Ashworth, 2012). This literature typically studies information revelation about types through policy choices or observable outcomes. I extend these ideas to settings where incumbents choose the *mode* of action precisely to manipulate the information voters receive. Delegation is not merely a principal-agent problem; it is a signaling device that allows incumbents to obscure their type to shape voters' beliefs about the state's strength.

My mechanism is related to models of strategic information transmission and media capture (Besley and Prat, 2006). In those settings, politicians manipulate information flows to influence voter beliefs. In my setting, politicians manipulate information by choosing actions whose outcomes are more or less diagnostic of their type. The innovation is that the informational properties of actions are themselves strategic objects.

Signaling and information in conflict: A growing literature studies how private information shapes conflict behavior (Fearon, 1995, Baliga and Sjöström, 2004, Powell, 2006, Jackson and Morelli, 2007, Fey and Ramsay, 2011). These papers typically consider bargaining between adversaries with private information about resolve or capability. I study a slightly different problem: how an incumbent's private information about its *own* strength shapes its choice of violence mode, with consequences for domestic rather than international audiences which eventually determine the survival of the incumbent regime itself.

Proxy warfare and non-state violence: An empirical literature documents patterns of state-proxy relationships, identifying correlates of militia formation, deployment, and behavior (Salehyan, 2010, Carey et al., 2015, Berman and Lake, 2019). This work provides rich descriptive evidence but lacks a unified theoretical framework for understanding when and why states delegate. I provide such a framework, generating predictions about the conditions under which delegation occurs and the factors that determine blowback risk.

Endogenous political competition: Standard models of electoral competition and political accountability take the set of political contestants as given. Egorov and Sonin (2011) study a dictator's choice of advisors who may become political threats, showing a competence-loyalty tradeoff: more capable agents are more useful but also more dangerous. My paper shares this "creating your own rival" logic but operates through a fundamentally different channel. In Egorov and Sonin (2011), the agent's threat arises from exogenous competence; in my model, the proxy becomes a rival *endogenously* through the information revealed by delegation. The closest antecedent to our mechanism is Powell (2013), who models a ruler's decision to arm agents who may later become rivals, showing that the inability to commit to future transfers creates inefficient violence. My paper differs in two respects. First, the state's decision to delegate is driven by *informational* considerations, hiding weakness from domestic audiences, rather than by a commitment problem over resource transfers. Second, blowback in my model operates through an electoral channel mediated by voter beliefs, whereas in Powell's framework rivalry arises from the coercive capacity that armed agents accumulate. More broadly, in my model, the proxy becomes a political rival *because of* delegation, the act of outsourcing violence creates the competitor. This mechanism of endogenous rival creation through information manipulation is, to my knowledge, novel in formal political economy.

The remainder of the paper proceeds as follows. Section 2 presents the model, defining players, information structure, timing, and payoffs. Section 3 characterizes equilibria, establishing existence conditions for separating, pooling, and semi-pooling configurations. Section 4 derives some important comparative statics and weak proxy paradox. Finally, section 5 concludes the paper.

2 The Model

2.1 Players and Environment

There are two strategic players and one passive actor:

- A **State** (S), the incumbent political power with private information about its strength.
- A **Proxy** (P), a non-state organization that can be enlisted for violence.
- A **Target** (T), a passive actor controlling resources that can be captured through violence.

In addition, a representative **Voter** (V) observes outcomes and decides whether to retain the incumbent. There is a fixed political-economic surplus (the "pie") that is normalized to size one. Initially, the target controls a share $\tau \in (0, 1)$ of the pie, while the state controls the remaining share $1-\tau$. The proxy controls no share initially. Following violence outcomes, voters observe public signals, update their beliefs, and decide whether to retain the incumbent or replace it. The game unfolds over three stages:

1. **Stage 0 (Nature):** Nature draws the state's strength $\theta \in \{\theta_L, \theta_H\}$ and the proxy's political activation cost c_A, both of which are private information to the respective players.
2. **Stage 1 (Attack Mode):** The state observes θ and chooses how to attack the target.
3. **Stage 2 (Violence):** The attack outcome is realized and publicly observed.
4. **Stage 3 (Political Contest):** If the proxy was used and succeeded, it may challenge the state; voters decide who governs.

I now describe each component of the model in detail.

2.2 Information Structure

2.2.1 State Strength

At the beginning of the game, nature draws the state's strength, which is a binary variable:

$$\theta \in \{\theta_L, \theta_H\}, \tag{2.1}$$

where θ_L denotes a *weak* state and θ_H denotes a *powerful* state. I normalize:

$$\theta_L = 0, \quad \theta_H = 1. \tag{2.2}$$

The prior probability that the state is powerful is:

$$\Pr(\theta = \theta_H) = p \in (0, 1). \tag{2.3}$$

The realization of θ is the **state's private information**. The proxy and voters share the common prior p.

2.2.2 Proxy Strength

While the state's strength is private, the proxy's strength is **publicly known** and denoted by:

$$\gamma = \frac{1}{1+\lambda}, \tag{2.4}$$

where $\lambda > 0$ is a parameter measuring proxy inefficiency. A higher λ means a weaker proxy.

Remark 1. *The proxy's strength γ lies in $(0,1)$. A fully competent proxy ($\lambda \to 0$) has $\gamma \to 1$; an incompetent proxy ($\lambda \to \infty$) has $\gamma \to 0$.*

2.2.3 Proxy's Private Information

The proxy has its own source of private information. If the proxy is used and the attack succeeds, the proxy may choose to *activate* politically (challenge the state). Political activation entails a cost c_A (drawn by nature at the beginning of the game) that is the **proxy's private information**:

$$c_A \sim H \text{ on } [\underline{c}, \bar{c}], \quad 0 < \underline{c} < \bar{c}, \tag{2.5}$$

where H is a continuously differentiable CDF with full support and density $h > 0$. I assume $\underline{c} < 1 - k\tau < \bar{c}$, ensuring $\pi \in (0,1)$

This two-sided private information, the state over its strength, the proxy over its activation cost, generates genuine strategic uncertainty on both sides of the delegation relationship.

2.3 Stage 1: Attack Mode

The state observes its type θ and chooses an attack mode:

$$m \in \{D, P\}, \tag{2.6}$$

where:

- $m = D$ (**Direct attack**): The state conducts a violent attack using its own forces, incurring a cost $c_D \geq 0$.

- $m = P$ (**Proxy attack**): The state delegates violence to the proxy, transferring resources $O = c_P$ to cover the proxy's operational cost $c_P > 0$.

An attack necessarily occurs; there is no option to refrain from violence. This reflects settings where the state has already committed to addressing the target. If the proxy is used and the attack succeeds, the state retains a share $(1 - k)$, and the proxy receives a share $k \in (0, 1)$ of the captured surplus τ, where k is a fixed and known parameter of the proxy contract.

2.4 Stage 2: Violence and Outcomes

The outcome of the attack depends on the mode chosen and, in the case of a direct attack, on the state's type.

2.4.1 Direct Attack

If the state chooses $m = D$, the attack succeeds with a probability that depends on the state strength:

$$p_D(\theta) = \begin{cases} 1 & \text{if } \theta = \theta_H, \\ \alpha & \text{if } \theta = \theta_L, \end{cases} \tag{2.7}$$

where $\alpha \in (0, 1)$. Thus, a powerful state always succeeds with a direct attack. A weak state, on the other hand, succeeds with probability $\alpha < 1$. If a direct attack succeeds, the share τ (initially held by the passive target) is transferred from the target to the state. If the attack fails, no redistribution occurs. Crucially, the dependence of success probability on θ makes direct attack outcomes *informative* about state strength.

2.4.2 Proxy Attack

If the state chooses $m = P$, the proxy attacks the target. The attack succeeds with probability:

$$p_P = \frac{c_P}{c_P + \lambda}. \tag{2.8}$$

Note that the proxy success probability is **independent of state strength**. It depends only on resources transferred (c_P) and proxy inefficiency (λ). This is crucial and implies that the proxy attack outcomes are uninformative about θ, essentially capturing the idea that the proxy delegation obfuscates the information about state type. If the proxy attack succeeds, the share τ is captured from the target. If the attack fails, no redistribution occurs.

Assumption 1 (Proxy Bounded Effectiveness)**.** *The proxy is strictly less effective than a powerful state:*

$$p_P = \frac{c_P}{c_P + \lambda} < 1 = p_D(\theta_H). \tag{2.9}$$

This is satisfied for any $\lambda > 0$.

2.5 Public Signals

After the violence outcome is realized, voters observe a public signal consisting of both the mode of attack and whether it succeeded:

$$\sigma \in \Sigma = \{(D, S), (D, F), (P, S), (P, F)\}, \tag{2.10}$$

where the first component indicates the attack mode and the second indicates success (S) or failure (F). Voters use this signal to update their beliefs about the state's type before the political contest.

2.6 Belief Updating

Let $s_\theta \in \{D, P\}$ denote the state's strategy (attack mode) as a function of type. Voters update their beliefs about the state's type using Bayes' rule. Define:

$$\Theta_D = \{\theta : s_\theta = D\}, \tag{2.11}$$
$$\Theta_P = \{\theta : s_\theta = P\}. \tag{2.12}$$

Definition 1 (Posterior Belief)**.** *The* ***posterior belief*** *after signal σ is:*

$$\mu(\sigma) = \Pr(\theta = \theta_H \mid \sigma). \tag{2.13}$$

Given the normalization $\theta_H = 1$, $\theta_L = 0$, the posterior expected strength equals the posterior belief:

$$\mathbb{E}[\theta \mid \sigma] = \mu(\sigma). \tag{2.14}$$

The following two lemmas establish the key informational asymmetry between attack modes.

Lemma 1 (Posteriors under Proxy Attack)**.** *If the state plays $m = P$, the attack outcome is uninformative about θ:*

$$\mu(P, S) = \mu(P, F) = \Pr(\theta_H \mid \theta \in \Theta_P). \tag{2.15}$$

Lemma 2 (Posteriors under Direct Attack)**.** *If both types play $m = D$, the posteriors satisfy:*

$$\mu(D, S) = \frac{p}{p + (1 - p)\alpha}, \tag{2.16}$$

$$\mu(D, F) = 0. \tag{2.17}$$

If only the weak type plays $m = D$: $\mu(D, S) = \mu(D, F) = 0$. If only the powerful type plays $m = D$: $\mu(D, S) = 1$, and $\mu(D, F)$ is off-path.

The proof for both lemmas is a straightforward application of Bayes' rule.

Corollary 1 (Informativeness)**.** *Direct attacks are informative: $\mu(D, S) \neq \mu(D, F)$ generically. Proxy attacks are uninformative: $\mu(P, S) = \mu(P, F)$.*

This asymmetry is the informational foundation of the model. Because proxy attacks do not update beliefs about state strength while direct attacks do, the choice of attack mode is itself a signal, which gives rise to the strategic considerations analyzed in the equilibrium section.

2.7 Stage 3: Political Contest

Stage 3 is reached after the signal σ is publicly observed. If the state attacked directly ($m = D$), no political contest occurs: the proxy was not involved in the violence, has acquired no organizational resources or public visibility from the operation, and therefore lacks both the legitimacy and the capacity to credibly challenge the incumbent. Similarly, if the proxy attack failed ($m = P,\ F$), the proxy lacks the resources or demonstrated competence and, more importantly, legitimacy to mount a viable political challenge. A political contest arises only when the proxy was used and succeeded: $\sigma = (P, S)$. In this case, the proxy has demonstrated military effectiveness, captured resources, and gained public salience; all of which provide the foundation for a credible political challenge.

2.7.1 Proxy's Activation Decision

If the proxy is used and succeeds, it must decide whether to accept the fixed share offered by the state or to fight the political contest. Formally, conditional on $\sigma = (P, S)$, the proxy privately observes c_A and chooses:

- **Inactivity** (I): Accept share $k \in (0, 1)$ of the captured surplus, yielding payoff $k\tau$.
- **Activation** (A): Challenge the state for control, incurring cost c_A.

2.7.2 Electoral Rule

If the proxy activates for the political contest, voters compare their perceived strength of the state to the proxy's known strength:

$$\text{Voters retain state} \iff \mu(\sigma) \geq \gamma. \tag{2.18}$$

- If $\mu(\sigma) \geq \gamma$: State wins, retains full pie.
- If $\mu(\sigma) < \gamma$: Proxy wins, gains full pie.

The microfoundation of this electoral rule follows from voters' payoff, which are presented in section 2.8.3.

Definition 2 (Political Prize)**.** *If the proxy wins the political contest, it replaces the state entirely as the governing authority and controls the full post-attack surplus:*

$$w = (1 - \tau) + \tau = 1. \tag{2.19}$$

This political prize definition essentially implies that the political victory confers complete governing authority. The proxy, having demonstrated both military effectiveness through the successful attack and political viability through the electoral contest, displaces the incumbent and inherits control over the entire political-economic surplus. This sharpens the state's tradeoff between delegation and political survival, especially when the political contest is decisive.[2]

2.7.3 Proxy's Activation Threshold

Since the proxy privately observes its political activation cost, even though its strength is publicly observable, the decision to fight a political contest remains unpredictable from the state's perspective. Even for the proxy, the political win may not be worthwhile if the political prize comes at a very high cost, c_A.

Lemma 3. *Proxy Activation Condition*
Let $\mu^* = \mu(P, S)$. The proxy activates iff:

$$c_A < \hat{c}(\mu^*) \equiv \begin{cases} w - k\tau = 1 - k\tau & \text{if } \mu^* \leq \gamma, \\ -k\tau & \text{if } \mu^* > \gamma. \end{cases} \tag{2.20}$$

Since $c_A > \underline{c} > 0$ and $-k\tau < 0$, the proxy never activates when $\mu^* > \gamma$.

Proof. If $\mu^* > \gamma$: activation leads to a certain political loss. Payoff is $0 - c_A < 0$, while inactivity gives $k\tau > 0$. If $\mu^* \leq \gamma$: activation leads to a certain political victory. Payoff is $w - c_A = 1 - c_A$. Proxy activates iff $1 - c_A > k\tau$, i.e., $c_A < 1 - k\tau$. □

Definition 3 (Activation Probability). *When $\mu^* < \gamma$, the probability that the proxy activates is:*

$$\pi = H(1 - k\tau). \tag{2.21}$$

When $\mu^* \geq \gamma$, $\pi = 0$.

The activation probability π is a key object in the analysis. It captures the likelihood of blowback conditional on proxy success and is determined by the distribution of the proxy's private activation cost.

[2] In practice, power transitions may be partial.

2.8 Payoffs

I now derive the expected payoffs for each player. These payoffs depend on the attack mode, violence outcome, and the realization of the political contest.

2.8.1 State's Payoff

Under direct attack ($m = D$): The attack costs c_D, state retains its initial share $(1-\tau)$, and captures τ with probability $p_D(\theta)$. There is no political contest after a direct attack. Hence:

$$U_S^D(\theta_H) = (1-\tau) - c_D + \tau = 1 - c_D, \tag{2.22}$$

$$U_S^D(\theta_L) = (1-\tau) - c_D + \alpha \cdot \tau = (1-\tau) - c_D + \alpha\tau. \tag{2.23}$$

Under proxy attack ($m = P$): The state pays c_P and the outcome depends on both the proxy's military success and, if successful, the subsequent political contest. Let $\mu^* = \mu(P, S)$ denote the posterior belief upon observing a successful proxy attack. There are two cases:

Case 1: $\mu^* > \gamma$ (no blowback). The proxy never activates (Lemma 3). With probability p_P the attack succeeds and the state retains share $(1-k)\tau$:

$$U_S^P = (1-\tau) - c_P + p_P(1-k)\tau. \tag{2.24}$$

Case 2: $\mu^* \le \gamma$ (blowback possible). With probability $(1-p_P)$, the attack fails and the state retains $(1-\tau) - c_P$. With probability $p_P(1-\pi)$, the attack succeeds, but the proxy stays inactive. With probability $p_P\pi$, the attack succeeds and the proxy activates and wins (since $\mu^* < \gamma$), leaving the state with $-c_P$:

$$\begin{aligned} U_S^P = {} & (1-p_P)[(1-\tau) - c_P] \\ & + p_P(1-\pi)[(1-\tau) - c_P + (1-k)\tau] \\ & + p_P \cdot \pi \cdot [-c_P]. \end{aligned} \tag{2.25}$$

Simplifying Case 2:

$$\begin{aligned} U_S^P &= (1-\tau) - c_P + p_P(1-\pi)(1-k)\tau - p_P\pi(1-\tau) \\ &= (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \end{aligned} \tag{2.26}$$

This expression is intuitive and it allows us to capture probable scenarios when the delegation is likely to destroy the value. The state's net expected gain from proxy success can be expressed as the difference between two competing forces: the expected resource capture when the proxy remains inactive, $(1-\pi)(1-k)\tau$, and the expected loss of the state's initial endowment when blowback occurs, $\pi(1-\tau)$. Delegation is beneficial only when the former exceeds the latter, which can be rearranged to:

$$\frac{1-k}{1} < \frac{\pi}{1-\pi} \cdot \frac{1-\tau}{\tau} \quad \Longleftrightarrow \quad \text{delegation destroys value for the state.} \tag{2.27}$$

The left-hand side captures the state's share of the spoils; the right-hand side is the product of the odds of blowback and the ratio of the state's initial endowment to the contested resources. Delegation is more likely to destroy value when blowback is probable (π high), the state's share of captured surplus is small (k high), and the state has more to lose from political displacement (τ low).

2.8.2 Proxy's Payoff

If $m = D$: $U_P = 0$ (the proxy is not involved). If $m = P$ and the attack fails: $U_P = O - c_P = c_P - c_P = 0$ (the transfer $O = c_P$ covers exactly the operational cost). If $m = P$ and attack succeeds:

- Inactive: $U_P = k\tau$.
- Activates and wins: $U_P = w - c_A = 1 - c_A$.
- Activates and loses: $U_P = -c_A$.

2.8.3 Voter's Payoff

Voters prefer to be governed by a stronger entity. Formally, the representative voter's payoff equals the strength of whoever holds power:

$$U_V = \begin{cases} \theta & \text{if the state governs,} \\ \gamma & \text{if the proxy governs.} \end{cases} \tag{2.28}$$

Given posterior belief $\mu(\sigma) = \Pr(\theta = \theta_H \mid \sigma)$, the voter's expected payoff from retaining the state is $\mathbb{E}[\theta \mid \sigma] = \mu(\sigma)$, while the expected payoff from electing the proxy is γ. Hence, voters optimally retain the state if and only if $\mu(\sigma) \geq \gamma$. This microfounds the electoral rule specified in Section 2.7.2.

2.9 Strategies and Equilibrium

Having described the model environment, I now define the solution concept. The equilibrium concept is *Perfect Bayesian Equilibrium* (PBE). Formally, strategy and the equilibrium are defined as follows.

Definition 4 (Strategy)**.** *A **strategy for the state** is a mapping $s : \{\theta_L, \theta_H\} \to \{D, P\}$. A **strategy for the proxy** is a mapping $\beta : \Sigma \times [\underline{c}, \bar{c}] \to \{I, A\}$.*

I focus on pure strategies; mixed strategies are considered in Section 3.5.

Definition 5 (Perfect Bayesian Equilibrium)**.** *A **Perfect Bayesian Equilibrium (PBE)** consists of strategies (α^*, β^*) and beliefs μ^* such that:*

1. *For each θ, $\alpha^*(\theta)$ maximizes the state's expected payoff.*
2. *For each σ and c_A, $\beta^*(\sigma, c_A)$ maximizes the proxy's expected payoff.*
3. *$\mu^*(\sigma)$ is derived via Bayes' rule for on-path signals. For off-path signals, beliefs are unrestricted (any $\mu \in [0, 1]$ is admissible)*

3 Equilibrium Analysis

I now characterize the Perfect Bayesian Equilibria of this game. With binary types, there are four candidate pure-strategy profiles for the state. I also examine whether semi-pooling equilibria, in which one type mixes, can arise. Before analyzing each equilibrium candidate, Lemma 4 establishes the proxy's best response, which applies across all equilibrium configurations.

3.1 Proxy's Best Response

The proxy's equilibrium strategy follows directly from the activation condition derived in Lemma 3. Since the proxy moves after observing the public signal and its private cost, its optimal behavior can be stated in closed form.

Lemma 4 (Proxy's Equilibrium Strategy)**.** *In any equilibrium, the proxy's strategy is:*

$$\beta^*(\sigma, c_A) = \begin{cases} A & \text{if } \sigma = (P, S),\ \mu(P, S) < \gamma,\ c_A < 1 - k\tau, \\ I & \text{otherwise.} \end{cases} \tag{3.1}$$

Proof. Immediate from Lemma 3. □

The lemma establishes that the proxy's behavior is fully determined by the posterior belief $\mu(P, S)$ and the realized cost c_A. The state's equilibrium problem therefore reduces to choosing an attack mode, taking as given that delegation creates blowback risk whenever $\mu(P, S) \leq \gamma$.

I now analyze each candidate equilibrium in turn, beginning with the separating equilibrium that is the paper's central object of interest.

3.2 Separating Equilibrium: $\theta_H \to D$, $\theta_L \to P$

The separating equilibrium is the most substantively important configuration. The powerful state signals strength through direct action, while the weak state hides behind the proxy, obfuscating its type. This is the equilibrium that generates the "snakes in the backyard" dynamic: weak states rationally create political rivals by delegating violence.

Proposition 1 (Separating Equilibrium Characterization)**.** *Consider the strategy profile* $(s(\theta_H) = D, s(\theta_L) = P)$*. This is a PBE if and only if the following conditions are simultaneously satisfied:*

$$c_D - c_P \leq \tau - p_P(1 - k)\tau \tag{3.2}$$

$$c_D - c_P \geq \alpha\tau - p_P[(1 - \pi)(1 - k)\tau - \pi(1 - \tau)] \tag{3.3}$$

$$\alpha \leq 1 - p_P\pi(1 - k) - p_P\pi\frac{1 - \tau}{\tau} \tag{3.4}$$

where $\pi = H(1 - k\tau)$ *is the blowback probability.*

Proof. In Appendix A.2 □

Condition (3.2) ensures that the powerful state prefers direct attack over delegation. Even in the best case for deviation, where the proxy does not activate, the powerful state must find direct attack more attractive. Since θ_H succeeds with certainty under D and captures

τ, while delegation yields at most $(1-k)\tau$ in expectation, this holds when $c_D - c_P$ is not too large. Condition (3.3) ensures that the weak state prefers delegation over direct attack despite the blowback risk. The weak state's direct attack payoff depends on α; when α is low, the weak state has little to gain from direct action and much to lose from revealing its type. Together, the two conditions require $c_D - c_P$ to lie in an intermediate range, which is non-empty when α is sufficiently small. Intuitively, the separating equilibrium exists precisely when weak states are militarily disadvantaged enough that the informational benefit of delegation outweighs the blowback cost.

3.3 Pooling on Proxy: $\theta_H \to P$, $\theta_L \to P$

When direct attack is sufficiently costly, both types may prefer to delegate. In this case, the posterior equals the prior, and the proxy's activation decision depends on how the prior compares to proxy strength.

Proposition 2 (Pooling on Proxy Equilibrium)**.** *Consider the strategy profile* $(s(\theta_H) = P, s(\theta_L) = P)$*. Under this profile:*

$$\mu(P, S) = \mu(P, F) = p. \tag{3.5}$$

This is a PBE if and only if:

1. $p > \gamma$ *(no blowback in equilibrium)*
2. $c_D - c_P > \tau - p_P(1-k)\tau$ *i.e.,* c_D *is sufficiently high favoring delegation*

Proof. In Appendix A.3 □

When both types delegate, the proxy attack conveys no information about θ, so the posterior remains at the prior p. If $p > \gamma$, voters perceive the state as stronger than the proxy, and the proxy never activates regardless of c_A. The equilibrium is therefore blowback-free: delegation is "safe" for both types. However, this requires two conditions. First, the state must be perceived as strong enough *a priori* ($p > \gamma$). Second, direct attack must be costly enough that even the powerful state, which would succeed with certainty, prefers the expected payoff from delegation. Notice that condition 2 is the strict reverse of condition (3.2): the parameter region supporting pooling on proxy is exactly the complement of the region where the powerful state is willing to separate.

3.4 Pooling on Direct: $\theta_H \to D$, $\theta_L \to D$

At the other extreme, when direct attack is cheap, both types may prefer to act directly, forgoing delegation entirely.

Proposition 3 (Pooling on Direct Equilibrium)**.** *Consider* $(s(\theta_H) = D, s(\theta_L) = D)$*. Under this profile:*

$$\mu(D, S) = \frac{p}{p + (1-p)\alpha}, \tag{3.6}$$

$$\mu(D, F) = 0. \tag{3.7}$$

This is a PBE if and only if :

$$c_D - c_P \leq \alpha\tau - p_P(1-k)\tau \tag{3.8}$$

Proof. In Appendix A.4 □

Under pooling on direct, both types attack using their own forces. The posterior $\mu(D, S)$ is interior, success raises the probability that the state is powerful, but does not fully reveal type since both types may succeed. The posterior $\mu(D, F) = 0$ because only the weak type can fail. The no-deviation condition requires that even the weak state, which benefits most from delegation's informational obscurity, finds direct attack preferable. This holds when $c_D - c_P$ is sufficiently low relative to the weak state's residual military advantage $\alpha\tau$. Note that condition (3.8) implies condition (3.2), so pooling on direct and pooling on proxy cannot coexist for the same parameter values.

3.5 Semi-Pooling Equilibrium

I now analyze equilibria in which one type plays a pure strategy while the other mixes. The most natural candidate is a *semi-pooling* equilibrium where the powerful state always attacks directly, while the weak state randomizes between direct and proxy attack. In this configuration, types partially pool on direct attack.

Proposition 4 (Semi-Pooling Equilibrium)**.** *Consider a strategy profile where θ_H plays D with probability 1, and θ_L plays D with probability $q \in (0,1)$ and P with probability $1-q$.*

Under this profile:

$$\mu(D,S) = \frac{p}{p+(1-p)q\alpha}, \tag{3.9}$$

$$\mu(D,F) = 0, \tag{3.10}$$

$$\mu(P,S) = \mu(P,F) = 0. \tag{3.11}$$

This constitutes a PBE if and only if:

$$c_D - c_P = \alpha\tau - p_P\left[(1-\pi)(1-k)\tau - \pi(1-\tau)\right], \tag{3.12}$$

where $\pi = H(1-k\tau)$ *is the activation probability.*

Proof. In Appendix A.5 □

Remark 2 (Non-Genericity)**.** *The semi-pooling equilibrium exists if and only if condition* (3.12) *holds exactly. This is a knife-edge condition with measure zero in the parameter space* $(c_D, c_P, \alpha, \tau, k, \lambda)$*. When the condition fails:*

- *If* $c_D - c_P < \alpha\tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)]$*: the weak state strictly prefers* D*, leading to pooling on direct attack.*
- *If* $c_D - c_P > \alpha\tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)]$*: the weak state strictly prefers* P*, leading to the separating equilibrium (Proposition 1).*

Remark 3 (Indeterminacy of the Mixing Probability)**.** *When condition* (3.12) *holds, the mixing probability* q *is indeterminate: any* $q \in (0,1)$ *is consistent with equilibrium. This arises because the state's payoff from direct attack does not depend on the posterior belief* $\mu(D,S)$*. There is no political contest following a direct attack, so the belief has no payoff consequence. Consequently, the mixing probability affects only the posterior* $\mu(D,S)$*, which is payoff-irrelevant, and cannot be pinned down by incentive compatibility.*

For completeness, I also analyze the fourth candidate equilibrium: powerful states delegate while weak states attack directly in Appendix A.1. The proposition 8 in Appendix A.1 formally proves the non-existence of this type of equilibrium.

The failure of the reverse separating equilibrium reflects a fundamental asymmetry in the model. Under this profile, only θ_H delegates, so $\mu(P,S) = 1 > \gamma$, and the proxy never activates. Delegation is therefore "safe." But this very safety creates a fatal incentive problem: if delegation carries no blowback risk and is attributed to the powerful type, the weak state has a strong incentive to mimic by also delegating, thereby being mistaken for θ_H. For the reverse separating equilibrium to hold, θ_L must prefer direct attack over safe delegation, which requires c_D to be low. Simultaneously, θ_H must prefer delegation over direct attack, which requires c_D to be high. These conditions are mutually incompatible, formally, they require $\alpha \geq 1$, contradicting our assumption that the weak state is militarily disadvantaged. A weak state cannot be induced to choose the risky, revealing action when a safe, pooling option is available.

4 Comparative Statics

The equilibrium analysis identifies three parameter regions, pooling on direct, separating, and pooling on proxy, governed by the cost differential $\Delta = c_D - c_P$ relative to thresholds that depend on α, τ, k, λ, and the distribution H. I now examine how changes in key parameters affect both the blowback probability within the separating equilibrium and the equilibrium regions themselves. Throughout this section, I focus on the separating equilibrium ($\theta_H \to D$, $\theta_L \to P$) unless stated otherwise.

4.1 Blowback Probability

The central endogenous object in the separating equilibrium is the blowback probability $\pi = H(1 - k\tau)$. Recall that the proxy activates if and only if $c_A < 1 - k\tau$: the proxy challenges the state when its private cost of political entry is below the net gain from winning power ($w = 1$) minus what it would earn by staying inactive ($k\tau$).

Proposition 5 (Comparative Statics on Blowback Probability)**.** *In the separating equilibrium, the blowback probability $\pi = H(1 - k\tau)$ satisfies:*

1. *The blowback probability (π) is decreasing in the proxy's offered share of captured surplus k.*
2. *The blowback probability (π) is decreasing in the target's share of the pie (τ).*

Proof. The activation threshold is $\hat{c} = 1 - k\tau$. Since H is continuously differentiable with density $h > 0$:

$$\frac{\partial \pi}{\partial k} = h(1 - k\tau) \cdot (-\tau) < 0, \tag{4.1}$$

$$\frac{\partial \pi}{\partial \tau} = h(1 - k\tau) \cdot (-k) < 0. \tag{4.2}$$

□

The intuition for both results operates through the proxy's opportunity cost of activation. A higher k means the proxy receives a larger share of captured resources if it remains inactive; this raises the opportunity cost of challenging the state, reducing blowback. A higher τ means more is at stake in the violence stage; since the proxy's inactive payoff is $k\tau$, a larger τ increases what the proxy certainly foregoes by activating. This again makes the outside option of staying inactive more attractive relative to the uncertain payoff from political entry. These results have a natural policy interpretation. States can reduce blowback risk by offering proxies more generous terms, a larger share k of the spoils. However, this comes at the cost of the state's own surplus from successful proxy operations: the state retains $(1 - k)\tau$ from a successful attack. There is thus a straightforward tradeoff between short-run resource extraction and long-run political stability.

4.2 The Role of Weak State's Military Capability

The parameter α, the weak state's probability of success under direct attack, plays a central role in determining whether the separating equilibrium exists.

Proposition 6 (Effect of Weak State's Capability). *1. The lower threshold for the separating equilibrium, $\underline{\Delta} = \alpha\tau - p_P[(1 - \pi)(1 - k)\tau - \pi(1 - \tau)]$, is strictly increasing in α.*

2. The upper threshold $\overline{\Delta} = \tau - p_P(1 - k)\tau$ is independent of α.

3. The width of the separating equilibrium region, $\overline{\Delta} - \underline{\Delta}$, is strictly decreasing in α: higher α shrinks the range of Δ supporting separation.

4. As $\alpha \to 0$, the separating equilibrium region is maximally wide. Moreover, the region is non-empty at $\alpha = 0$ if and only if $\tau > p_P\pi(1 - k\tau)$.

Proof. (1) $\frac{\partial \underline{\Delta}}{\partial \alpha} = \tau > 0$. (2) $\overline{\Delta}$ does not contain α. (3) Follows from (1) and (2). (4) At $\alpha = 0$, $\underline{\Delta} = -p_P[(1 - \pi)(1 - k)\tau - \pi(1 - \tau)]$. □

This result is central to the paper's narrative. The separating equilibrium, the configuration generating endogenous blowback, arises precisely when weak states are militarily disadvantaged. A weak state with very low α gains little from direct violence (its expected capture is only $\alpha\tau$) and risks exposing its type through failure. Delegation offers a superior lottery: the proxy succeeds with probability p_P regardless of θ, and although blowback may follow, the weak state at least avoids the certainty of being revealed as weak. As α increases, the weak state's direct attack becomes more attractive, eventually making delegation unnecessary.

4.3 Proxy Inefficiency

The parameter λ simultaneously affects the proxy's military effectiveness (p_P), its political strength (γ), and the equilibrium structure.

Proposition 7 (Effect of Proxy Inefficiency)**.** *As proxy inefficiency λ increases:*

1. *Proxy attack success probability $p_P = \frac{c_P}{c_P+\lambda}$ decreases.*
2. *Proxy political strength $\gamma = \frac{1}{1+\lambda}$ decreases.*
3. *The range of c_D supporting the separating equilibrium shrinks, and pooling on D becomes more likely.*

Proof. (1) and (2) follow directly from the definitions. For (3), note that as $p_P \to 0$, the benefit of proxy attack vanishes: the expected surplus capture $p_P\tau \to 0$. This makes the proxy attack less attractive for θ_L, pushing toward pooling on D. Formally, as $p_P \to 0$:

$$U_S^P(\theta_L; \mu = 0) \to (1 - \tau) - c_P, \tag{4.3}$$

while $U_S^D(\theta_L) = (1 - \tau) - c_D + \alpha\tau$. For low α, this becomes $U_S^D(\theta_L) \approx (1 - \tau) - c_D$. The comparison depends on $c_D - c_P$, but the benefit of delegation disappears. □

Remark 4 (Weak Proxy Paradox)**.** *A very weak proxy (λ large) might seem to pose less blowback risk since $\gamma \to 0$ makes the proxy less electorally threatening. However, in the separating equilibrium, the proxy's electoral threat is invariant to its strength, the posterior $\mu(P, S) = 0$ always loses to any $\gamma > 0$, regardless of how small γ is. The relevant effect of high λ is therefore not on the electoral outcome but on the* value *of delegation: a weak* *proxy is militarily ineffective ($p_P \to 0$), reducing the state's incentive to delegate in the first place. The paradox, therefore, is that the relevant margin for blowback risk is not the proxy's* political *weakness but its* military *competence. States should worry less about how strong a proxy might become politically and more about whether the proxy is effective enough to justify delegation in the first place.*

5 Conclusion

This paper develops a formal model of why states empower domestic violent actors despite the risk of future political competition. The key mechanism is informational: direct violence reveals state capacity to domestic audiences, while delegated violence obscures it. When voters condition political support on perceived strength, weak states face incentives to delegate even at the cost of creating political rivals. The separating equilibrium, in which powerful states attack directly and weak states delegate, captures the self-undermining logic of coercive delegation: empowering proxies hides weakness today but creates competitors tomorrow. The model also characterizes pooling equilibria in which both types choose the same mode, and shows that semi-pooling equilibria are non-generic and that the reverse separating configuration is ruled out entirely.

The analysis yields several empirically relevant insights. Proxy delegation should be more prevalent among states for which revealing weakness is politically costly, not merely among states that lack military capacity. Blowback is endogenous to the delegation decision and can be mitigated through more generous resource-sharing arrangements with the proxy, though at the cost of the state's own surplus. The model also highlights a weak proxy paradox: less capable proxies pose no less political risk conditional on activation, since voters know that delegators are weak, but reduce the military value of delegation, eventually making it unattractive. These predictions distinguish the theory from existing explanations based on efficiency, deniability, or capacity constraints, and are in principle testable using data on state-proxy relationships and subsequent political transitions.

Several extensions merit future investigation. Endogenizing the sharing rule k would allow analysis of the state's optimal contract design, formalizing the tradeoff between resource extraction and blowback management. Introducing heterogeneity among voters or allowing the proxy to strategically choose its level of political visibility would enrich the political contest stage. Finally, dynamic extension in which the proxy accumulates political capital over repeated interactions could capture the escalation of rivalry and reputational concerns in political competition observed in empirical settings and states like Lebanon or Palestine.

A Appendix

A.1 The Reverse Separating Equilibrium: $\theta_H \to P$, $\theta_L \to D$

Proposition 8 (Non-Existence of Reverse Separating Equilibrium)**.** *The strategy profile* $(s(\theta_H) = P, s(\theta_L) = D)$ *cannot constitute a PBE for any parameter values satisfying our maintained assumptions.*

Proof. **Beliefs:** Under the proposed strategy:

- Only θ_H plays P. Hence $\mu(P, S) = \mu(P, F) = 1$.
- Only θ_L plays D. Hence $\mu(D, S) = \mu(D, F) = 0$.

Blowback: Since $\mu(P, S) = 1 > \gamma$ for any $\gamma < 1$, the proxy never activates. The powerful state faces zero blowback risk. **On-path payoffs:** For θ_H (playing P with no blowback):

$$U_S^P(\theta_H) = (1 - \tau) - c_P + p_P(1 - k)\tau. \tag{A.1}$$

For θ_L (playing D):

$$U_S^D(\theta_L) = (1 - \tau) - c_D + \alpha\tau. \tag{A.2}$$

θ_L's deviation incentive: If θ_L deviates to P, voters observe a proxy signal and believe $\mu = 1$ (thinking the state is θ_H). Since $\mu = 1 > \gamma$, no blowback occurs. The deviation payoff is:

$$U_S^P(\theta_L; \mu = 1) = (1 - \tau) - c_P + p_P(1 - k)\tau. \tag{A.3}$$

For θ_L not to deviate:

$$(1 - \tau) - c_D + \alpha\tau \geq (1 - \tau) - c_P + p_P(1 - k)\tau \tag{A.4}$$

$$c_D \leq c_P + [\alpha - p_P(1 - k)]\tau. \tag{A.5}$$

θ_H's deviation incentive: If θ_H deviates to D, it succeeds with probability 1 and captures τ. Although voters would believe $\mu(D,S)=0$, this has no payoff consequence since there is no political contest following direct attack. The deviation payoff is:

$$U_S^D(\theta_H) = 1 - c_D. \tag{A.6}$$

For θ_H not to deviate:

$$(1-\tau) - c_P + p_P(1-k)\tau \geq 1 - c_D \tag{A.7}$$

$$c_D \geq c_P + \tau[1 - p_P(1-k)]. \tag{A.8}$$

Compatibility of conditions: For the equilibrium to exist, both (A.5) and (A.8) must hold simultaneously. This requires:

$$c_P + \tau[1 - p_P(1-k)] \leq c_D \leq c_P + [\alpha - p_P(1-k)]\tau. \tag{A.9}$$

The interval is non-empty only if:

$$\tau[1 - p_P(1-k)] \leq [\alpha - p_P(1-k)]\tau. \tag{A.10}$$

Dividing by $\tau > 0$:

$$1 - p_P(1-k) \leq \alpha - p_P(1-k), \tag{A.11}$$

which simplifies to $1 \leq \alpha$. Since $\alpha < 1$ and $0 < p_P < 1$ by Assumption 1, this condition can never be satisfied. The interval of feasible c_D values is empty. □

A.2 Proof of Proposition 1

Proof. **Beliefs:** Note that under the separating strategy:

- Only θ_H plays D, and θ_H always succeeds. So $\mu(D,S)=1$.
- The signal (D,F) is off-path (impossible under this strategy).
- Only θ_L plays P. So $\mu(P,S)=\mu(P,F)=0$.

Blowback: Since $\mu(P,S) = 0 < \gamma$ (as $\gamma > 0$), the proxy will activate if $c_A < 1 - k\tau$. The activation probability is $\pi = H(1-k\tau) \in (0,1)$.
Payoffs: Since the powerful state wins with certainty, the powerful state's payoff under D:

$$U_S^D(\theta_H) = 1 - c_D. \tag{A.12}$$

If θ_H deviates to P, voters observe a proxy signal and hold equilibrium belief $\mu(P,S) = 0$. The deviation payoff, accounting for blowback risk, is:

$$U_S^P(\theta_H; \mu^* = 0) = (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)]. \tag{A.13}$$

Since the most favorable scenario for the deviant is when the proxy does not activate (yielding $(1-\tau) - c_P + p_P(1-k)\tau$), a sufficient condition for θ_H not to deviate is:

$$\begin{aligned} 1 - c_D &\geq (1-\tau) - c_P + p_P(1-k)\tau \\ c_D - c_P &\leq \tau - p_P(1-k)\tau \end{aligned}$$

Condition (3.2) ensures θ_H doesn't deviate even in the best case for deviation. Since the actual deviation payoff (accounting for blowback) is strictly lower than the best-case payoff, condition (3.2) is sufficient. Moreover, it is the tightest condition arising from θ_H's incentives: the weak state's no-deviation condition (3.3) provides the binding constraint from below, while (3.2) provides the binding constraint from above.
Now, the weak state's payoff under P with $\mu(P,S) = 0$:

$$U_S^P(\theta_L; \mu^* = 0) = (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)]. \tag{A.14}$$

The weak state's payoff if it deviates to D: It succeeds with probability α and captures τ:

$$U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau. \tag{A.15}$$

If θ_L deviates to D and succeeds, voters observe (D,S) and update $\mu(D,S) = 1$ (since only θ_H is supposed to play D). To support the equilibrium, we can assign off-path belief $\mu(D,F) = 0$ (or any belief). The weak state faces a failure probability $1-\alpha$. For θ_L not to deviate:

$$U_S^P(\theta_L; \mu^* = 0) \geq U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau. \tag{A.16}$$

This is condition (3.3) upon the simplification:

$$\begin{aligned}(1-\tau)-c_D+\alpha\tau &\leq (1-\tau)-c_P+p_P[(1-\pi)(1-k)\tau-\pi(1-\tau)]\\ c_D-c_P &\geq \alpha\tau-p_P[(1-\pi)(1-k)\tau-\pi(1-\tau)]\end{aligned}$$

Thus, for separating equilibrium we need c_D-c_P to lie in the interval $\big(\ \alpha\tau-p_P[(1-\pi)(1-k)\tau-\pi(1-\tau)],\ \ \tau-p_P(1-k)\tau\ \big)$. This interval will be nonempty if $\alpha\leq 1-p_P(1-k)+p_P(1-\pi)(1-k)-p_P\pi\frac{1-\tau}{\tau}$. □

A.3 Proof of Proposition 2

Proof. **Beliefs and blowback:** If both types play P, then $\mu(P,S)=\mu(P,F)=p$ (prior). If $p>\gamma$: Since $\mu(P,S)=p>\gamma$, the proxy never activates. The payoff for both types is:

$$U_S^P=(1-\tau)-c_P+p_P(1-k)\tau. \tag{A.17}$$

Deviation incentives: Consider θ_H deviating to D. The signal (D,S) is off-path. Let off-path belief be $\mu(D,S)=\tilde{\mu}$. θ_H's deviation payoff:

$$U_S^D(\theta_H)=1-c_D. \tag{A.18}$$

For θ_H not to deviate:

$$(1-\tau)-c_P+p_P(1-k)\tau \geq 1-c_D \tag{A.19}$$

$$c_D-c_P \geq \tau-p_P(1-k)\tau \tag{A.20}$$

Similarly, θ_L doesn't deviate if:

$$(1-\tau)-c_P+p_P(1-k)\tau \geq (1-\tau)-c_D+\alpha\tau \tag{A.21}$$

$$c_D-c_P \geq \alpha\tau-p_P(1-k)\tau \tag{A.22}$$

Hence, from equation (A.20) and (A.22) we need $c_D-c_P>\max\{\tau(1-p_P(1-k)),\ \ \alpha\tau-p_P(1-k)\tau\}$, that is condition 2: $c_D-c_P>\tau-p_P(1-k)\tau$.
Note that if $p\leq\gamma$: the proxy activates with probability $\pi>0$ on the equilibrium path. The state's equilibrium payoff becomes $(1-\tau)-c_P+p_P[(1-\pi)(1-k)\tau-\pi(1-\tau)]$, which is strictly lower than the no-blowback payoff. While this does not immediately rule out

the equilibrium, it alters the deviation conditions and, more importantly, the blowback-free pooling characterization no longer applies. The case $p \leq \gamma$ with both types delegating is substantially different and is not analyzed here. □

A.4 Proof of Proposition 3

Proof. Note that $m = P$ is off path.

Payoffs on path:

θ_H: $U_S^D(\theta_H) = 1 - c_D$.

θ_L: $U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau$.

Deviation to P: If a type deviates to P, signals (P, S) and (P, F) are off-path. Let off-path belief be $\mu(P, S) = \tilde{\mu}$. If $\tilde{\mu} > \gamma$: no blowback. Deviation payoff is:

$$U_S^P(\text{no blowback}) = (1-\tau) - c_P + p_P(1-k)\tau. \tag{A.23}$$

For θ_H not to deviate:

$$1 - c_D \geq (1-\tau) - c_P + p_P(1-k)\tau \tag{A.24}$$

$$c_D - c_P \leq \tau - p_P(1-k)\tau \tag{A.25}$$

For θ_L not to deviate:

$$(1-\tau) - c_D + \alpha\tau \geq (1-\tau) - c_P + p_P(1-k)\tau, \tag{A.26}$$

$$c_D - c_P \leq \alpha\tau - p_P(1-k)\tau \tag{A.27}$$

Combining equation (A.25) and (A.27), we get a sufficient condition 3.8 for pooling equilibrium.

If $\tilde{\mu} \leq \gamma$: blowback occurs with probability π. Deviation becomes less attractive. Since we verify the no-deviation condition under most favorable beliefs, the condition holds for any off-path beliefs $\tilde{\mu} \leq \gamma$ that would trigger blowback and further reduce the deviation payoff. □

A.5 Proof of Proposition 4

Proof. **Beliefs:** Under the proposed strategy:

- Signal (D, S): Both types may play D. The powerful type plays D and succeeds with probability 1; the weak type plays D with probability q and succeeds with probability α. By Bayes' rule:

$$\mu(D, S) = \frac{p \cdot 1}{p \cdot 1 + (1-p) \cdot q \cdot \alpha} = \frac{p}{p + (1-p)q\alpha}. \tag{A.28}$$

- Signal (D, F): Only θ_L can fail, so $\mu(D, F) = 0$.

- Signals (P, S) and (P, F): Only θ_L plays P, so $\mu(P, S) = \mu(P, F) = 0$.

Blowback: Since $\mu(P, S) = 0 < \gamma$, the proxy will challenge the state if it activates. By Lemma 3, the proxy activates whenever $c_A < 1 - k\tau$, so the activation probability is $\pi = H(1 - k\tau)$.

Payoffs:

Strong state (θ_H): On-path payoff with direct attack is $U_S^D(\theta_H) = 1 - c_D$. If the strong state deviates to P, voters observe a proxy signal and believe $\mu(P, S) = 0$ (since only weak types are supposed to play P). The state therefore faces full blowback risk, and the deviation payoff is:

$$U_S^P(\theta_H) = (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{A.29}$$

To prevent this deviation, we need

$$1 - c_D \geq (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{A.30}$$

$$c_D - c_P \leq \tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{A.31}$$

Weak state (θ_L): On-path payoff under direct attack:

$$U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau \tag{A.32}$$

Weak state's payoff under proxy attack with $\mu(P, S) = 0$:

$$U_S^P(\theta_L) = (1-\tau) - c_P + p_P(1-\pi)(1-k)\tau - p_P\pi(1-\tau) \tag{A.33}$$

Indifference condition for mixing: For θ_L to mix with interior probability $q \in (0, 1)$, we require indifference:

$$U_S^D(\theta_L) = U_S^P(\theta_L) \tag{A.34}$$

$$(1-\tau) - c_D + \alpha\tau = (1-\tau) - c_P + p_P(1-k)(1-\pi)\tau - p_P\pi(1-\tau) \tag{A.35}$$

$$c_D - c_P = \alpha\tau - p_P\left[(1-k)(1-\pi)\tau - \pi(1-\tau)\right]. \tag{A.36}$$

Verify that θ_H strictly prefers D: Since the proxy attack payoff is independent of the state's type (p_P does not depend on θ), we have $U_S^P(\theta_H) = U_S^P(\theta_L)$. Using the indifference condition (A.36):

$$U_S^P(\theta_H) = U_S^P(\theta_L) = U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau. \tag{A.37}$$

We need $U_S^D(\theta_H) > U_S^P(\theta_H)$:

$$1 - c_D > (1-\tau) - c_D + \alpha\tau \tag{A.38}$$

$$1 > 1 - \tau(1-\alpha). \tag{A.39}$$

Since $\tau > 0$ and $\alpha < 1$, we have $\tau(1-\alpha) > 0$, so $1 - \tau(1-\alpha) < 1$. The inequality holds strictly, confirming that θ_H strictly prefers D. Note that this also implies condition (A.31) is automatically satisfied whenever the indifference condition (A.36) holds. □

References

Acemoglu, D., Robinson, J. A., and Santos, R. J. (2013). The monopoly of violence: Evidence from colombia. *Journal of the European Economic Association*, 11(suppl_1):5–44.

Ahram, A. I. (2011). *Proxy warriors: the rise and fall of state-sponsored militias*. Stanford University Press.

Ashworth, S. (2012). Electoral accountability: Recent theoretical and empirical work. *Annual Review of Political Science*, 15(1):183–201.

Baliga, S. and Sjöström, T. (2004). Arms races and negotiations. *The Review of Economic Studies*, 71(2):351–369.

Berman, E. and Lake, D. A. (2019). *Proxy wars: Suppressing violence through local agents*. Cornell University Press.

Besley, T. (2006). *Principled agents?: The political economy of good government*. Oxford University Press.

Besley, T. and Persson, T. (2010). State capacity, conflict, and development. *Econometrica*, 78(1):1–34.

Besley, T. and Persson, T. (2011). Pillars of prosperity: The political economics of development clusters. In *Pillars of Prosperity*. Princeton University Press.

Besley, T. and Prat, A. (2006). Handcuffs for the grabbing hand? media capture and government accountability. *American economic review*, 96(3):720–736.

Carey, S. C., Colaresi, M. P., and Mitchell, N. J. (2015). Governments, informal links to militias, and accountability. *Journal of Conflict Resolution*, 59(5):850–876.

Carey, S. C., Mitchell, N. J., and Lowe, W. (2013). States, the security sector, and the monopoly of violence: A new database on pro-government militias. *Journal of Peace Research*, 50(2):249–258.

Egorov, G. and Sonin, K. (2011). Dictators and their viziers: Endogenizing the loyalty–competence trade-off. *Journal of the European Economic Association*, 9(5):903–930.

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

Fey, M. and Ramsay, K. W. (2011). Uncertainty and incentives in crisis bargaining: Game-free analysis of international conflict. *American Journal of Political Science*, 55(1):149–169.

Hughes, G. and Tripodi, C. (2009). Anatomy of a surrogate: historical precedents and implications for contemporary counter-insurgency and counter-terrorism. *Small Wars & Insurgencies*, 20(1):1–35.

Jackson, M. O. and Morelli, M. (2007). Political bias and war. *American Economic Review*, 97(4):1353–1373.

Kalyvas, S. N. (2006). *The logic of violence in civil war*. Cambridge University Press.

Lyall, J. (2010). Are coethnics more effective counterinsurgents? evidence from the second chechen war. *American Political Science Review*, 104(1):1–20.

Mazzei, J. (2009). *Death Squads or Self-Defense Forces? How Paramilitary Groups Emerge and Challenge Democracy in Latin America*. University of North Carolina Press, Chapel Hill, NC.

Mitchell, N. J., Carey, S. C., and Butler, C. K. (2014). The impact of pro-government militias on human rights violations. *International Interactions*, 40(5):812–836.

Myerson, R. B. (2008). The autocrat's credibility problem and foundations of the constitutional state. *American Political Science Review*, 102(1):125–139.

North, D. C., Wallis, J. J., and Weingast, B. R. (2009). *Violence and social orders: A conceptual framework for interpreting recorded human history.* Cambridge University Press.

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

Powell, R. (2013). Monopolizing violence and consolidating power. *The Quarterly Journal of Economics*, 128(2):807–859.

Salehyan, I. (2010). The delegation of war to rebel organizations. *Journal of conflict resolution*, 54(3):493–515.

Stanton, J. A. (2015). Regulating militias: Governments, militias, and civilian targeting in civil war. *Journal of Conflict Resolution*, 59(5):899–923.

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

Weber, M. (1946). Politics as a vocation. In Gerth, H. H. and Mills, C. W., editors, *From Max Weber: Essays in Sociology*, pages 77–128. Oxford University Press, New York, NY.