Snakes in the Backyard*

Ravi Ashok Satpute †

March 30, 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 decouples attack outcomes from state strength but endogenously creates a political claimant who may later contest power with private information about the cost of doing so. I characterize PBE and show that in the key equilibrium, powerful states attack directly while weak states delegate, accepting blowback risk because direct action's type-dependent success probability makes it a costly gamble relative to accepting the risk of 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

†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 deploy armed non-state groups that later contest political power, extract concessions, or become 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, 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 regimes: 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 some of the democracies have 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 influencing elections and becoming autonomous actors. In most of these settings, the proxy relationship created risks that materialized over time: groups turned against patrons, demanded political concessions, 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 also makes them dangerous for the state.

The literature offers three plausible explanations across diverse fields for state reliance on violent proxies. First, proxies are *militarily efficient* in a sense 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,

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

2009, Lyall, 2010). This explanation captures real advantages but cannot account for why states use and 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 with reduced liability by 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 avoid accountability, escape sanctions, reputational costs, or intervention (Mitchell et al., 2014, Stanton, 2015, Carey et al., 2015). This account has merit for explaining the use of proxies in international conflicts and as an alternate foreign policy, 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 to 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?

This paper provides a formal theory of endogenous rival creation through proxy delegation, rooted in a simple observation: violence modes differ in their informational content. Direct attacks reveal state capacity to domestic audiences, while delegated attacks do not. I show that this informational asymmetry generates a signaling motive for proxy delegation that is distinct from military efficiency, deniability, or capacity constraints. In the key equilibrium, weak states rationally empower proxies that may later contest political power, because direct action's low success probability for weak states makes it materially inferior to accepting future competition from an empowered rival whose success is state-type independent. This mechanism, in which the act of delegation itself creates a political competitor by concealing information from voters, is new to the formal political economy literature. I characterize the full equilibrium structure, establishing existence and non-existence results across separating, pooling, and semi-pooling configurations, and derive comparative statics that yield testable predictions on when delegation occurs, when blowback follows, and how proxy contracts mediate this tradeoff.

The underlying logic operates through the domestic political information channel. When the state uses violence, the outcome reveals information about the state's capacity to domestic audiences: successful violence signals strength, and failure signals weakness. Voters observe these outcomes and update their beliefs about the state's ability to maintain order and project power. These updated beliefs, in turn, shape political survival. Crucially, the informativeness of violence depends on its mode. When the state attacks directly, outcomes are diagnostic: success reflects capacity, and failure reflects incapacity. When the state delegates to a proxy, outcomes are uninformative about state capacity because they depend on the proxy's effectiveness rather than the state's own strength. The delegation, in other words, obscures the mapping from violence outcomes to beliefs about the state strength.

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 state strength. 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.

The equilibrium analysis yields 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 important insights. A more generously compensated proxy has a higher opportunity cost of political activation, reducing the likelihood of the proxy's political entry. Moreover, as proxy inefficiency increases, the proxy becomes militarily weaker, reducing the value of delegation for the state, and shrinking 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, where revealing weakness is politically costly, 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 used 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 generous shares of captured resources have weaker incentives to contest political power, suggesting a tradeoff between short-run resource extraction and long-run political stability.

The paper engages with 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 state 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, Powell, 2006, Jackson and Morelli, 2007, Fey and Ramsay, 2011, Baliga and Sjöström, 2004, 2012). 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 ultimately determine the incumbent regime's survival.

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 (Downs, 1957, Alesina, 1988, Persson et al., 1997, Lizzeri and Persico, 2001). A related strand on endogenous candidacy allows citizens to self-select into political competition (Osborne and Slivinski, 1996, Besley and Coate, 1997). 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 due to *proxy delegation*. Another closest antecedent to my mechanism is Powell (2013), who studies an infinite-horizon model in which a ruler decides whether and how quickly to consolidate power by disarming an armed opposition. In Powell's framework, the ruler cannot credibly commit to future transfers: once the opposition's military capacity is reduced, the ruler has an incentive to renege on promised concessions. This commitment problem over resource-sharing generates inefficient fighting, and the outcome depends on whether "contingent spoils" (economic gains) from monopolizing violence are large enough to justify a costly conflict rather than a peaceful buyout. My paper differs in two aspects. First, the driving friction is *informational*, not contractual: the state delegates to hide weakness from domestic audiences, not because it cannot commit to 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.

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. While Section 4 derives some important comparative statics, Section 5 discusses extensions, endogenous surplus sharing with optimal contracts, and an outside option. Finally, Section 6 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 violent 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, denoted by:

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

where $\lambda > 0$ measures how incompetent or ineffective the proxy is. Thus, a higher λ means an ineffective and 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 $\pi \in (0, 1)$, where k is the post-violence share offered by the state to the proxy (discussed in section 2.7.1).

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

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.[2] 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[3].

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 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 $p_P = \frac{c_P}{c_P + \lambda}$. 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 proxy attack outcomes are uninformative about θ, essentially capturing the idea that proxy delegation decouples attack outcomes from state strength. If the proxy attack succeeds, the share τ is captured from the target. If the attack fails, no redistribution occurs.

[2] In Section 5.2, I discuss a no attack option.

[3] In Section 5.1, I discuss endogenous k with optimal contracts.

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

$$p_P = \frac{c_P}{c_P + \lambda} < 1 = p_D(\theta_H). \tag{2.8}$$

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:

$$\sigma \in \Sigma = \{(D, S), (D, F), (P, S), (P, F)\}, \tag{2.9}$$

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:

$$\Theta_D = \{\theta : s_\theta = D\}, \tag{2.10}$$

$$\Theta_P = \{\theta : s_\theta = P\}. \tag{2.11}$$

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

$$\mu(\sigma) = \Pr(\theta = \theta_H \mid \sigma). \tag{2.12}$$

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

$$\mathbb{E}[\theta \mid \sigma] = \mu(\sigma). \tag{2.13}$$

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

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

$$\mu(P, S) = \mu(P, F) = \Pr(\theta_H \mid \theta \in \Theta_P). \tag{2.14}$$

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

$$\mu(D, S) = \frac{p}{p + (1 - p)\alpha}, \tag{2.15}$$

$$\mu(D, F) = 0. \tag{2.16}$$

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 proofs of both lemmas are straightforward applications 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 model's informational foundation. As proxy attacks do not update beliefs about state strength, whereas direct attacks do, the choice of attack mode is itself a signal, giving rise to the strategic considerations analyzed in the equilibrium.

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 relevant political contest (in this context) 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.

The absence of a political contest after a direct attack is not an assumption of convenience; it reflects the model's core asymmetry: only delegation creates the organizational platform and demonstrated competence from which a credible challenge can be mounted. Direct attack outcomes are informative about state strength, but this information is consequential only to the extent that it shapes the political contest triggered by delegation.

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:

$$\text{Voters retain state} \iff \mu(\sigma) \geq \gamma. \tag{2.17}$$

- 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:*

$$w = (1 - \tau) + \tau = 1. \tag{2.18}$$

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.[4]

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:

$$c_A < \hat{c}(\mu^*) \equiv \begin{cases} w - k\tau = 1 - k\tau & \text{if } \mu^* \leq \gamma, \\ -k\tau & \text{if } \mu^* > \gamma. \end{cases} \tag{2.19}$$

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 politically is:*

$$\pi = H(1 - k\tau). \tag{2.20}$$

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.

[4] In practice, power transitions may be partial. I adopt the deterministic rule for tractability to isolate the informational mechanism as cleanly as possible. Under a probabilistic contest technology (e.g., the proxy wins with probability increasing in γ/μ or $\mu < \gamma$), the proxy's activation threshold would become belief-dependent, and the stark outcome where any activating proxy defeats a weak delegator (Remark 4) would be attenuated. The qualitative structure would still survive, but the all-or-nothing electoral outcome would be replaced by graded risk.

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 following a direct attack.

$$U_S^D(\theta_H) = (1-\tau) - c_D + \tau = 1 - c_D, \tag{2.21}$$

$$U_S^D(\theta_L) = (1-\tau) - c_D + \alpha \cdot \tau = (1-\tau) - c_D + \alpha\tau. \tag{2.22}$$

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 $(1-k)\tau$:

$$U_S^P = (1-\tau) - c_P + p_P(1-k)\tau. \tag{2.23}$$

Case 2: $\mu^* \leq \gamma$ (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 $-c_P$:

$$\begin{aligned} U_S^P &= (1-p_P)[(1-\tau) - c_P] \\ &\quad + p_P(1-\pi)[(1-\tau) - c_P + (1-k)\tau] \\ &\quad + p_P \cdot \pi \cdot [-c_P] \qquad (2.24) \\ &= (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \qquad (2.25) \end{aligned}$$

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:

$$\frac{1-k}{1} < \frac{\pi}{1-\pi} \cdot \frac{1-\tau}{\tau} \iff \text{delegation destroys value for the state.} \tag{2.26}$$

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:

$$U_V = \begin{cases} \theta & \text{if the state governs,} \\ \gamma & \text{if the proxy governs.} \end{cases} \tag{2.27}$$

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 (s^*, β^*) and beliefs μ^* such that:*

1. *For each θ, $s^*(\theta)$ 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 delegates to the proxy, exploiting the type-independent success probability of proxy violence. This is the equilibrium that generates the "snakes in the backyard" dynamic: weak states rationally create political rivals by delegating violence. Note that in the separating equilibrium, delegation does not conceal the state's type from voters; voters correctly infer that delegators are weak. The strategic logic is that weak states prefer the type-independent lottery of proxy violence (success probability p_P) to the type-dependent lottery of direct violence (success probability α), even though the former creates blowback risk. It is the differential informativeness of direct attack across types that makes it unattractive for the weak type, thereby sustaining separation.

Proposition 3.1 (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:*

$$c_D - c_P \leq \tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{3.2}$$
$$c_D - c_P \geq \alpha\tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{3.3}$$

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 a 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 a 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 3.2 (Pooling on Proxy Equilibrium)**.** *Consider the strategy profile* $(s(\theta_H) = P, s(\theta_L) = P)$*. Under this profile:*

$$\mu(P, S) = \mu(P, F) = p. \tag{3.4}$$

This is a PBE *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. This is essentially a *blowback-free* delegation. 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 coincides with $\Delta > \overline{\Delta}_P$ (defined in the section 3.6), while the separating equilibrium requires $\Delta \leq \overline{\Delta}_S$ (condition (3.2)). Since $\overline{\Delta}_P < \overline{\Delta}_S$, there is an overlap region $\Delta \in (\overline{\Delta}_P, \overline{\Delta}_S]$ where both the pooling-on-proxy and separating equilibria coexist (see Section 3.6).

Remark 2. *When $p \leq \gamma$, a pooling-on-proxy equilibrium with on-path blowback may exist: $\mu(P, S) = p \leq \gamma$ implies the proxy activates with probability $\pi > 0$. The deviation conditions change substantially, as both types face blowback on the equilibrium path. Analysis of this configuration, which requires the prior to be sufficiently pessimistic, is beyond the scope of the present paper.*

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

$$\mu(D, S) = \frac{p}{p + (1-p)\alpha}, \tag{3.5}$$

$$\mu(D, F) = 0. \tag{3.6}$$

This is a PBE *if:*

$$c_D - c_P \leq \alpha\tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] = \underline{\Delta} \tag{3.7}$$

under off-path beliefs $\mu(P, S) \leq \gamma$, which are consistent with the Intuitive Criterion (Cho and Kreps, 1987).

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 (3.7) 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 3.4 (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:*

$$\mu(D,S) = \frac{p}{p+(1-p)q\alpha}, \tag{3.8}$$

$$\mu(D,F) = 0, \tag{3.9}$$

$$\mu(P,S) = \mu(P,F) = 0. \tag{3.10}$$

This constitutes a PBE *if and only if:*

$$c_D - c_P = \alpha\tau - p_P\left[(1-\pi)(1-k)\tau - \pi(1-\tau)\right], \tag{3.11}$$

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

Proof. In Appendix A.5 □

Remark 3 (Non-Genericity)**.** *The semi-pooling equilibrium exists if and only if condition (3.11) 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 3.1).*

Remark 4 (Indeterminacy of the Mixing Probability)**.** *When condition (3.11) 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 (Appendix A.1): powerful states delegate while weak states attack directly. Proposition A.1 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 relatively safer, pooling option is available.

3.6 Equilibrium Mapping

Collecting the results from Propositions 3.1 - 3.4 and Appendix A.1, the equilibrium structure is governed by a single sufficient statistic: the cost differential $\Delta \equiv c_D - c_P$ between direct and delegated violence. I now formalize how the equilibrium regions partition the Δ-line. Define three threshold values of Δ:

$$\underline{\Delta} \equiv \alpha\tau - p_P \left[(1-\pi)(1-k)\tau - \pi(1-\tau)\right], \tag{3.12}$$

$$\overline{\Delta}_P \equiv \tau - p_P(1-k)\tau, \tag{3.13}$$

$$\overline{\Delta}_S \equiv \tau - p_P \left[(1-\pi)(1-k)\tau - \pi(1-\tau)\right]. \tag{3.14}$$

Assumption 2 (Sufficient Type Separation)**.** *The military gap between state types exceeds the blowback wedge (Γ):*

$$\tau(1-\alpha) > \Gamma. \tag{3.15}$$

where $\Gamma \equiv p_P\pi(1-k\tau) > 0$

The left side of the inequality measures the relative military advantage of the strong state. The right-side, blowback wedge (Γ), measures the difference in expected cost between facing blowback (under separating beliefs) and avoiding it (under pooling beliefs). It captures the payoff consequence of voter expectations: when voters interpret delegation as a signal of weakness, the state suffers an expected loss of Γ relative to the case where delegation is uninformative. This assumption requires that the weak state is sufficiently militarily disadvantaged relative to the political cost of blowback. When the assumption fails, the separating equilibrium still exists, but it coexists with a pooling equilibrium; the "snakes in the backyard" mechanism remains operative but no longer generates a unique prediction.

Under Assumption 2, the three thresholds are strictly ordered: $\underline{\Delta} < \overline{\Delta}_P < \overline{\Delta}_S$. To see this, note that $\overline{\Delta}_S - \underline{\Delta} = \tau(1-\alpha) > 0$ and $\overline{\Delta}_S = \overline{\Delta}_P + \Gamma$, so $\overline{\Delta}_P < \overline{\Delta}_S$. The remaining ordering $\underline{\Delta} < \overline{\Delta}_P$ follows from $\overline{\Delta}_P - \underline{\Delta} = \tau(1-\alpha) - \Gamma > 0$, which holds by Assumption 2. The following proposition provides a complete equilibrium characterization.

Proposition 3.5 (Equilibrium Characterization)**.** *Suppose $p > \gamma$ and assumption 2 holds. The cost differential Δ partitions the parameter space into four generic regions ($p \leq \gamma$ is discussed in Remark 2):*

*(i) $\Delta < \underline{\Delta}$: **Pooling on Direct** is the unique pure-strategy PBE.*

*(ii) $\underline{\Delta} < \Delta \leq \overline{\Delta}_P$: **Separating** is the unique pure-strategy PBE.*

*(iii) $\overline{\Delta}_P < \Delta \leq \overline{\Delta}_S$: Both the **Separating** and **Pooling on Proxy** equilibria coexist.*

(iv) $\Delta > \overline{\Delta}_S$*:* ***Pooling on Proxy*** *is the unique pure-strategy PBE.*

At the boundary $\Delta = \underline{\Delta}$*, the Pooling on Direct, Separating, and Semi-Pooling equilibria all coexist.*

Proof. In Appendix A.6 □

Note that the unique separating region (ii) has width $\tau(1-\alpha) - \Gamma$. As blowback risk increases (π rises), the unique separating region shrinks, and the multiplicity region expands. The "snakes in the backyard" mechanism is most sharply predictive when blowback is *moderate*, large enough to matter but not so large as to make voter expectations decisive. The multiplicity region $(\overline{\Delta}_P, \overline{\Delta}_S]$ exhibits a self-fulfilling prophecy structure. In this range, two self-consistent equilibria coexist, distinguished not by fundamentals but by how voters interpret the state's choice of violence mode.

The key property is that expectations are self-validating. If voters *expect* delegation to signal weakness, only weak states delegate, blowback follows, and the expectation is confirmed. If voters *expect* delegation to be uninformative, both types delegate, no blowback occurs, and the expectation is again confirmed. Whether the state creates a political rival depends not on the conflict environment, which is identical across equilibria, but on the prevailing interpretation of state violence.

This multiplicity is robust to standard refinements. In both equilibria, the only off-path signal that could be refined is the direct attack outcome. But direct attack triggers no political contest, so off-path beliefs after $m = D$ are payoff-irrelevant. Neither the Intuitive Criterion (Cho and Kreps, 1987) nor D1 selects between the two equilibria. Equilibrium selection depends on how domestic audiences coordinate their interpretation of proxy violence, an inherently political determination, external to the model, that could vary across countries, historical periods, or institutional environments.

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 4.1 (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 4.2 (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 $\overline{\Delta}_S = \tau - p_P[(1 - \pi)(1 - k)\tau - \pi(1 - \tau)]$ is independent of α.*
3. *The width of the separating equilibrium region, $\overline{\Delta}_S - \underline{\Delta} = \tau(1 - \alpha)$, 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 with width τ. The region is non-empty for all $\alpha < 1$.*

Proof. (1) $\frac{\partial \underline{\Delta}}{\partial \alpha} = \tau > 0$. (2) $\overline{\Delta}$ does not contain α. (3) Follows from (1) and (2). (4) The width is $\overline{\Delta}_S - \underline{\Delta} = \tau(1-\alpha)$. At $\alpha = 0$, width is $\tau > 0$. Since $\alpha < 1$, the width is strictly positive for all permissible α. □

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 4.3 (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. *Both thresholds $\underline{\Delta}$ and $\overline{\Delta}_S$ shift downward as λ increases. The* width $\tau(1-\alpha)$ *is invariant to λ, but for a fixed cost differential Δ, the separating region shifts leftward and pooling on D becomes more likely.*

Proof. (1) and (2) follow directly from the definitions. For (3), note that $\partial\underline{\Delta}/\partial p_P < 0$ and $\partial\overline{\Delta}_S/\partial p_P < 0$, so both thresholds decrease as p_P falls (i.e., as λ rises). The width $\overline{\Delta}_S - \underline{\Delta} = \tau(1-\alpha)$ is independent of p_P. As $p_P \to 0$, both thresholds converge to $\alpha\tau$ and τ respectively, and the entire separating region shifts left. For a fixed Δ, the state is more likely to fall in the pooling-on-direct region. □

Remark 5 (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 Extensions

5.1 Endogenous Surplus Sharing

In the baseline model, the proxy's share k of captured surplus is exogenous. I now endogenize this sharing rule by allowing the state to make a take-it-or-leave-it offer to the proxy following a successful proxy attack. This extension formalizes the tradeoff identified in the comparative statics: a larger k reduces blowback but also reduces the state's surplus from delegation.

5.1.1 Modified Timing

The timing of the game is modified as follows. After a successful proxy attack ($\sigma = (P, S)$), a new sub-stage is inserted:

- **Stage 2.5 (Surplus Division):** The state makes a take-it-or-leave-it offer $k \in [0, 1]$ to the proxy, specifying the proxy's share of the captured surplus τ.

The proxy then observes k together with its private cost c_A and decides whether to accept (yielding payoff $k\tau$) or reject and activate politically (yielding payoff $1 - c_A$ if it wins the political contest, which occurs with certainty when $\mu(P, S) < \gamma$). If the proxy rejects but activation is unprofitable ($c_A \geq 1$), the proxy receives 0; however, since $k\tau > 0$, such types always accept the offer. All other aspects of the game remain unchanged.

5.1.2 The Proxy's Acceptance Decision

The proxy's decision rule is identical in structure to the baseline. Given offer k, the proxy accepts if and only if the offer exceeds the net gain from political activation:

$$k\tau \geq 1 - c_A \iff c_A \geq 1 - k\tau. \tag{5.1}$$

The proxy rejects the offer and activates with probability:

$$\pi(k) = H(1 - k\tau). \tag{5.2}$$

This is the same activation probability as in the baseline model. The difference is that k is now a choice variable.

5.1.3 The State's Optimal Offer

I focus on the separating equilibrium, where the proxy's activation decision is the only relevant consideration (since $\mu(P, S) = 0 < \gamma$ and the proxy wins with certainty if it activates). At Stage 2.5, the cost c_P is already sunk. The state chooses k to maximize its expected continuation payoff after proxy success:

$$V_S(k) = \underbrace{[1 - H(1 - k\tau)]}_{\text{prob. proxy accepts}} \cdot \underbrace{(1 - k\tau)}_{\text{state's surplus if accepted}} + \underbrace{H(1 - k\tau)}_{\text{prob. proxy activates}} \cdot 0. \tag{5.3}$$

The second term reflects that if the proxy activates and wins, the state loses the entire pie. Note that although c_P is relevant for payoffs, it is irrelevant for the optimization decision as it appears on both sides.

Proposition 5.1 (Optimal Surplus Sharing)**.** *Suppose H has a strictly increasing hazard rate on $[\underline{c}, \bar{c}]$. Then the state's optimal offer k^* is uniquely characterized as follows. Define $\hat{c}^*$ as the unique solution to:*

$$\frac{1 - H(\hat{c}^*)}{h(\hat{c}^*)} = \hat{c}^*, \tag{5.4}$$

where $\hat{c}^* \in (\underline{c}, \bar{c})$. The optimal sharing rule is:

$$k^* = \frac{1 - \hat{c}^*}{\tau}, \tag{5.5}$$

provided $k^* \leq 1$ (i.e., $\hat{c}^* \geq 1 - \tau$). If $\hat{c}^* < 1 - \tau$, the state sets $k^* = 1$, offering the entire captured surplus to the proxy.

Proof. In Appendix A.7 □

The characterization has a natural economic interpretation. Condition (5.4) equates the *inverse hazard rate* of the proxy's activation cost distribution, evaluated at the threshold $\hat{c}^*$, to the threshold itself. The inverse hazard rate $(1 - H(\hat{c}))/h(\hat{c})$ measures the "mass of remaining types" per unit density at the margin: it captures how many additional proxy types the state can buy off by marginally increasing its offer. Thus, the state is directly choosing the cutoff proxy type $\hat{c}$ it wants to appease. The state balances two forces: a higher offer (lower $\hat{c}$) buys off more proxy types, reducing blowback, but also reduces the state's surplus from each bought-off type. The optimum trades off the marginal reduction in blowback against the marginal decrease in surplus, precisely the condition captured by (5.4). Formally, this is equivalent to a monopolist's pricing problem: the state sets a price $\hat{c}$ for political peace, facing demand $1 - H(\hat{c})$ from proxy types willing to accept that price, and maximizes expected revenue $\hat{c} \cdot [1 - H(\hat{c})]$.

Corollary 2 (Uniform Distribution)**.** *If $c_A \sim Uniform[0, 1]$, then $\hat{c}^* = \frac{1}{2}$, and:*

1. *The optimal sharing rule is $k^* = \frac{1}{2\tau}$, which is interior (i.e., $k^* \leq 1$) iff $\tau \geq \frac{1}{2}$.*
2. *When the interior solution obtains ($\tau \geq \frac{1}{2}$), the equilibrium blowback probability is $\pi^* = \frac{1}{2}$. At the corner ($\tau < \frac{1}{2}$, $k^* = 1$), the blowback probability is $\pi^* = 1 - \tau > \frac{1}{2}$.*
3. *The state's optimal continuation payoff after proxy success is $V_S^* = \frac{1}{4}$.*

Proof. For $H(x) = x$ on $[0, 1]$, the FOC becomes $(1 - \hat{c})/1 = \hat{c}$, yielding $\hat{c}^* = 1/2$. Then $k^* = (1 - 1/2)/\tau = 1/(2\tau)$, $\pi^* = H(1/2) = 1/2$, and $V_S^* = (1/2)(1/2) = 1/4$. □

The uniform case illustrates a sharp result: when the interior solution obtains ($\tau \geq 1/2$), the state accepts exactly a 50% blowback probability regardless of τ. When $\tau < 1/2$, the state hits the corner $k^* = 1$ and the blowback probability becomes $\pi^* = 1 - \tau$, which exceeds 50% and depends on τ. What varies with τ is the generosity of the offer: a larger contested resource allows the state to buy off the proxy with a smaller share.

5.1.4 Properties of the Optimal Sharing Rule

Proposition 5.2 (Comparative Statics on Optimal Sharing)**.** *Under the conditions of Proposition 5.1:*

1. *For $k^* < 1$ (interior solution), the blowback probability $\pi^* = H(\hat{c}^*)$ depends only on the distribution H and not on target value τ, proxy capability λ, or any other model parameter. At corner $k^* = 1$, $\pi^* = H(1-\tau)$ depends on τ.*
2. *The optimal share $k^* = (1-\hat{c}^*)/\tau$ is strictly decreasing in τ: when the contested resource is larger, the state can afford to offer a smaller share.*
3. *The state's ex ante expected payoff from delegation (in the separating equilibrium) with optimal surplus sharing is:*
$$U_S^P(\theta_L; k^*) = (1-p_P)(1-\tau) + p_P \cdot V_S^* - c_P, \tag{5.6}$$
where $V_S^* = [1-H(\hat{c}^*)] \cdot \hat{c}^*$ is a constant determined entirely by H.

Proof. Proof in Appendix A.8 □

Property (1) is a key result. It states that the *probability of blowback* is invariant to the economic environment once the state optimizes its offer. The state always chooses the same threshold type $\hat{c}^*$ to separate accepting and rejecting proxies, and hence the same blowback probability. What adjusts in response to changes in τ is the actual share k^* offered, not the equilibrium level of political instability. This result is reminiscent of optimal auction design, where the seller's optimal reserve price depends on the distribution of buyer valuations but not on the seller's opportunity cost.

5.1.5 Implications for Equilibrium Existence

The separating equilibrium conditions in the baseline model (Proposition 3.1) required $c_D - c_P$ to lie in an intermediate range determined by the fixed k. With endogenous surplus sharing, the weak state's payoff from delegation becomes $U_S^P(\theta_L; k^*) = (1-p_P)(1-\tau) + p_P V_S^* - c_P$. The equilibrium conditions, appropriately modified, become:

$$c_D - c_P \leq \tau - p_P V_S^* + p_P(1-\tau), \tag{5.7}$$
$$c_D - c_P \geq \alpha\tau - p_P V_S^* + p_P(1-\tau). \tag{5.8}$$

The interval width remains $\tau(1-\alpha)$, exactly identical to the exact baseline formulation. Endogenizing k simply shifts the equilibrium region upwards since the optimized expected state payoff increases, but it does not expand or shrink it: the informational tradeoff between revealing type and creating a rival is unchanged; only the proxy's compensation adjusts optimally.

5.2 Endogenous Violence - The No-Attack Option

In the baseline model, the state must attack; it chooses only the mode. This captures settings where the state has already committed to addressing the target. In practice, however, states can also choose inaction: tolerating the target's control over τ and avoiding the costs and risks of violence altogether. I now introduce this option and characterize when the "snakes in the backyard" logic operates, and when it breaks down.

5.2.1 Modified Strategy Space

The state's strategy set is expanded to:

$$m \in \{D, P, \varnothing\}, \tag{5.9}$$

where $m = \varnothing$ denotes *inaction*. Under *inaction*, no violence occurs, no resources are redistributed, and no signal is generated. The state retains its initial share $(1 - \tau)$ and incurs no cost. Thus,

$$U_S^{\varnothing}(\theta) = 1 - \tau \qquad \text{for both } \theta \in \{\theta_L, \theta_H\}. \tag{5.10}$$

5.2.2 The Participation Condition

The introduction of inaction imposes an additional constraint on each equilibrium: each type must prefer its equilibrium action to doing nothing. For the powerful state, direct attack is preferred to inaction whenever $1 - c_D \geq 1 - \tau$, i.e.:

$$c_D \leq \tau. \tag{5.11}$$

This is a simpler condition saying that the cost of a direct attack must be lower than the share of the pie held by the target (T). If $c_D > \tau$, the powerful state also prefers inaction, so no violence occurs regardless of type. For the weak state, the relevant comparison is between delegation and inaction. Now, the *net expected value of delegation* (say $\mathcal{V}$) is given by

$$\mathcal{V} \equiv p_P \left[(1 - \pi)(1 - k)\tau - \pi(1 - \tau)\right] - c_P. \tag{5.12}$$

Now, the weak state's payoff from delegation is $(1 - \tau) + \mathcal{V}$, so it prefers delegation to inaction if and only if $\mathcal{V} \geq 0$.

Proposition 5.3 (Active vs. Passive Separating Equilibrium)**.** *Suppose $c_D \leq \tau$.*

1. *If $\mathcal{V} \geq 0$: the* active separating equilibrium *($\theta_H \to D$, $\theta_L \to P$) from Proposition 3.1 survives, subject to the baseline incentive compatibility conditions* (3.2)-(3.3) *and the additional requirement $c_D \leq \tau$.*
2. *If $\mathcal{V} < 0$ and $c_D \geq \alpha\tau$: a* passive separating equilibrium *($\theta_H \to D$, $\theta_L \to \varnothing$) exists, supported by the off-path belief $\mu(P, S) = 0$.*

Proof. In Appendix A.9 □

Now, the participation condition $\mathcal{V} \geq 0$ can be rewritten as:

$$\underbrace{p_P(1-\pi)(1-k)\tau}_{\text{expected capture if proxy stays inactive}} \geq \underbrace{c_P + p_P\pi(1-\tau)}_{\text{delegation cost + expected blowback loss}} . \tag{5.13}$$

The "snakes in the backyard" mechanism is operative only when the expected spoils from delegation exceed its combined costs. The mechanism shuts down, and the weak state prefers inaction, when: *(i)* blowback is highly likely increasing the expected loss $p_P\pi(1-\tau)$, *(ii)* the contested resource is small (τ low), reducing the gains from violence while increasing the state's exposure to political displacement, *(iii)* delegation is too expensive (c_P high), or *(iv)* when the state offers generous rewards (k high).

The no-attack option introduces a participation constraint that bounds the domain of the paper's central mechanism. The active separating equilibrium, in which delegation generates endogenous blowback, arises only when the expected value of proxy violence justifies the attendant political risk. When it does not, the model predicts a qualitatively different outcome: powerful states reveal themselves through action while weak states reveal themselves through inaction. Both equilibria are informationally separating, but they differ in their consequences. The active equilibrium produces endogenous political instability; the passive equilibrium produces political stagnation. This distinction reveals potential empirically testable consequences. The model predicts that proxy delegation should be observed in environments with valuable targets, reasonably effective proxies, and moderate blowback risk. When conditions are unfavorable, volatile political environments, low-value targets, or unreliable proxies, weak states should refrain from violence altogether rather than gamble on delegation. This is one of the potential reasons why some weak states with access to potential proxies nevertheless do not use them.

6 Conclusion

This paper develops a formal model explaining why states empower domestic violent actors despite the risk of future political competition. The key mechanism is informational: direct violence has type-dependent success probabilities, while delegated violence does not, decoupling outcomes from state strength. Because voters condition political support on perceived strength, weak states face incentives to delegate even at the cost of creating political rivals. The separating equilibrium 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 the reverse separating configuration is ruled out entirely.

The analysis yields several empirically relevant insights. First, proxy delegation should be more prevalent among states for which revealing weakness is politically costly, rather than merely among those lacking military capacity. Second, blowback is endogenous to the delegation decision; it can be mitigated through more generous resource-sharing arrangements, though at the expense of the state's surplus. The model also highlights a weak proxy paradox: inherently 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, providing testable implications for state-proxy relationships and subsequent political transitions.

Several extensions merit future investigation. The political contest could very well be non-electoral to formally capture the violent state takeovers by armed groups. Introducing voter heterogeneity or allowing the proxy to strategically choose its level of political visibility would enrich the political contest stage. Replacing the deterministic electoral rule with a probabilistic contest technology would moderate the severity of blowback while preserving the core signaling mechanism. Finally, a 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 such as Colombia, Lebanon, and Palestine.

A Appendix

A.1 The Reverse Separating Equilibrium (Non-existence): $\theta_H \to P$, and $\theta_L \to D$

Proposition A.1 (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:

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

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 3.1

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, powerful state's payoff under D:

$$U_S^D(\theta_H) = 1 - c_D. \tag{A.11}$$

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:

$$U_S^P(\theta_H; \mu^* = 0) = (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)]. \tag{A.12}$$

For θ_H not to deviate, we require $U_S^D(\theta_H) \geq U_S^P(\theta_H; \mu^* = 0)$:

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

This is condition (3.2). Note that this is the *exact* no-deviation condition: the deviation payoff is fully determined because the proxy signal is on-path under the separating strategy ($\mu(P, S) = 0$ follows from Bayes' rule), so no off-path beliefs are involved. 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 $\mu(P, S) = 0$:

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

If θ_L deviates to D: It succeeds with probability α and captures τ:

$$U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau. \tag{A.14}$$

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:

$$U_S^P(\theta_L; \mu^* = 0) \geq U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau. \tag{A.15}$$

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 $\big(\, \alpha\tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)],\ \ \tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \,\big)$. This interval has width $\tau(1-\alpha) > 0$, so it is non-empty for all $\alpha < 1$. □

A.3 Proof of Proposition 3.2

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:

$$U_S^P = (1-\tau) - c_P + p_P(1-k)\tau. \tag{A.16}$$

Deviation incentives: Consider θ_H deviating to D. Since θ_H always succeeds, no proxy relevant political contest arises. The signal (D, S) is off-path. θ_H's deviation payoff:

$$U_S^D(\theta_H) = 1 - c_D. \tag{A.17}$$

For θ_H not to deviate:

$$(1-\tau) - c_P + p_P(1-k)\tau \geq 1 - c_D \tag{A.18}$$

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

Similarly, θ_L doesn't deviate if:

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

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

Hence, from equation (A.19) and (A.21) 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. □

A.4 Proof of Proposition 3.3

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 $\mu(P, S) = \tilde{\mu}$.

Case 1: $\tilde{\mu} > \gamma$ (no blowback on deviation). Deviation payoff is:

$$U_S^P(\text{no blowback}) = (1-\tau) - c_P + p_P(1-k)\tau. \tag{A.22}$$

For θ_L not to deviate:

$$(1-\tau) - c_D + \alpha\tau \geq (1-\tau) - c_P + p_P(1-k)\tau, \tag{A.23}$$

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

Case 2: $\tilde{\mu} \leq \gamma$ (blowback on deviation). Deviation payoff reduced by blowback risk:

$$U_S^P(\text{blowback}) = (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)]. \tag{A.25}$$

For θ_L not to deviate:

$$c_D - c_P \leq \alpha\tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)]. \tag{A.26}$$

Case 2 is strictly weaker (larger RHS), so it supports the equilibrium for a wider range of Δ. Under the *Intuitive Criterion* (Cho and Kreps, 1987), the weak type has a stronger incentive to deviate to proxy (to hide information), so voters should infer that a deviating state is weak, supporting $\mu(P, S) \leq \gamma$. Under pessimistic off-path beliefs consistent with the *Intuitive Criterion*, the pooling-on-direct equilibrium exists for all $\Delta \leq \underline{\Delta}$. In both cases, θ_H's no-deviation condition is implied by θ_L's (since $\alpha < 1$). □

A.5 Proof of Proposition 3.4

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:
$$\mu(D, S) = \frac{p \cdot 1}{p \cdot 1 + (1-p) \cdot q \cdot \alpha} = \frac{p}{p + (1-p)q\alpha}. \tag{A.27}$$
- 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:

$$U_S^P(\theta_H) = (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{A.28}$$

To prevent this deviation, we need

$$1 - c_D \geq (1-\tau) - c_P + p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{A.29}$$
$$c_D - c_P \leq \tau - p_P[(1-\pi)(1-k)\tau - \pi(1-\tau)] \tag{A.30}$$

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

$$U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau \tag{A.31}$$

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

$$U_S^P(\theta_L) = (1-\tau) - c_P + p_P(1-\pi)(1-k)\tau - p_P\pi(1-\tau) \tag{A.32}$$

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

$$U_S^D(\theta_L) = U_S^P(\theta_L) \tag{A.33}$$
$$(1-\tau) - c_D + \alpha\tau = (1-\tau) - c_P + p_P(1-k)(1-\pi)\tau - p_P\pi(1-\tau) \tag{A.34}$$
$$c_D - c_P = \alpha\tau - p_P\left[(1-k)(1-\pi)\tau - \pi(1-\tau)\right]. \tag{A.35}$$

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.35):

$$U_S^P(\theta_H) = U_S^P(\theta_L) = U_S^D(\theta_L) = (1-\tau) - c_D + \alpha\tau. \tag{A.36}$$

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

$$1 - c_D > (1-\tau) - c_D + \alpha\tau \tag{A.37}$$
$$1 > 1 - \tau(1-\alpha). \tag{A.38}$$

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.30) is automatically satisfied whenever the indifference condition (A.35) holds. □

A.6 Proof of Proposition 3.5

Proof. The separating equilibrium exists if and only if $\Delta \in [\underline{\Delta}, \overline{\Delta}_S]$ (Proposition 3.1). The blowback-free pooling-on-proxy equilibrium exists if and only if $\Delta > \overline{\Delta}_P$ (Proposition 3.2). Pooling on direct is sustained for $\Delta \leq \underline{\Delta}$ under pessimistic off-path beliefs (Proposition 3.3). Under Assumption 2, $\underline{\Delta} < \overline{\Delta}_P < \overline{\Delta}_S$. In region (i), only pooling on direct survives: $\Delta < \underline{\Delta}$ excludes separating and $\Delta < \overline{\Delta}_P$ excludes pooling on proxy. In region (ii), $\Delta > \underline{\Delta}$ excludes pooling on direct and $\Delta \leq \overline{\Delta}_P$ excludes pooling on proxy; only separating remains. In region (iii), both $\Delta \in [\underline{\Delta}, \overline{\Delta}_S]$ and $\Delta > \overline{\Delta}_P$ are satisfied. In region (iv), $\Delta > \overline{\Delta}_S$ excludes separating. The boundary claim follows from the coincidence of conditions at $\Delta = \underline{\Delta}$ and proposition 3.4. □

A.7 Proof of Proposition 5.1

Proof. Substituting $\hat{c} = 1 - k\tau$ (so that $k = (1-\hat{c})/\tau$ and the state's surplus conditional on acceptance is $\hat{c}$), the state's problem becomes:

$$\max_{\hat{c}} \quad V_S(\hat{c}) = [1 - H(\hat{c})] \cdot \hat{c}. \tag{A.39}$$

Note that $V_S(\underline{c}) = [1 - H(\underline{c})] \cdot \underline{c} > 0$ and $V_S(\bar{c}) = 0$, and V_S is continuous, so an interior maximum exists. The first-order condition is:

$$\frac{dV_S}{d\hat{c}} = [1 - H(\hat{c})] - \hat{c} \cdot h(\hat{c}) = 0, \tag{A.40}$$

$1 - H(\hat{c}^*) = \hat{c}^* \cdot h(\hat{c}^*)$ which yields condition (5.4).

Uniqueness. Rewrite the FOC as:

$$\frac{1 - H(\hat{c})}{h(\hat{c})} = \hat{c}. \tag{A.41}$$

The left-hand side is the inverse hazard rate, which is strictly decreasing under the increasing hazard rate (IHR) assumption. The right-hand side is strictly increasing. Hence, there is at most one intersection in $(\underline{c}, \bar{c})$. Existence follows from the intermediate value theorem: at $\hat{c} = \underline{c}$, the LHS equals $(1 - H(\underline{c}))/h(\underline{c}) > 0 = \underline{c}$ only if $\underline{c}$ is small enough, and at $\hat{c} = \bar{c}$, the LHS equals $0 < \bar{c}$. Since V_S achieves a positive maximum on a compact set, the interior solution exists.

Second-order condition.

$$\frac{d^2V_S}{d\hat{c}^2} = -2h(\hat{c}) - \hat{c} \cdot h'(\hat{c}). \tag{A.42}$$

Under IHR with log-concave density ($h' \geq -h^2/(1 - H)$, which log-concave distributions satisfy), this is strictly negative at the critical point. Hence, $\hat{c}^*$ is a strict global maximum. □

A.8 Proof of Proposition 5.2

Proof. (1) The characterizing equation (5.4) involves only H and h. Neither τ nor any other model parameter appears. (2) $k^* = (1 - \hat{c}^*)/\tau$, so $\partial k^*/\partial\tau = -(1 - \hat{c}^*)/\tau^2 < 0$ since $\hat{c}^* < 1$. (3) The state's ex ante payoff from delegation, anticipating the optimal post-success offer, is:

$$\begin{aligned} U_S^P(\theta_L; k^*) &= (1 - p_P) \underbrace{[(1 - \tau) - c_P]}_{\text{attack fails}} + p_P \underbrace{[V_S^* - c_P]}_{\text{attack succeeds}} \\ &= (1 - p_P)(1 - \tau) + p_P \cdot V_S^* - c_P. \end{aligned} \tag{A.43}$$

Since $\hat{c}^*$ depends only on H, V_S^* is a constant with respect to all parameters except H. □

A.9 Proof of Proposition 5.3

Proof. **Part 1.** The weak state's payoff from delegation exceeds its payoff from inaction when $\mathcal{V} \geq 0$. Combined with the baseline conditions ensuring neither type deviates between D and P, and the participation constraint (5.11) for θ_H, the active separating equilibrium is sustained.

Part 2. Consider candidate profile $(\theta_H \to D,\ \theta_L \to \varnothing)$ with off-path belief $\mu(P, S) = 0$.

Beliefs: On path, only θ_H plays D and succeeds with certainty, so $\mu(D, S) = 1$. All other signals are off-path.

θ_H does not deviate:

- To $\varnothing$: $1 - c_D \geq 1 - \tau$ holds by assumption (5.11).
- To P: Under the off-path belief $\mu(P, S) = 0$, the deviation payoff is $(1 - \tau) + \mathcal{V}$. Since $\mathcal{V} < 0$, this is strictly less than $(1 - \tau) \leq 1 - c_D$.

θ_L does not deviate:

- To D: $(1 - \tau) \geq (1 - \tau) - c_D + \alpha\tau$ holds since $c_D \geq \alpha\tau$.
- To P: Under $\mu(P, S) = 0$, the deviation payoff is $(1 - \tau) + \mathcal{V} < (1 - \tau)$ since $\mathcal{V} < 0$.

Note that the off-path belief $\mu(P, S) = 0$ is not uniquely necessary to support this equilibrium. Any off-path belief $\tilde{\mu} \leq \gamma$ yields the same deviation payoff, since the proxy's activation threshold $\hat{c} = 1 - k\tau$ and hence the blowback probability π depend only on whether $\tilde{\mu} \leq \gamma$ (blowback region) or $\tilde{\mu} > \gamma$ (no blowback), not on the precise value of $\tilde{\mu}$. The equilibrium is therefore supported by any off-path belief in $[0, \gamma]$. □

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.

Alesina, A. (1988). Credibility and policy convergence in a two-party system with rational voters. *The American Economic Review*, 78(4):796–805.

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.

Baliga, S. and Sjöström, T. (2012). The strategy of manipulating conflict. *American Economic Review*, 102(6):2897–2922.

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 Coate, S. (1997). An economic model of representative democracy. *The quarterly journal of economics*, 112(1):85–114.

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.

Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.

Downs, A. (1957). An economic theory of democracy. *Happer & Row*.

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.

Lizzeri, A. and Persico, N. (2001). The provision of public goods under alternative electoral incentives. *American Economic Review*, 91(1):225–239.

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.

Osborne, M. J. and Slivinski, A. (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics*, 111(1):65–96.

Persson, T., Roland, G., and Tabellini, G. (1997). Separation of powers and political accountability. *The Quarterly Journal of Economics*, 112(4):1163–1202.

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.