The Dictator's Balancing Act

Jack Paine*

April 28, 2018

Abstract

Elites and outsiders threaten dictators. How do autocrats manage these threats, and what are the consequences for regime survival? This paper analyzes a repeated interaction in which a dictator chooses whether or not to share power with another elite, and then the dictator and elite bargain over state revenues. In the baseline model without outsider threats, the dictator faces predatory incentives to exclude and only shares power if elites pose a strong rebellion threat—in which case institutional quality and threat capabilities are non-monotonically related to equilibrium coup propensity because of survival-based exclusion incentives. Introducing an outsider threat further encourages power-sharing by creating incentives to "hang together" to prevent "hanging separately." Although the hang-together effect may raise coup likelihood by creating a guardianship dilemma, it also implies that outsider threats may underpin inclusive and durable authoritarian regimes in weakly institutionalized environments.

*Assistant Professor, Department of Political Science, University of Rochester, jackpaine@rochester.edu. I thank Alex Debs, Tiberiu Dragu, and Milan Svolik for helpful comments on previous drafts.

Dictators face diverse survival threats. Most imminent, treasonous elites can attempt a coup d'etat to overthrow the ruler. This risk is indeed omnipresent in authoritarian regimes: successful coups accounted for 68% of nonconstitutional leadership removals in authoritarian regimes between 1945 and 2002 (Svolik, 2009, 478). One possible strategy to counteract coup attempts is to narrow the ruling coalition by excluding threatening elites from power. However, discontent elites often violently rebel against narrowly constructed regimes, as Roessler (2016), Roessler and Ohls (2018), and Cederman et al. (2013) demonstrate for ethnic civil wars, and Goodwin and Skocpol (1989) for social revolutions. These dual elite threats create a strategic coup-civil war tradeoff that affects dictators' power-sharing decisions.

Non-elite outsiders also threaten dictators' survival. In cases ranging from the Hittites in Ancient Egypt to mass demonstrators in Tahrir Square in 2011, and from Mongols in pre-modern China to communists in 1949, outsiders have directly participated in overthrowing authoritarian rulers. Furthermore, outsider threats also exert indirect effects that alter the elite power-sharing calculus. On the one hand, outsider threats may hasten regime overthrow by creating a "guardianship dilemma" (Finer, 2002; Acemoglu, Vindigni and Ticchi, 2010; Besley and Robinson, 2010; Svolik, 2013). Rulers often choose to build bigger militaries and to create more inclusive elite coalitions in reaction to outsider threats, which enables insiders to overthrow the ruler via a coup d'ètat, as in Egypt in 2011 or in Iraq prior to the rise of Saddam Hussein. On the other hand, outsider threats may *facilitate* power-sharing deals by causing a dictator and elites to "hang together" to prevent "hanging separately," thus promoting inclusive coalitions with low coup risk (McMahon and Slantchev, 2015). For example, in pre-1994 South Africa, British and Dutch descendants formed a ruling coalition among whites—despite stark regional differences that triggered civil war in the late colonial period—to counter the perceived African threat from below (Lieberman, 2003).

How do autocrats manage these threats from elites and outsiders, and what are the consequences for regime survival? Existing research has mostly examined the strategic coup-civil war tradeoff, the guardianship dilemma, and the hang-together mechanism in isolation. However, we lack a unified framework for understanding how dictators manage this omnipresent yet delicate balancing act among elites and outsiders.

This paper studies these contending strategic considerations by analyzing a game theoretic model. The baseline model studies an interaction between a dictator and elite faction without an outsider threat. In every period of an infinite-horizon game, the dictator decides whether to share power at the center with elites (include) or not (exclude), followed by a bargaining interaction in which the elite faction can accept a proposed division of state revenues or fight. The dictator has imperfect ability to commit to current-period bargaining offers, and the fighting technology for an included elite is denoted as a "coup" and for an excluded elite as a "rebellion." The key power-sharing tradeoff is that although sharing power raises the likelihood that the dictator can commit to its bargaining offer, power-sharing also enables elites to attempt a coup—which are assumed to succeed at higher rates than if elites are excluded from power and rebel. The unique stationary equilibrium in which the dictator shares power in every period features zero probability of a coup attempt in equilibrium if commitment ability under inclusion is high enough, and a mixed strategy for the elite with a strictly positive probability of coup attempts if commitment ability is low.

The baseline model without outsider threats shows that the dictator faces predatory incentives to exclude and only shares power if elites pose a strong rebellion threat—in which case institutional quality and threat capabilities are non-monotonically related to equilibrium coup propensity because of survival-based exclusion incentives. Specifically, the elite's rebellion threat is necessary to compel power-sharing. The dictator faces *predatory* incentives to exclude the elite and to accrue larger rents. Even if coup attempts occur with zero probability under inclusion, the dictator will still refuse to share power—therefore raising the equilibrium probability of overthrow—if the probability of a rebellion under exclusion is low. If instead the elite poses a relatively strong rebellion threat, then the dictator's ability to commit to deals (i.e., institutional quality) relates non-monotonically to the equilibrium probability of a coup attempt. If institutions are very weak, then the dictator excludes the elite for *political survival* reasons to prevent coup attempts. Medium-strength institutions induce the dictator to share power to prevent a rebellion, but the equilibrium probability of a coup attempt under inclusion is positive unless institutional strength reaches a high level. Finally, although relatively high threat capability is necessary to eliminate the predatory motive for exclusion, the severity of the elites' threat capability only conditionally affects prospects for power-sharing because threatening traits that enable coups to succeed relative to rebellions foster survival-based exclusion.

The model then incorporates an outsider threat. In every period, there is a positive probability that a (non-strategic) outsider will gain political power in the next period and that the two strategic players will consume 0 in all future periods, but peaceful power-sharing at the center decreases this probability. A more severe outsider threat encourages power-sharing by creating incentives to "hang together" to prevent "hanging separately" because neither the dictator nor want to trigger outsider takeover. The hang-together effect generates two main consequences. First, more severe outsider threats ambiguously affect coup propensity because the hang-together effect can compel power-sharing arrangements that yield a positive probability of a coup threat—the standard guardianship dilemma—even though, conditional on sharing power, the hang-together effect decreases coup propensity. Only in cases where power-sharing is self-enforcing absent an outsider threat—a consequence of the strategic coup-civil war tradeoff from the baseline model—do stronger outsider threats unambiguously decrease coup risk.

Second, despite the ambiguous coup effect, the hang-together effect can contribute to stable power-sharing and durable authoritarian regimes in weakly institutionalized environments. Although the direct effect of stronger outsider threats increases the probability of overthrow, indirect effects that cause the dictator and elites to band together can decrease the overall probability of overthrow (i.e., by either the elite or the outsider) relative to a counterfactual scenario without an outsider threat.

1 Contributions to Existing Research

1.1 Power-Sharing in the Shadow of Rebellion

This paper contributes to three main research agendas. First, the baseline model contributes to a large comparative politics literature on causes of civil war and revolution. Research on social revolutions argues that exclusionary, personalist regimes often leave society "no other way out" than revolution to challenge the regime (Goodwin and Skocpol, 1989; Goodwin, 2001; Snyder, 1998). Cederman et al. (2013) provide systematic evidence from a large sample of ethnic groups since 1945 that ethnic groups that lack access to power in the central government are more likely to initiate ethnic civil wars. However, most existing analyses do not scrutinize strategic government behavior. Why do governments pursue exclusionary policies if this triggers costly societal violence? The present model provides insight into this key question.

The baseline model most closely relates to Philip Roessler's research on the coup-civil tradeoff in Africa. Roessler (2011, 2016) frames the general tradeoff, arguing that dictators in newly independent African countries often feared members of other ethnic groups because European generals could no longer provide security guarantees for the incumbent. This underlying weakness in state institutions created exclusionary incentives to prevent the more imminent threat of a coup, often at the cost of facing armed rebellions. However, despite these challenges, some African countries have constructed durable power-sharing arrangements and have avoided civil war, such as Benin and Ghana. Roessler and Ohls (2018) argue that in countries where multiple ethnic groups have high threat capability (specifically, large size and close proximity to capital city) that the mutual threat of an excluded group rebelling makes power-sharing self-enforcing, despite weak state institutions.

However, important questions remain. First, although the insight that *high* threat capacity can facilitate power-sharing is intriguing, formally modeling the interaction can help to explicate its logical foundations. If coups pose a more imminent threat than outsider rebellions, then why does higher threat capability necessitate inclusion rather than exclusion? Second, the research agenda on ethnic civil wars and social revolutions has remained largely separate from the formal literature on outsider threats and coup risks. The present model shows how an outsider that threatens the strategic dictator and elite actors changes the power-sharing calculus. Whereas Roessler and Ohls (2018, 11) argue for weak states that power-sharing only occurs when actors have mutual threat capabilities and that these regimes necessarily exhibit high coup risk, the present analysis shows how outsider threats can foster power-sharing absent mutual threat capabilities and diminish coup risk.

1.2 Outsider Threats and Coup Risk

This paper also contributes to formal theoretic research that examines how outsider threats affect prospects for military coups. Acemoglu, Vindigni and Ticchi (2010), Acemoglu, Ticchi and Vindigni (2010), Besley and Robinson (2010), Leon (2013), and Svolik (2013) present different ways of modeling the classic "guardianship dilemma."[1] A regime that faces outsider threats needs a military to combat the threat, but increasing the size of the military increases its ability to stage a coup and overthrow the regime. This forces the government to trade off among loyalty, efficiency, and cost (Finer, 1997, 18).[2] To ensure loyalty, the government can build a small military that does not pose a coup threat, although this renders the government relatively feeble against the outsider threat. To improve efficiency, the government can build a larger military better able to deal with the threat, but this imposes a direct cost of resources needed for the military and an indirect cost from increasing the military's ability to stage a coup—which may occur in equilibrium in these

[1]Related contributions analyze other aspects of agency problems in dictatorships (Debs, 2016; Dragu and Przeworski, 2017; Gehlbach and Keefer, 2011; Tyson, 2017).

[2]The guardianship dilemma is widely discussed in other non-formal research as well, such as Feaver (1999) and Huntington (1957).

models either because of commitment problems or incomplete contracting. Finally, the government may be able to simultaneously counteract the outsider threat and solve its commitment or contracting problem by building a very large army (Acemoglu, Ticchi and Vindigni, 2010; Svolik, 2013), but this is very expensive and may undermine the ability of the civilian regime to pursue desired policies.

McMahon and Slantchev (2015), by contrast, argue that the core logic behind the guardianship dilemma is flawed. In previous models and non-formal conceptualizations of the guardianship dilemma, the severity of the outsider threat is assumed to not affect the military's expected utility of attempting a coup. McMahon and Slantchev (2015) show that a military is only willing to pay the cost associated with staging a coup if its prospects of surviving in power are sufficiently high conditional on winning. Therefore, although stronger outsider threats induce the government to create a larger and more capable military, larger threats *decrease* the expected utility of attempting a coup even if coup attempts are likely to succeed.

However, this debate has yet to address two important questions to which the present model provides insight, conceptualizing the elite challenger as similar to the military in these models. First, although outsider threats may deter military coups conditional on having a military (i.e., sharing power with the elite), outsider threats also may compel sharing power with the elite in circumstances in which the ruler would otherwise exclude and therefore eliminate its coup threat—the very premise of the guardianship dilemma. It is unclear which effect will dominate under which circumstances.

Second, existing models assume that a dictator can eliminate threats by other elites simply by choosing not to build a military (whatever the consequences of that choice for dealing with an outsider threat). However, this leaves open the question of whether the effect of an outsider threat is the same when imposing the more substantively relevant assumption that no dictator exists in a vacuum and therefore *always* faces a threat of overthrow by other elites. The present model evaluates outsider threats in a model featuring a permanent threat by other elites, therefore more directly considering how dictators balance elite and outsider challenges. Whereas in existing models the government will never devote resources to the military absent an external threat, the current model elucidates conditions under which the dictator shares power even with no external threat present.

1.3 Additional Related Formal Literature

The analysis also relates to other strands of the formal literature. It draws from and extends formal bargaining models of war that focus on commitment problems (Fearon, 2004; Powell, 2004; Acemoglu and Robinson, 2006; Krainin, 2017). This literature provides the key theoretical insight that costly fighting can occur in equilibrium if the government's ability to credibly commit to promises is low. To capture a tradeoff between insider coups and external rebellions, the present model allows the government to choose between two institutional settings in which to conduct bargaining, whereas commitment ability is exogenous in most existing formal work. Thus, whereas low commitment ability is often referred to as a commitment *problem* in existing research, the present model shows why the government may not deem expected low commitment ability toward the challenger, which occurs under exclusion, as problematic—considering the alternative of sharing power with the challenger and increasing its probability of winning. The flip-side to this consideration is that the elite actor's inability to commit not to exploit its stronger bargaining leverage if included (or to attempt a coup) drives the dictator's exclusion incentives.

The recursive dynamic structure of the present model also exhibits intriguing theoretical properties, such as the possibility of the challenger mixing in equilibrium between accepting and fighting—despite the game having complete and perfect information. Kalandrakis (2016), Gibilisco (2017), and Acemoglu and Robinson (2017) show how mixing can arise in related bargaining frameworks.

Finally, Francois et al. (2015) also formally examine dictators' power-sharing challenge and produce the intriguing theoretical result and supporting empirical evidence—relative to existing characterizations of "big man" politics in Africa—that rulers tend to allocate cabinet portfolios in proportion to ethnic group size to prevent coups and civil war. They develop a rich coalitional structure with N players to generate this implication. By contrast, the main regions of the parameter space examined here are precisely those in which the ruler faces a *tradeoff* between facing coups and rebellions, and the key results evaluate comparative statics for parameters that affect this tradeoff, including institutional quality, threat capabilities, and outsider threats. Furthermore, in other models that indirectly consider aspects of power-sharing, the set of actions that relate to this choice do not capture the key tradeoffs studied here. For example, in Acemoglu and Robinson (2006), elites can respond to mass threats by not sharing any power (temporary concessions or repression) or by handing the keys to the executive office to the masses (democratization), and do not consider intermediate power-sharing arrangements.

2 Baseline Model and Bargaining Analysis

This section presents a baseline model in which a dictator and elite interact without an outsider threat. After presenting the setup and equilibrium concept, it characterizes their bargaining interaction. The next section fully characterizes the equilibrium of the baseline model by analyzing the dictator's optimal power-sharing choice, followed by analyzing the consequences of outsider threats.

2.1 Setup

Two players interact in an infinitely repeated game with time denoted as $t \in \mathbb{Z}_+$ and future consumption discounted exponentially by a common factor $\delta \in (0, 1)$. The players are symmetric except with regard to which one controls the government at the onset of any period, yielding labels of dictator D_t and elite E_t. In each period, D_t first makes a power-sharing choice, $\mu_t \in [0, 1]$, where $\mu_t = 1$ denotes a pure strategy with inclusive power-sharing at the center, $\mu_t = 0$ denotes a pure exclusion strategy, and $\mu_t \in (0, 1)$ denotes mixing. To provide examples of power-sharing, the two major clusters of ethnic groups in Uganda, Baganda and northern groups, shared power at independence.[3] The prime minister and most army officers were northerners and the president and most cabinet positions were held by Baganda. For more than a decade after 1966, however, Baganda were excluded from power after the northern president arrested their cabinet ministers and assumed the title of president for himself. In a non-ethnic example, power was tightly concentrated around Fulgencio Batista and a small cadre of military officers prior to the Cuban Revolution, excluding other elites (large landowners and businesspeople) from positions of power.[4]

After the power-sharing choice, D_t offers E_t a share of exogenously generated per-period government spoils, $x_t \in [0, 1]$. E_t responds to D_t's offer by accepting with probability $\alpha_t \in [0, 1]$. If E_t accepts,[5]

[3] The codebook for Cederman et al.'s (2013) dataset provides relevant citations.

[4] In reality, historical factors affect a dictator's ability to incorporate or to exclude other social groups from power. However, for the purposes of understanding a dictator's *strategy*, assuming that D_t can freely choose inclusion or exclusion in every period is more illuminating.

[5] That is, either by deterministic choice or by the outcome of the Nature draw following a mixed acceptance choice.

then it only probabilistically receives the promised transfer because D_t cannot perfectly commit to offers, as formalized below. Following acceptance, D_t remains dictator in the next period and E_t remains as the elite, i.e., $D_{t+1} = D_t$ and $E_{t+1} = E_t$. Alternatively, E_t can fight in response to D_t's offer. If E_t fights, then neither player consumes in the current period but there are no permanent costs of fighting. If E_t wins a fight in period t, then it becomes the dictator at time $t+1$ and D_t becomes the elite, i.e., $D_{t+1} = E_t$ and $E_{t+1} = D_t$. If instead D_t survives the fight, then both players retain their positions in the next period: $D_{t+1} = D_t$ and $E_{t+1} = E_t$. If E_t is included in power at time t, then its probability of winning equals p_i and its fight is a "coup." If E_t is excluded from power at time t, then its probability of winning equals p_e and its fight is a "rebellion."

The following assumptions generate D_t's power-sharing dilemma. The benefit of sharing power with E_t is to improve D_t's (expected) ability to commit to a deal. If E_t is included, then with probability $\theta_t = \theta_i \in (0,1]$, the deal goes through and D_t and E_t consume their respective shares $1 - x_t$ and x_t. With complementary probability $1 - \theta_i$, Nature allows D_t to renege on the deal, in which case D_t and E_t respectively consume 1 and 0. If instead E_t is excluded, then θ_t is lower in expectation. Formally, $\theta_t\big|_{\mu_t = \mu^e} = \theta_i$ with probability $\sigma \in (0,1)$ and $\theta_t\big|_{\mu_t = \mu^e} = 0$ with probability $1 - \sigma$.[6]

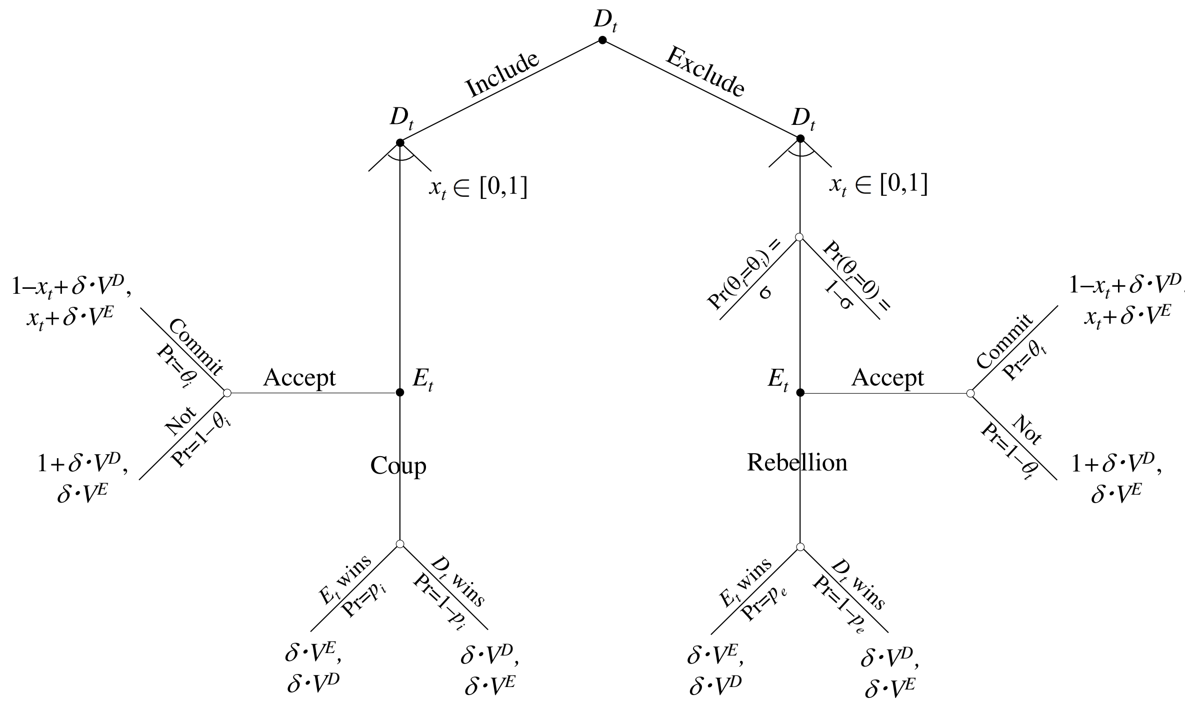
However, the cost of sharing power is that E_t's probability of succeeding at a coup exceeds its probability of winning a rebellion: $0 < p_e < p_i \leq 1$. Roessler (2016, 37) distinguishes between coup conspirators' partial control of the state and insurgents' lack of such power access and need to build a private military organization. "This organizational distinction helps to account for why coups are often much more likely to displace rulers from power than rebellions."

Finally, from the perspective of any period t, V^D denotes the future continuation value for the period t dictator and V^E the continuation value for the period t elite.[7] Figure 1 presents the tree of the stage game.

[6]Because the Nature draw from the Bernoulli distribution summarized by σ is not realized until after D_t's power-sharing decision, D_t is uncertain what its commitment ability toward E_t will be if E_t is excluded. This captures a general inability for a dictator to anticipate the bargaining environment if the other group is distant from power at the center, whereas E_t observes the contemporaneous commitment probability before accepting or fighting.

[7]These continuation values could be written as a function of the history of play, but below I restrict attention to stationary strategies.

Figure 1: Tree of Baseline Stage Game



Notably, although none of the variables in the baseline game evolve over time, the dynamic recursive structure of the game affects the analysis by endogenizing the value of holding office. D_t and E_t must consider the likelihood of future challenges when making their respective power-sharing and fighting choices in each period, as opposed to fixing an exogenous value to remaining as or to becoming ruler. One consequence of the endogenizing the value of holding office is to generate E_t's equilibrium mixing strategy, described below. Although many conflict bargaining models exhibit exogenous shifts in the distribution of power over time to generate commitment problems, given the present setup, adding that feature would complicate the model without adding important theoretical insights.

2.2 Equilibrium Concept

The paper analyzes the conditions under which there exists a (stationary) Markov perfect equilibrium, henceforth simply "stationary equilibrium," in which D_t includes E_t with probability 1 in every period. The only difference in payoff-relevant states depends on which actor controls office at the start of the period. Formally, a "power-sharing equilibrium" entails $\mu_t = 1$ for all t, and the analysis solves backwards on the stage game for the conditions under which a power-sharing equilibrium exists. The focus on stationary strategies ensures that, for any period t, optimal actions do not depend on actions in periods prior to t, and therefore the subscript t will usually be dropped from D_t and E_t to minimize notation. The appendix proves all formal statements, and choice variables accompanied by asterisks are equilibrium choices.[8]

2.3 Bargaining Analysis

The analysis begins by assessing the bargaining interaction after D has made its power-sharing choice in period t. If a power-sharing equilibrium exists, then the respective continuation values for the dictator and elite equal:

$$V^D = \alpha_i^* \cdot \Big[1 - \theta_i \cdot x_i^* + \delta \cdot V^D\Big] + (1-\alpha_i^*) \cdot \delta \cdot \Big[p_i \cdot V^E + (1-p_i) \cdot V^D\Big] \tag{1}$$

$$V^E = \alpha_i^* \cdot \Big[\theta_i \cdot x_i^* + \delta \cdot V^E\Big] + (1-\alpha_i^*) \cdot \delta \cdot \Big[p_i \cdot V^D + (1-p_i) \cdot V^E\Big] \tag{2}$$

If a power-sharing equilibrium exists, then D's offer in every period is denoted as x_i^* and E's probability of acceptance as α_i^*, where the subscript "i" denotes *inclusion*. If E accepts the equilibrium offer in period t, which occurs with probability α_i^*, then D's expected consumption in period t is $1 - \theta_i \cdot x_i^*$ and E's is $\theta_i \cdot x_i^*$, and in period $t+1$ both players retain their period t positions with probability 1. If instead E attempts a coup, which occurs with probability $1 - \alpha_i^*$, then neither player consumes in period t. In period $t+1$, the players switch positions with probability p_i and retain their period t positions with complementary probability.

To solve for E's acceptance constraint, if E accepts an offer x_t, then its expected lifetime consumption equals $\theta_i \cdot x_t + \delta \cdot V^E$. By contrast, fighting yields lifetime expected utility $\delta \cdot \Big[p_i \cdot V^D + (1-p_i) \cdot V^E\Big]$. Therefore, E will only accept with positive probability offers for which $\theta_i \cdot x_t \geq \delta \cdot p_i \cdot \big(V^D - V^E\big)$ and, in equilibrium, D will make E is indifferent between accepting and fighting. Therefore, the optimal offer and

[8]Future drafts will characterize the full range of equilibrium possibilities, such D excluding E in every period or D mixing between inclusion and exclusion. Note that in parameter ranges where a power-sharing equilibrium exists, it may not be the unique stationary equilibrium.

probability of acceptance terms satisfy:[9]

$$\Omega\big(x_i^*, \alpha_i^*\big) \equiv \underbrace{\theta_i \cdot x_i^*}_{\text{Current opp. cost}} \underbrace{- \delta \cdot p_i \cdot \overbrace{\overbrace{\frac{1-\delta}{1-\delta\cdot(1-2\cdot p_i)}}^{\in(0,1)} \cdot \overbrace{\frac{\alpha_i^*}{1-\delta}}^{=V^D+V^E}}^{V^D-V^E}}_{\text{Expected future gains from coup}} = 0 \tag{3}$$

E's incentives to possibly fight arise because the governing actor consumes more than the elite: $V^D > V^E$. E consumes a lower percentage of the joint surplus in equilibrium because (1) E's bid to take the government may fail ($p_i < 1$) and (2) even conditional on winning, E does not enjoy the fruits of governing until the next period ($\delta > 0$). The total surplus $V^D + V^E$ equals the present discounted value of the infinite per-period revenue stream, $\frac{1}{1-\delta}$, multiplied by the equilibrium percentage of periods in which consumption occurs, α_i^*.

There are two possible bargaining outcomes in an inclusionary strategy profile.

Peaceful bargaining. If θ_i is high enough, then $\big(x_i^*, \alpha_i^*\big) = \big(\tilde{x}_i^*, 1\big)$, for $\Omega\big(\tilde{x}_i^*, 1\big) = 0$, which yields the explicit solution:

$$\tilde{x}_i^* = \frac{1}{\theta_i} \cdot \frac{\delta \cdot p_i}{1-\delta\cdot(1-2p_i)} \tag{4}$$

This offer and acceptance probability follow from standard results in conflict bargaining models: D will induce acceptance with probability 1 if possible because fighting is costly and, by virtue of making the bargaining offers, D pockets the surplus saved from not fighting. The costliness of fighting in this model follows from no consumption in a fighting period.

Probabilistic coups. If instead θ_i is low, then E will fight with positive probability in equilibrium because D has low ability to commit to its bargaining offers, which Lemma 1 formalizes.

Lemma 1 (Probabilistic coup threshold)**.** *There does not exist an offer that E will accept with probability 1 if and only if:*

$$\theta_i < \overline{\theta} \equiv \frac{\delta \cdot p_i}{1-\delta\cdot(1-2p_i)} \in (0,1)$$

[9] Solving for Equation 3 requires substituting $\theta_i \cdot x_i^* = \delta \cdot p_i \cdot \big(V^D - V^E\big)$ into Equations 1 and 2 to solve for $V^D - V^E$ as a function of α_i^*.

The intuition is that if θ_i is low, then E is sufficiently unlikely to receive the bargaining offer that it can profit from attempting a coup with positive probability. Even if D offers the entire per-period revenue stream of 1, E's expected consumption if it accepts the offer is low. Therefore, if $\theta_i < \overline{\theta}$, then $\alpha_i^* = 1$ is not possible.

However, even if $\theta_i < \overline{\theta}$, E cannot accept $x_t = 1$ with probability 0. To see why, if this was possible, then fighting would occur in every period and no player would ever consume. This generates a contradiction because E would strictly prefer to deviate to accept $x_t = 1$, yielding lifetime expected consumption of θ_i rather than 0 if it attempts a coup (because the strategy profile states that the elite will attempt a coup in all future periods). Formally, Equation 3 shows that $\Omega\big(x_i^* = 1, \alpha_i^* = 0\big) = \theta_i > 0$.

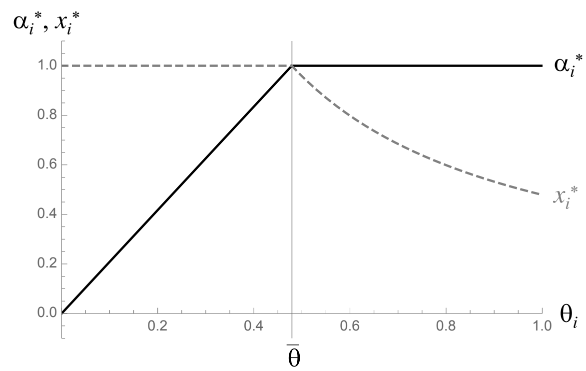
Instead, if $\theta_i < \overline{\theta}$, then E accepts $x_t = 1$ with a unique probability strictly bounded between 0 and 1. From the perspective of E_t, knowing that future elites will mix is necessary is necessary to generate indifference between accepting $x_t = 1$ and fighting to possibly become dictator in period $t+1$. Equation 3 shows that the anticipated future gains of fighting strictly increase in the equilibrium probability of acceptance. Higher α_i^* increases lifetime joint consumption, which creates stronger incentives for E to fight to become the player that consumes relatively more of this larger surplus. If θ_i is small enough that Equation 3 cannot be satisfied with $\alpha_i^* = 1$ (because $x_t \leq 1$), then a lower value of α_i^* will satisfy the equation.

These considerations highlight a subtle and counterintuitive tradeoff. A higher probability of acceptance in future periods increases E's expected utility to fighting in period t, which narrows the range of offers that E will accept in period t. Therefore, choices by *future* elites must be calibrated to make the *current* elite indifferent between accepting or attempting a coup. Formally, if $\theta_i < \overline{\theta}$, then $\big(x_i^*, \alpha_i^*\big) = \big(1, \tilde{\alpha}_i^*\big)$, for $\Omega\big(1, \tilde{\alpha}_i^*\big) = 0$, which yields the explicit solution:

$$\tilde{\alpha}_i^* = \frac{1 - \delta \cdot (1 - 2p_i)}{\delta \cdot p_i} \cdot \theta_i \tag{5}$$

Figure 2 plots equilibrium bargaining outcomes as a function of θ_i. The equilibrium acceptance probability is the solid black line and the equilibrium patronage offer is the dashed gray line. If $\theta_i > \overline{\theta}$, then the patronage offer is interior and accepted with probability 1. Furthermore, the offer strictly decreases in θ_i, and equals 1 at $\theta_i = \overline{\theta}$. If $\theta_i < \overline{\theta}$, then the patronage offer is 1 and the probability of acceptance is interior. The probability of acceptance strictly increases in θ_i, achieves its lower bound of 0 at $\theta_i = 0$, and its upper bound of 1 at $\theta_i = \overline{\theta}$.

Figure 2: Equilibrium Bargaining Outcomes Under Power-Sharing



Notes: Figure 2 uses parameter values $\delta = 0.95$ and $p_i = 0.6$.

Figure 2 also shows how x_t and α_t act as substitutes in equilibrium. As θ_i decreases, at least one of these two choice variables must change in equilibrium to preserve Equation 3. The costliness of fighting implies that D will always make a high enough offer to induce acceptance with probability 1 if possible, and only if θ_i is low enough to make that impossible does the equilibrium acceptance probability drop below 1. Lemma 2 formally presents the unique bargaining outcome in a power-sharing equilibrium.

Lemma 2 (Bargaining under power-sharing)**.** *Assume a strategy profile in which D shares power with E in every period, and D follows the strategy profile in period t. If D's commitment ability is sufficiently high, then D chooses the unique x_t that satisfies with equality E's acceptance constraint. E accepts any offer at least that large with probability 1, and any lower offer with probability 0. In equilibrium, E accepts with probability 1. If D's commitment ability is lower, then D offers $x_t = 1$. E accepts $x_t = 1$ with probability strictly bounded between 0 and 1, and accepts any $x_t < 1$ with probability 0. In equilibrium, E accepts with probability strictly bounded between 0 and 1. Formally:*

- *If $\theta_i > \overline{\theta}$, then D offers $x_t = \tilde{x}_i^*$, for $\overline{\theta}$ defined in Lemma 1 and $\tilde{x}_i^*$ defined in Equation 4. E's equilibrium probability of acceptance function is:*

$$\alpha_i^*(x_t) = \begin{cases} 0 & \textit{if } x_t < \tilde{x}_i^* \\ 1 & \textit{if } x_t \geq \tilde{x}_i^*, \end{cases}$$

and the equilibrium acceptance probability is $\alpha_i^* = 1$.

- *If $\theta_i < \overline{\theta}$, then D offers $x_t = 1$. E's equilibrium probability of acceptance function is:*

$$\alpha_i^*(x_t) = \begin{cases} 0 & \textit{if } x_t < 1 \\ \tilde{\alpha}_i^* & \textit{if } x_t = 1, \end{cases}$$

for $\tilde{\alpha}_i^* \in (0, 1)$ defined in Equation 5, and the equilibrium acceptance probability is $\alpha_i^* = \tilde{\alpha}_i^*$.

Deviation to exclusion. If instead D_t deviates to exclude E_t, then there is a $1 - \sigma$ percent chance that a rebellion will occur because D cannot commit to deliver any spoils to E. With complementary probability, D has positive commitment ability and can induce E to accept with probability 1.[10] Appendix Lemmas A.1 and A.2 formally characterize bargaining with an excluded elite, and Remark 1 summarizes the most important implication.

> **Remark 1** (Frequency of rebellion under exclusion)**.** *When D makes its power-sharing decision, it knows that there is a $1 - \sigma$ percent chance that E will launch a rebellion if excluded.*

3 The Dictator's Balancing Act without Outsiders

In the baseline model without outsider threats, the dictator faces predatory incentives to exclude and only shares power if elites pose a strong rebellion threat—in which case institutional quality and threat capabilities are non-monotonically related to equilibrium coup propensity because of survival-based exclusion incentives.

3.1 Power-Sharing Incentive Compatibility Constraint

An equilibrium with permanent inclusion exists if and only if D cannot profitably deviate to excluding E. This occurs only if the equilibrium probability of a coup attempt under inclusion is sufficiently low relative to the equilibrium probability of a rebellion under exclusion. Comparing D's expected consumption under inclusion (see Equation 1) with the single deviation of excluding yields:

$$\alpha_i^* \cdot \Big[1 - \theta_i \cdot x_i^* + \delta \cdot V^D\Big] + (1 - \alpha_i^*) \cdot \delta \cdot \Big[p_i \cdot V^E + (1 - p_i) \cdot V^D\Big] \geq$$

$$\sigma \cdot \Big[1 - \theta_i \cdot x_e^* + \delta \cdot V^D\Big] + (1 - \sigma) \cdot \delta \cdot \Big[p_e \cdot V^E + (1 - p_e) \cdot V^D\Big],$$

for x_e^* defined in Appendix Equation A.8. Substituting in the equilibrium values of V^D and V^E and then re-arranging yields:

[10] Because bargaining under exclusion is a single deviation from the posited inclusionary strategy profile, accepting an offer under exclusion with probability 0 does not affect future consumption. Therefore, the future acceptance probabilities cannot be calibrated to generate an equilibrium mixed strategy for E_t under temporary exclusion—because, along the equilibrium path, no future-period elite actors will be excluded.

$$\overbrace{1-\alpha_i^*(\theta_i)}^{\text{Pr(coup | incl.)}} \leq \frac{1}{1-\kappa\cdot\underbrace{(p_i-p_e)}_{\text{Poli. surv. motive}}}\cdot\left[\overbrace{1-\sigma}^{\text{Pr(rebel | excl.)}} -\kappa\cdot\underbrace{(p_i-p_e)}_{\text{Predation motive}}\right], \tag{6}$$

for $\kappa \equiv \frac{\delta}{1-\delta\cdot(1-2p_i)} > 0$. This term expresses D's incentive compatibility constraint in terms of the equilibrium probability of a coup attempt under inclusion relative to the equilibrium probability of a rebellion under exclusion.

3.2 Predatory Exclusion

D faces two distinct types of incentives to exclude E. The first occurs for *predatory* reasons. Excluding E decreases its probability of winning a fight (i.e., rebellions succeed with lower probability than coup attempts) and therefore decreases the amount that E can credibly demand in the bargaining interaction. This creates a predatory incentive for exclusion by enabling D to accrue more rents. Equation 6 shows this effect formally: the larger is $p_i - p_e$, the lower is the range of parameter values in which D will share power.

Therefore, E's rebellion threat is necessary to compel D to share power in the baseline model. Even if the equilibrium probability of a coup attempt under inclusion is zero, Equation 6 shows that D will not share power if the probability of a rebellion under exclusion is sufficiently low because of the predation motive. Equation 6 also shows that $p_i > p_e$ is necessary and sufficient for the predation motive. Proposition 1 formally states this result in terms of E's ability to threaten D with rebellion if excluded.

> **Proposition 1** (Predatory exclusion)**.** *If the likelihood of a rebellion under exclusion is low enough, then D will exclude E. Formally, there exists a unique $\hat{\sigma} \in (0,1)$ such that if $1-\sigma < 1-\hat{\sigma}$, then there does not exist a stationary equilibrium in which $\mu^* = 1$ for any $\alpha^* \in [0,1]$.*

Notably, this predatory incentive to exclude exists even if $1-\sigma > 1-\hat{\sigma}$, although in that parameter range the predatory motive is not sufficient to ensure exclusion. Subsequent references to "predatory exclusion" refer to the case $1-\sigma < 1-\hat{\sigma}$, in which the predatory motives are sufficient for exclusion.

3.3 Survival-Based Exclusion

The second possible incentive to exclude E is for *political survival* reasons. Even if the likelihood of rebellion under exclusion exceeds $1-\hat{\sigma}$, D might still exclude E because it fears a coup attempt. Although

D will tolerate some coup risk to avoid a likely rebellion, it will deviate to exclusion if the equilibrium coup probability is high enough. Crucially, D fears coups more than it fears civil wars because $p_i > p_e$. Therefore, D will not share power if a coup attempt by an included E is equally likely as a rebellion by an excluded E. Overall, whereas a minimal rebellion threat induces predatory exclusion, a large enough coup threat induces survival-based exclusion.

Formally, if $1 - \sigma > 1 - \hat{\sigma}$, then the probability of a coup attempt under inclusion, $1 - \alpha_i^*$ (see Equation 5), must be sufficiently lower for D to share power. Proposition 2 characterizes a threshold $\underline{\alpha} \in (\hat{\sigma}, 1)$ such that D will share power if $1 - \alpha_i^* < 1 - \underline{\alpha}$ and exclude E otherwise because the risk of a coup is too great. Equation 6 demonstrates the political survival effect that D is less tolerant of coup attempts than rebellions, which follows from $\frac{1}{1-\kappa\cdot(p_i - p_e)} > 1$, and that $p_i > p_e$ is necessary and sufficient for this mechanism. Additionally, because the equilibrium probability of a coup strictly decreases in θ_i for $\theta_i \in \left(0, \overline{\theta}\right)$, the $\underline{\alpha}$ threshold can equivalently be stated in terms of a unique $\underline{\theta} \in \left(0, \overline{\theta}\right)$ threshold such that D shares power if $\theta > \underline{\theta}$ and not otherwise. There does not exist a power-sharing equilibrium for $\theta_i \to 0$ because the equilibrium probability of a coup attempt under inclusion goes to 1, as Equation 5 shows.

Proposition 2 (Survival-based exclusion)**.** *Even if the probability of a rebellion under exclusion is high enough that D will not exclude E for predatory reasons, if the probability of a coup under inclusion is high enough, then D will exclude E. Formally, suppose $1 - \sigma > 1 - \hat{\sigma}$, for $\hat{\sigma}$ defined in Proposition 1. Then:*

Part a. *There exists a unique $\underline{\alpha} \in (\hat{\sigma}, 1)$ such that there does not exist a stationary equilibrium in which $\mu^* = 1$ if $1 - \alpha_i^* > 1 - \underline{\alpha}$.*

Part b. *There exists a unique $\underline{\theta} \in \left(0, \overline{\theta}\right)$, for $\overline{\theta}$ defined in Lemma 1, such that:*

- *If $\theta_i < \underline{\theta}$, then $1 - \alpha_i^* > 1 - \underline{\alpha}$.*
- *If $\theta_i > \underline{\theta}$, then $1 - \alpha_i^* < 1 - \underline{\alpha}$.*

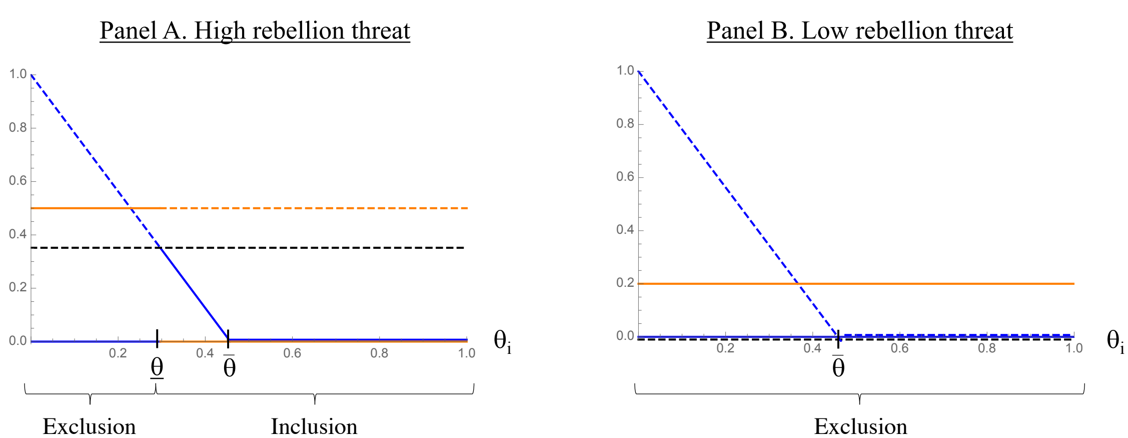
The predation and political survival motives for exclusion enable characterizing the conditions under which a stationary equilibrium exists in which D shares power with E in every period.

Proposition 3 (Power-sharing equilibrium)**.** *If and only if $1 - \sigma > 1 - \hat{\sigma}$ and $1 - \alpha_i^* < 1 - \underline{\alpha}$, for $\hat{\sigma}$ defined in Proposition 1 and $\underline{\alpha}$ defined in Proposition 2, then there exists a stationary equilibrium in which $\mu_t = 1$ for all t. Lemmas 2 and A.2 characterize optimal bargaining behavior.*

3.4 Effects of Institutional Quality

The remainder of this section examines comparative statics on two key parameters to understand prospects for power-sharing and for elite in-fighting (i.e., fighting between the dictator and other elites). Figure 3 plots the relationship between the dictator's commitment ability θ_i—which corresponds to institutional quality more generally—and equilibrium fighting probabilities, with coup attempts in blue and rebellions in orange. For parameter ranges in which a power-sharing equilibrium exists, solid lines depict expected outcomes premised on D_t choosing to include E_t and dashed lines depict expected outcomes premised on D_t's single deviation to excluding E_t. For parameter ranges in which a power-sharing equilibrium does not exist, solid lines depict expected outcomes premised on D_t's single deviation to excluding E_t and dashed lines depict expected outcomes premised on D_t choosing to include E_t.[11] The black curve depicts the highest possible coup probability that D will tolerate when contemplating whether to share power with E, $1 - \underline{\alpha}$. The two panels differ in the probability of rebellion under exclusion. In Panel A, the threat of rebellion is high: $1 - \sigma > 1 - \hat{\sigma}$. In Panel B, the threat of rebellion is low: $1 - \sigma < 1 - \hat{\sigma}$.

Figure 3: Commitment Ability, Power-Sharing, and Elite In-Fighting



Black: Survival-based exclusion threshold • Solid blue: Pr(coup) • Dashed blue: Pr(coup | included) •
Solid orange: Pr(rebellion) • Dashed orange: Pr(rebellion | excluded)

Notes: Both panels use the parameter values $\delta = 0.9$, $p_i = 0.6$, and $p_e = 0.3$. In Panel A, $\sigma = 0.5$. In Panel B, $\sigma = 0.8$. Under these parameter values, $\hat{\sigma} \approx 0.77$.

Panel A demonstrates an intriguing non-monotonic relationship between institutional quality and the equilibrium probability of a coup attempt, contrary to the common presumption that these variables should

[11] In all cases, future-period behavior is premised on power-sharing in every period.

exhibit a strictly decreasing relationship. If institutions are weak, then, there is *no* risk of a coup because the political survival incentive generates strategic substitution: D strategically excludes E because it fears coup attempts more than rebellions. The dashed blue line for $\theta_i > \underline{\theta}$ shows that the coup risk would be severe if D instead included E.

The equilibrium probability of a coup attempt exhibits a discrete upward jump at $\theta_i = \underline{\theta}$. At this point, the probability of a coup attempt by an included elite is low enough that D is willing to tolerate positive coup risk to prevent a more likely rebellion. The equilibrium probability of a coup attempt under inclusion must be lower than the equilibrium probability of a rebellion under exclusion (shown in orange) at the point where D shifts to including E because, again, E fears coup attempts more than rebellions.

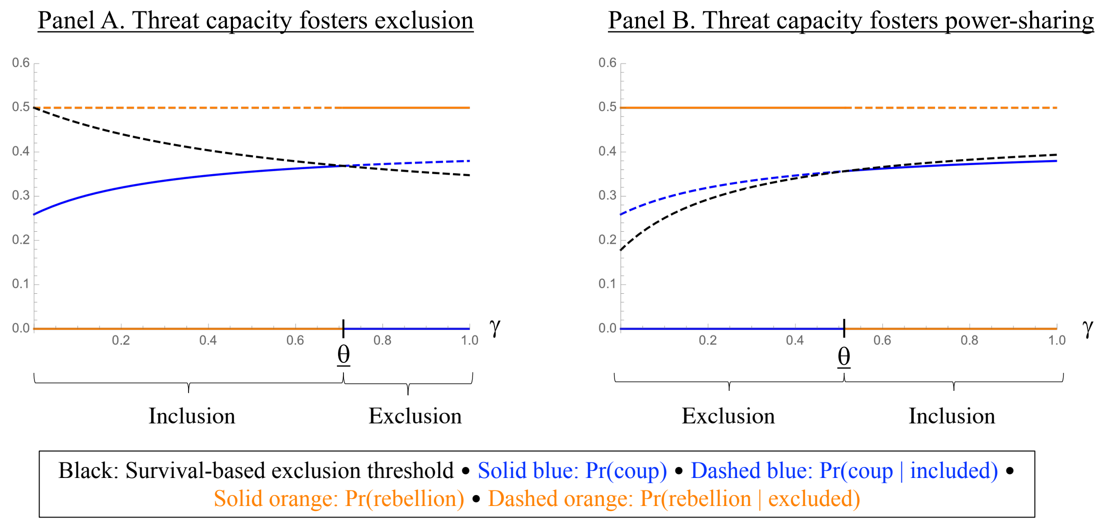
Only if institutional quality exhibits a larger increase, to at least $\theta_i = \overline{\theta}$, does the equilibrium probability of a coup attempt drop to 0. In this range, despite sharing power with E, D's promised concessions are sufficiently credible that E does not threaten a coup with positive probability.

Contrasting these findings with the patterns in Panel B demonstrates the necessity of a high rebellion threat for incentivizing D to share power with E. The only difference between Panels A and B is that, in Panel B, the equilibrium probability of rebellion under exclusion is sufficiently low to foster predatory exclusion. Even if θ_i is large—making the equilibrium probability of a coup attempt under inclusion low or even 0—the equilibrium probability of a rebellion under exclusion is low enough that D prefers to strengthen its bargaining position by excluding E. Intriguingly, for high enough θ_i, it is *possible* for D to drive down the equilibrium probability of a challenge by E (i.e., either a coup or rebellion) to 0 by sharing power. However, D instead chooses to exclude E to accrue the larger rents associated with E's diminished bargaining leverage under exclusion—despite creating a positive rebellion possibility.

3.5 Effects of Threat Capabilities

The model also enables studying the effects of threat capabilities on prospects for power-sharing and for elite in-fighting. The conventional wisdom is that a dictator will exclude more threatening elites from power because they fear coup attempts more than rebellions. However, Roessler and Ohls (2018) posit an alternative ethnic geography framework that yields the opposite conclusion. The present model provides a useful framework for evaluating these competing arguments. Suppose that p_i and p_e are each functions of underlying threat capability $\gamma \in [0, 1]$, specifically, $p_i = \beta_{0,i} + \beta_{1,i} \cdot \gamma$ and $p_e = \beta_{0,e} + \beta_{1,e} \cdot \gamma$.[12] Figure 4 shows that the effect of threat capability on prospects for power-sharing depends on whether γ more greatly affects E's ability to succeed at a coup or a rebellion.

Figure 4: Threat Capabilities and Power-Sharing



Notes: Both panels use the parameter values $\delta = 0.9$, $\theta_i = 0.29$, and $\sigma = 0.5$. In Panel A, $\beta_{0,i} = 0.2$, $\beta_{0,e} = 0.2$, $\beta_{1,i} = 0.6$, and $\beta_{1,e} = 0.2$. In Panel B, $\beta_{0,i} = 0.2$, $\beta_{0,e} = 0$, $\beta_{1,i} = 0.6$, and $\beta_{1,e} = 0.5$.

E's probability of succeeding in a coup or winning a rebellion strictly increases in γ in both panels. Related, the equilibrium probability of a coup attempt strictly increases in γ in both panels. The main difference between Panels A and B is the magnitude of the effect of threat capability on the probability that a rebellion succeeds. The slope is fairly flat in Panel A (0.2) but relatively steep in Panel B (0.5). Consequently, the effect of γ on D's tolerance to face coup attempts under inclusion differs in the two panels: higher γ strictly decreases this tolerance in Panel A but strictly raises it in Panel B. Consequently, D shares power in Panel A if and only if threat capacity is low, whereas D shares power in Panel B if and only if threat capacity is high.

Understanding which is these two scenarios is likely to hold in a given substantive setting depends on the nature of the attribute affecting threat capability. Horowitz's (1985) discussion of colonial military policies and "split domination" at independence likely corresponds to Panel A. Countries in which different ethnic groups controlled the military and civilian political institutions at independence exhibited split domination

[12] Assume the intercepts are weakly positive, the slopes are strictly positive, and the functions satisfy $p_i > p_e$ for all $\gamma \in [0, 1]$.

and high prospects for coup attempts and ethnic narrowing at the center. For example, Britain favored Karens in the military and bureaucracy in colonial Burma, but the ethnic majority Burmese controlled key political institutions after Britain regained control of the colony following World War II. Therefore, Karen officers exhibited high threat capacity, but more directly in the form of coup attempts given their favorable colonial position rather than a rebellion that could topple the regime, especially given their small size (7% of the population compared to 68% for Burmans).

By contrast, Panel B corresponds to Roessler and Ohls's (2018) conceptualization of high threat capabilities in the sense of large numerical size and close proximity to the capital. In particular for distance to the capital, it is quite plausible that this would affect a group's ability to rebel more than to stage a coup, which would cause high threat capacity to foster power-sharing rather than to trigger exclusion—despite the relatively high risk of a coup attempt among parameter values in which power-sharing occurs in Panel B. By disentangling scope conditions, the present model enables evaluating these heterogeneous aspects of threat capabilities and how they affect prospects for power-sharing.

4 The Dictator's Balancing Act with Outsider Threats

Introducing an outsider threat encourages power-sharing by creating incentives to "hang together" to prevent "hanging separately." Although the hang-together effect may raise coup likelihood by creating a guardianship dilemma, it also implies that outsider threats may underpin inclusive and durable authoritarian regimes in weakly institutionalized environments.

4.1 Setup

This section introduces an exogenous outsider threat to the baseline model. In any period t, if D excludes E and/or if E fights, then with probability $q \in (0, 1)$ a non-strategic outsider actor takes power in period $t + 1$. By contrast, the exogenous takeover probability is $\Delta \cdot q$ if E is included and accepts D's offer, for $\Delta \in (0, 1)$. If the outsider takes over following consumption in period t, then D and E each consume 0 in all periods after t. This captures that D and E prefer rule by either of them—in which each consumes a strictly positive amount in each period—over the outsider. Appendix Section B provides analogous formal results as those from the baseline model that include these additional terms.

This setup builds on two important ideas. First, actors in society can be differentiated based on whether they are "elites" or "outsiders." This could correspond to many different possible settings. One possibility is that D and E are co-ethnics and the exogenous threat is from a different ethnic group (for example, elite Malay factions and the Chinese-dominated communist movement in post-colonial Malaysia). Therefore, rather than *assuming* a dictator necessarily has the support of its co-ethnics as do other models of ethnic conflict (Padró-i Miquel, 2007), the present analysis characterizes conditions under which co-ethnics will choose to form cohesive regimes and to peacefully bargain in the shadow of an out-group threat. This distinction is also sensible in cases like pre-1994 South Africa, where there were two important elite European factions (British and Dutch descendants) and, from their perspective, a mass of African outsiders. Considering elites more broadly, Ansell and Samuels (2014) distinguish between agricultural and industrial elites and discuss their interaction with the masses. Yet another possibility is that D and E represent two elite families. Even in cases where one family is excluded from power (i.e., $\mu_t = 0$), it is still closer to power than outsiders such as peasants. For example, in 15th century England, the House of Lancaster and the House of York—which collectively composed the royal House of Plantagenet—contested for control of the government in the War of the Roses.

Of course, the elite-outsider distinction is not perfect in all substantive circumstances. For example, in cases of ethnic power-sharing discussed by Roessler (2011) and Cederman et al. (2013), the ruler strategically interacts with other ethnic groups and one possible outcome of their strategic action is to make the other group an "outsider." However, the broader idea that a more distant threat to the ruler's hold on power may affect its interaction with a different group is substantively relevant in a broad range of circumstances.[13]

The second consequential assumption is that disruptions at the center, and narrowly constructed regimes with minimal societal support, create openings for outsiders to control the government—whereas these openings are less likely to arise if groups at the center present a united front. For example, Goodwin (2001) argues that ruling elites who undermine their military and state capacity by coup-proofing their regimes create openings for revolutionary social movements (49). Snyder (1998, 56) claims that sultanistic regimes in Haiti, Nicaragua, and Romania successfully co-opted a broad range of societal elites for long periods and that the regimes fell to societal uprisings amid an "increase in the exclusion of political elites." Harkness

[13]How a ruler balances between two strategic factions that can form alliances with each other is also a relevant theoretical consideration, but poses many additional technical difficulties that obscure from the main takeaway points here, and therefore is left as an important consideration for future research.

(2016, 588) argues: "Compelling evidence exists that coups also ignite insurgencies by weakening the central government and thereby opening up opportunities for rebellion . . . In the midst of Mali's March 2012 coup, for example, Tuareg rebels launched a powerful military offensive. They and Islamic rebel groups proceeded to capture much of the country before French intervention forces drove them back." Earlier in African history, internecine warfare in the 19th century between different factions of the same kingdom frequently generated a pretense for outsider European intervention.

Given the altered setup with an exogenous outsider threat, the rest of the analysis takes comparative statics on q to understand its consequences for power-sharing and for regime survival.

4.2 The Ambiguous Guardianship Dilemma

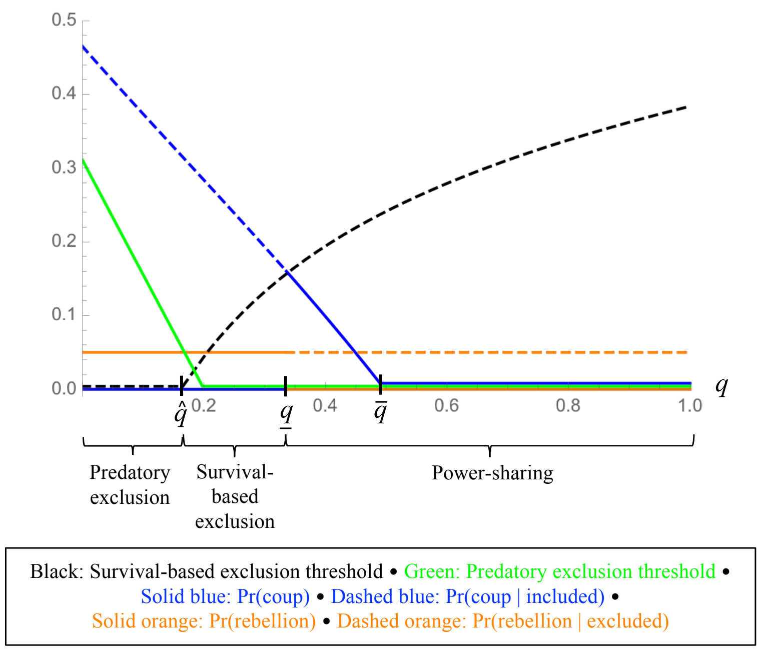
Stronger outsider threats create incentives for the dictator and elites to band together, as McMahon and Slantchev (2015) argue. However, outsider threats can also compel power-sharing arrangements that yield a positive probability of a coup threat, which captures the standard guardianship logic (Acemoglu, Ticchi and Vindigni, 2010; Acemoglu, Vindigni and Ticchi, 2010; Besley and Robinson, 2010; Leon, 2013; Svolik, 2013). Therefore, even when accounting for the endogenous effects of outsider threats on the value of holding office, the overall effect of outsider threats is ambiguous. Only in cases where power-sharing is self-enforcing absent an outsider threat—which the baseline model shows is possible only if excluded elites pose a strong rebellion threat—do outsider threats unambiguously decrease coup risk.

4.2.1 Hang Together or Hang Separately

Sufficiently large outsider threats create a "hang together or hang separately" logic for (1) E not to attempt coups if included and (2) D to share power with E. Both effects occur because the actors fear triggering outsider removal. First, conditional on inclusion, E's coup choice determines whether the probability of outsider takeover is q or $\Delta \cdot q$. In effect, the outsider threat decreases E's expected utility to attempting a coup by raising the probability that it will be unable to rule upon winning its coup and will consume 0, potentially counteracting the gains from attempting a coup described in the baseline model analysis. Second, D knows that the probability of outsider takeover is q if E is excluded and $\left[1 - (1 - \Delta) \cdot \alpha_i^*(q)\right] \cdot q < q$ if E is included. Therefore, D's possible gains from excluding E, described in the baseline model analysis, may be mitigated by raising the probability of outsider takeover.

Figure 5 depicts three aspects of these tradeoffs as a function of q, and Proposition 4 formalizes the main results.[14] The first finding from Figure 5 is that large enough outsider threats deter predatory exclusion. At $q = 0$, E's rebellion threat $1 - \sigma$ (orange line) is low enough that even if the equilibrium probability of a coup under exclusion was 0, D would exclude E for predatory reasons. The green line depicts the predatory exclusion threshold $1 - \hat{\sigma}(q)$. It intersects the orange line at $\hat{q} \equiv \hat{\sigma}^{-1}(\sigma)$, and therefore predatory exclusion occurs for any $q < \hat{q}$. However, large enough outsider threats raise the expected cost of exclusion. In fact, where the green line hits 0, the predatory motive is not sufficient for exclusion even if E was unable to stage a rebellion when excluded. This contrasts with the baseline model in which a rebellion threat is necessary for D to share power.

Figure 5: The Ambiguous Guardianship Dilemma



Notes: Figure 5 uses the following parameter values: $\delta = 0.7$, $p_i = 0.6$, $p_e = 0.1$, $\sigma = 0.95$, $\Delta = 0.2$, $\theta_i = 0.2$.

Second, higher q makes D more tolerant of coup attempts, which the black curve $1 - \underline{\alpha}(q)$ shows. Consistent with the discussion of the predatory exclusion range, this curve equals 0 for any $q < \hat{q}$. For higher values of q, the minimum probability of a coup attempt under inclusion that D is willing to tolerate in order to share

[14] The appendix characterizes analogs for key expressions from the baseline model: $1 - \alpha^*(q)$ is the equilibrium probability of a coup attempt under inclusion (Equation B.11), $1 - \hat{\sigma}(q)$ is the predatory exclusion threshold (Proposition B.1), and $1 - \underline{\alpha}(q)$ is the survival-based exclusion threshold (Proposition B.2).

power strictly increases in q, which part b of Proposition 4 formalizes.[15] For severe-enough outsider threats, D's tolerance for coup attempts under inclusion exceeds its tolerance for rebellions under exclusion, which can be seen in the range where the black line exceeds the orange line. This is impossible in the baseline model because $p_i > p_e$, but the outsider threat causes D to fear coup attempts compared to the alternative of excluding and raising the probability of outsider takeover.

Third, higher q strictly decreases the equilibrium probability of a coup attempt under inclusion until q is large enough that this probability hits 0, which a threshold value $\overline{q}$ denotes (part c of Proposition 4). The blue curve $1 - \alpha^*(q)$ depicts this effect. Whereas the first two effects operated by encouraging D to share power, this effect occurs because the *elite* does not want to raise the probability of outsider takeover.

Combining the effects that higher q increases D's tolerance for facing coup attempts and decreases E's probability of staging a coup if included implies that D will share power with E if q is large enough, formally, larger than a unique threshold value $\underline{q}$. Figure 5 shows that if $q \in \left(\hat{q}, \underline{q}\right)$, then although D will tolerate positive coup probabilities under inclusion, the equilibrium probability of a coup attempt under inclusion is sufficiently high that D excludes for political survival-based reasons. However, if $q > \underline{q}$, then the equilibrium probability of a coup attempt under inclusion is low enough relative to D's tolerance for facing coup attempts that D shares power, which part d of Proposition 4 formalizes.[16]

[15] Importantly, $1-\underline{\alpha}(q)$ is calculated independently of the actual equilibrium probability of a coup attempt, which the blue line captures.

[16] Any of the q thresholds in Proposition 4 may be negative, in which case the behavior stated for sufficiently high values of q are true for all $q \in [0, 1]$. Such parameter values are identical to various cases in the baseline model. If $\hat{q} < 0$, then D does not exclude E for predatory reasons, as in Panel A of Figure 3. If $\underline{q} < 0$, then D shares power with E, which is true for $\theta_i > \underline{\theta}$ in Panel A of Figure 3. If $\overline{q} < 0$, then coup attempts occur under inclusion with probability 0, which is true for $\theta_i > \overline{\theta}$ in both panels of Figure 3.

Proposition 4 (Hang together or hang separately)**.**

- ***Part a.*** *If the outsider threat is sufficiently severe, then predatory motives are not sufficient for D to exclude E. Formally, there exists a unique $\hat{q} < 1$ such that if $q > \hat{q}$, then $1 - \sigma > 1 - \hat{\sigma}(q)$.*
- ***Part b.*** *In the parameter range in which D will not exclude E for purely predatory reasons, D's tolerance to face coup attempts by an included E strictly increases in q. Formally, if $q > \hat{q}$, then $\frac{d}{dq}\big[1 - \underline{\alpha}(q)\big] > 0$.*
- ***Part c.*** *The equilibrium probability of a coup attempt under inclusion strictly decreases in the severity of the outsider threat until the point where the threat is sufficiently ominous that an included E attempts coups with probability 0. Formally, there exists a unique $\overline{q} < 1$ such that:*
  - *If $q < \overline{q}$, then $1 - \alpha^*(q) > 0$ and $\frac{d}{dq}\big[1 - \alpha^*(q)\big] < 0$.*
  - *If $q > \overline{q}$, then $1 - \alpha^*(q) = 0$.*
- ***Part d.*** *If the outsider threat is sufficiently severe, then D will share power with E. Formally, there exists a unique $\underline{q} \in (\hat{q}, 1)$ such that if and only if $q > \underline{q}$, then there exists an equilibrium in which $\mu_t = 1$.*

4.2.2 Whither the Guardianship Dilemma?

Although outsider threats create incentives to hang together, Figure 5 also shows that stronger outsider threats do not necessarily decrease the equilibrium probability of a coup attempt. At the threshold $\underline{q}$ where D changes from excluding to including E, the equilibrium probability of a coup attempt jumps from 0 to positive. This is, in essence, the guardianship logic: outsider threats encourage building a military (or, here, power-sharing) that generates coup risk. Because of the hanging together logic, large enough outsider threats enable power-sharing without coup risk ($q > \overline{q}$). However, a moderately large increase in the severity of the outsider threat raises the equilibrium probability of a coup attempt by making D more tolerant of facing coup attempts, despite also diminishing E's incentives to attempt a coup conditional on inclusion. Also note that this conclusion is identical when considering the broader concept of an elite challenge, i.e., either a rebellion or a coup attempt. This probability exhibits a discrete upward jump at $q = \underline{q}$, as shown by the positive distance between the blue and orange lines.

This is not the first model to generate a non-monotonic relationship between outsider threats and equilibrium coup probabilities, although the logic differs by evaluating the standard guardianship logic in combination with allowing the outsider threat to endogenously affect the value of holding office. Acemoglu, Ticchi and Vindigni (2010) assume that governments can commit to continually pay large enough militaries, which result in equilibrium from large threats. Svolik (2013) shows that the contracting problem between a government and its military dissipates as the military becomes large—which again results from large threats—because the military can control policy without actually intervening (what he calls a "military tutelage" regime). In both models, more extreme outsider threats increase the military's bargaining leverage, whereas here the non-monotonic relationship emerges specifically because greater outsider threats endogenously affect the value of challenging and *diminish* elites' bargaining leverage.

4.2.3 How Permanent Rebellion Threats Can Eliminate the Guardianship Dilemma

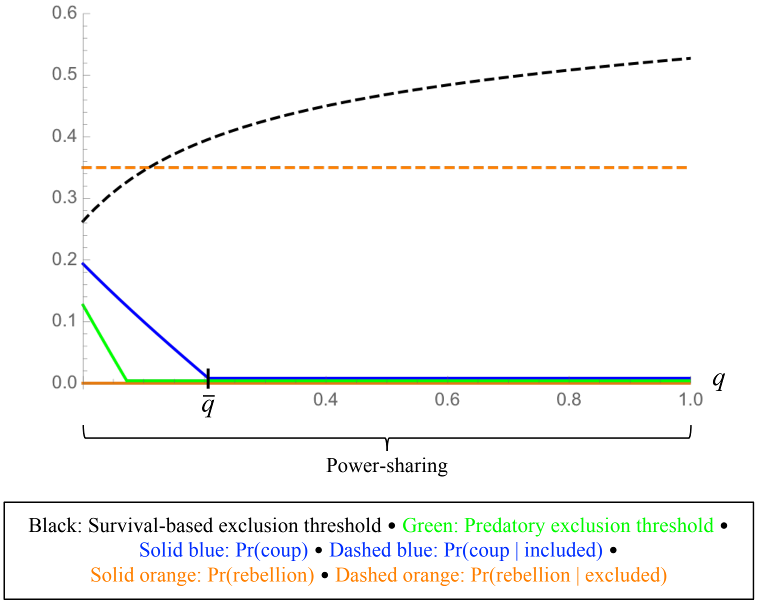
Instead, the presence of a permanent elite threat can make the guardianship logic irrelevant. Figure 6 modifies the equilibrium probability of rebellion under exclusion and the probability of a rebellion succeeding compared to Figure 5. The stronger rebellion threat causes D to share power with E even absent an outsider threat, which is equivalent to the $\theta_i > \underline{\theta}$ range in Panel A of Figure 3. The positive gap between the black and blue lines at $q = 0$ in Figure 6 depicts this calculation. Consequently, the severity of the outsider threat does not affect D's power-sharing decision, which eliminates the mechanism by which a moderately large increase in the outsider threat caused the equilibrium probability of a coup attempt to rise in Figure 5. Instead, the only effect of the outsider threat in Figure 6 is to diminish D's equilibrium probability of a coup attempt under inclusion—because of the hanging together logic—until $q = \underline{q}$, where this probability hits 0.[17]

The necessary condition for eliminating the guardianship dilemma is a strong enough threat of rebellion by E that D shares power absent an outsider threat. This result is not possible in existing models of coups, either those highlighting the guardianship logic (e.g., Besley and Robinson, 2010) or arguing against it (McMahon and Slantchev, 2015): if there is no outsider threat, then the dictator can simply choose not to build a military and therefore faces no risk of overthrow by insiders.[18] By contrast, the present model

[17]There is also a trivial case in which outsider threats exert no effect on equilibrium power-sharing or the equilibrium probability of a coup attempt because θ_i is high enough that, at $q = 0$, D includes E and E attempts coups with probability 0 (see the $\theta > \overline{\theta}$ range in Panel A of Figure 3).

[18]In McMahon and Slantchev (2015), this would entail the ruler not delegating national defense to a specialized military agent. Although possible in their model, they explicitly only analyze parameter values

Figure 6: Permanent Elite Threats and the Guardianship Dilemma



Notes: Figure 5 uses the following parameter values: $\delta = 0.7$, $p_i = 0.6$, $p_e = 0.4$, $\sigma = 0.65$, $\Delta = 0.2$, $\theta_i = 0.3$.

presumes that a dictator always faces a threat from other elites, and shows that the threat of a rebellion can be sufficient to compel power-sharing—despite creating a coup risk—absent an outsider threat.

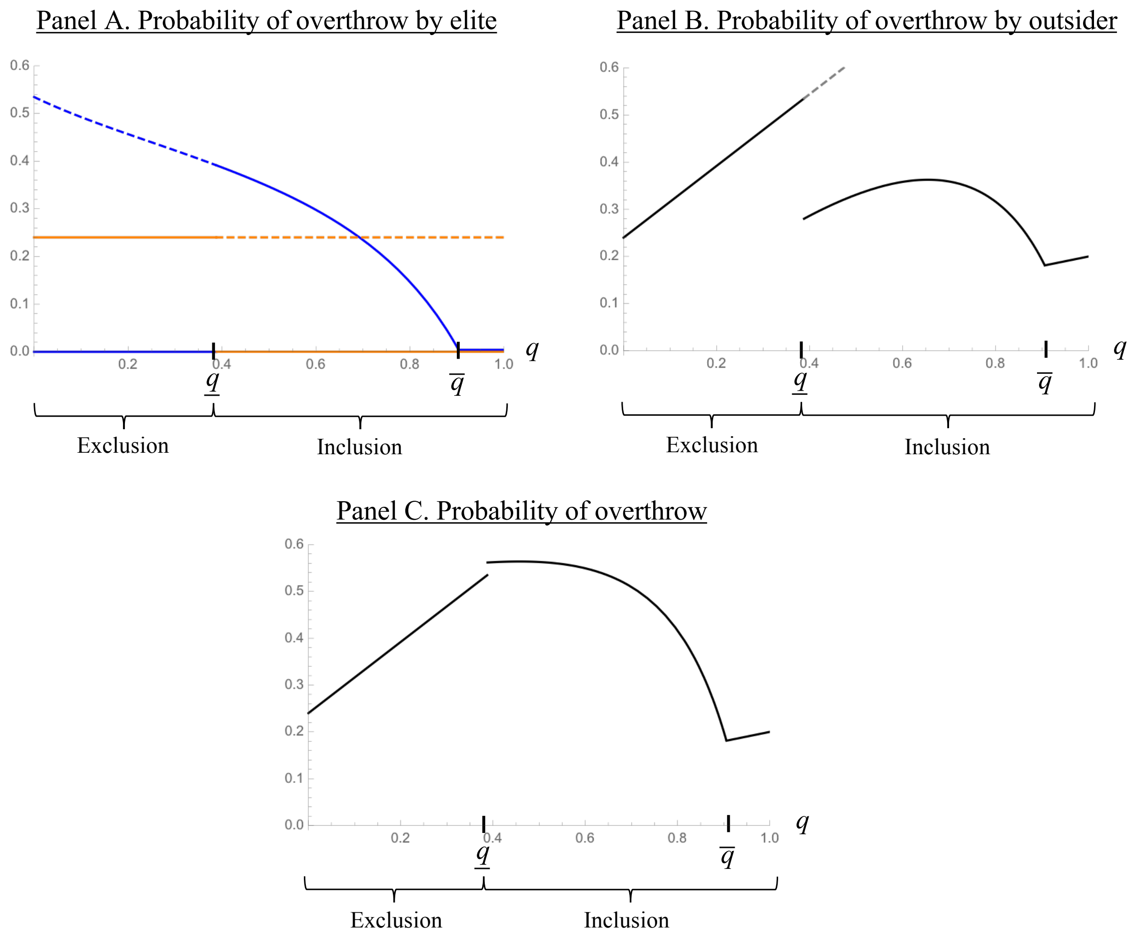
4.3 Outsider Threats and Durable Authoritarian Regimes

Although outsider threats ambiguously affect the equilibrium probability of a coup attempt, it is possible for strong outsider threats to not only facilitate peaceful power-sharing among D and E, but to lower the *overall* probability of overthrow. This is striking when considering that the direct effect of an outsider threat is to raise the exogenous probability of regime overthrow, and provides a possible path to durable and peaceful regimes in weakly institutionalized states.

Figure 7 depicts the equilibrium probability of overthrow by the elite either via coup or rebellion (Panel A), by the outsider (Panel B), or by either (Panel C). In Panel A, the color scheme follows that in the previous figures, with blue corresponding to coups and orange to rebellions. It generates identical insights as Figure 5, except that it depicts the equilibrium probabilities of *successful* overthrow rather than equilibrium attempts.

in which the outsider threat is sufficiently large that the ruler optimally chooses to delegate to a military agent, which creates positive coup risk.

Figure 7: Equilibrium Probability of Overthrow



Notes: $\delta = 0.9$, $p_i = 0.6$, $p_e = 0.4$, $\sigma = 0.4$, $\Delta = 0.2$, $\sigma = 0.9$, and $\theta_i = 0.05$.

Panel B depicts the equilibrium per-period probability that the outsider will overthrow the regime. For $q < \underline{q}$, the overthrow probability strictly increases in q because q exerts a direct effect of increasing the likelihood of overthrow and, in this parameter range, the coalitional dynamics between D and E are constant. The overthrow probability exhibits a discrete drop at $q = \underline{q}$ because D moves from excluding E to sharing power, which decreases the probability of outsider overthrow from q to $\left[1 - (1 - \Delta) \cdot \alpha_i^*(q)\right] \cdot q < q$. For $q \in \left(\underline{q}, \overline{q}\right)$, there are two countervailing effects. First, q exerts the same overthrow-inducing direct effect. However, higher q also strictly decreases the probability of a coup attempt by E (and therefore of a successful coup attempt; see Panel A). These countervailing effects explain the non-monotonic effect of q in this parameter range. Finally, for $q > \overline{q}$, the threat is large enough that the equilibrium probability of a coup attempt is 0, and after this point the only effect of a stronger outsider threat is the direct effect that increases overthrow likelihood.

Panel C combines the previous two panels by presenting the equilibrium per-period probability that either the elite or the outsider will overthrow D. For $q < \underline{q}$, this probability strictly increases in q because the probability that E overthrows D in a rebellion remains constant whereas the outsider becomes increasingly likely to overthrow the regime. At $q = \underline{q}$, the overall probability shifts in response to the discrete increase in the probability of overthrow by E (Panel A; this is the guardianship logic) and the discrete drop in the probability of outsider overthrow (Panel B). Because the probability of overthrow by E strictly decreases for $q \in \left(\underline{q}, \overline{q}\right)$—the hang together effect—and this effect is large in magnitude relative to the effect of q on outsider overthrow in this range, the overall probability of overthrow strictly decreases. Finally, the overall overthrow probability rises again for $q > \overline{q}$ because this range only exhibits the direct effect of the outsider threat.

Although there are other parameter values in which $q = 0$ minimizes the total overthrow probability, Figure 7 presents the striking finding that strong outsider threats can underpin durable and peaceful regimes that share power among elites even in weakly institutionalized states that would otherwise would create ripe conditions for factional splits. Contrary to Roessler and Ohls's (2018) argument that only strong rebellion threats by the elite can compel power-sharing and that power-sharing regimes in weak states are necessarily highly vulnerable to coups, outsider threats offer an alternative path to power-sharing and can cause the dictator and elite to hang together to prevent hanging separately.

South Africa prior to 1994 illustrates this dynamic. The Union of South Africa gained independence in 1910 as an amalgam of four regionally distinct colonies. Among the European population, two regions were dominated by British descendants and two by Dutch descendants. Despite sharing European heritage, South Africa faced political divisions at independence between British and Boer. In fact, the territory experienced a major internal war less than a decade prior between the British and Boer factions. "When South Africans spoke of the 'race question' in the early part of the [20th] century, it was generally accepted that they were referring to the division between Dutch or Afrikaners on the one hand and British or English-speakers on the other" (Lieberman, 2003, 76). This division created debates among English settlers (who were victorious in the Boer War) about how widely to share power with Afrikaners when writing the foundational constitution. In terms of the model, commitment ability θ_i was relatively low. However, whites also faced a grave potential threat from the African majority that composed roughly 80% of the population at independence, which corresponds with high q. Furthermore, despite their numerical deficiency, Europeans' military superiority implied that if they banded together, the likelihood of external overthrow was low (at least in the medium-term), implying low Δ.[19] This exemplifies how outsider threats can facilitate power-sharing in a case that otherwise might have featured factional conflict among British and Boers.

5 Conclusion

Elites and outsiders threaten dictators. How do autocrats manage these threats, and what are the consequences for regime survival? This paper analyzes a repeated interaction in which a dictator chooses whether or not to share power with another elite, and then the dictator and elite bargain over state revenues. In the baseline model without outsider threats, the dictator faces predatory incentives to exclude and only shares power if elites pose a strong rebellion threat—in which case institutional quality and threat capabilities are non-monotonically related to equilibrium coup propensity because of survival-based exclusion incentives. Introducing an outsider threat further encourages power-sharing by creating incentives to "hang together" to prevent "hanging separately." Although the hang-together effect may raise coup likelihood by creating a guardianship dilemma, it also implies that outsider threats may underpin inclusive and durable authoritarian regimes in weakly institutionalized environments.

References

Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*.

Acemoglu, Daron and James A. Robinson. 2017. "Why Did the West Expand the Franchise? A Correction.". Mimeo. Available at https://economics.mit.edu/files/12738. Accessed 4/25/18.

Acemoglu, Daron, Andrea Vindigni, and Davide Ticchi. 2010. "Persistence of Civil Wars." *Journal of the European Economic Association* 8(2-3):664–676.

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni. 2010. "A Theory of Military Dictatorships." *American Economic Journal: Macroeconomics* 2(1):1–42.

Ansell, Ben W. and David J. Samuels. 2014. *Inequality and Democratization: An Elite Competition Approach*. Cambridge, UK: Cambridge University Press.

Besley, Timothy and James A. Robinson. 2010. "Quis Custodiet Ipsos Custodes? Civilian Control Over the Military." *Journal of the European Economic Association* 8(2-3):655–663.

Cederman, Lars-Erik, Kristian Skrede Gleditsch, and Halvard Buhaug. 2013. *Inequality, Grievances, and Civil War*. Cambridge University Press.

[19]Although high repression costs eventually compelled whites to share power with Africans in 1994, this occurred 84 years after independence.

Debs, Alexandre. 2016. "Living by the Sword and Dying by the Sword? Leadership Transitions in and out of Dictatorships." *International Studies Quarterly* 60(1):73–84.

Dragu, Tiberiu and Adam Przeworski. 2017. "Preventive Repression: Two Types of Moral Hazard." Working Paper, Department of Political Science, New York University.

Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer Than Others?" *Journal of Peace Research* 41(3):275–301.

Feaver, Peter D. 1999. "Civil-Military Relations." *Annual Review of Political Science* 2(1):211–241.

Finer, Samuel E. 2002. *The Man on Horseback: The Role of the Military in Politics*. Transaction Publishers.

Finer, Samuel Edward. 1997. *The History of Government from the Earliest Times: Ancient monarchies and empires*. Vol. 1 Oxford University Press, USA.

Francois, Patrick, Ilia Rainer, and Francesco Trebbi. 2015. "How is Power Shared in Africa?" *Econometrica* 83(2):465–503.

Gehlbach, Scott and Philip Keefer. 2011. "Investment Without Democracy: Ruling-Party Institutionalization and Credible Commitment in Autocracies." *Journal of Comparative Economics* 39(2):123–139.

Gibilisco, Michael. 2017. "Decentralization and the Gamble for Unity.". Working paper, Division of the Humanities and Social Sciences, California Institute of Technology. Available at https://www.dropbox.com/s/h7s0shnbmbloy3x/grievanceGibilisco.pdf?dl=0. Accessed 8/17/17.

Goodwin, Jeff. 2001. *No Other Way Out: States and Revolutionary Movements, 1945-1991*. Cambridge University Press.

Goodwin, Jeff and Theda Skocpol. 1989. "Explaining Revolutions in the Contemporary Third World." *Politics & Society* 17(4):489–509.

Harkness, Kristen A. 2016. "The Ethnic Army and the State: Explaining Coup Traps and the Difficulties of Democratization in Africa." *Journal of Conflict Resolution* 60(4):587–616.

Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. University of California Press.

Huntington, Samuel P. 1957. *The Soldier and the State: The theory and Politics of Civil-Military Relations*. Harvard University Press.

Kalandrakis, Tasos. 2016. "Pareto Efficiency in the Dynamic One-Dimensional Bargaining Model." *Journal of Theoretical Politics* 28(4):525–536.

Krainin, Colin. 2017. "Preventive War as a Result of Long Term Shifts in Power." *Political Science Research and Methods* 5(1):103–121.

Leon, Gabriel. 2013. "Soldiers or politicians? Institutions, conflict, and the military?s role in politics." *Oxford Economic Papers* 66(2):533–556.

Lieberman, Evan S. 2003. *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*. Cambridge, UK: Cambridge University Press.

McMahon, R. Blake and Branislav L. Slantchev. 2015. "The Guardianship Dilemma: Regime Security Through and from the Armed Forces." *American Political Science Review* 109(2):297–313.

Padró-i Miquel, Gerard. 2007. "The Control of Politicians in Divided Societies: The Politics of Fear." *Review of Economic Studies* 74(4):1259–1274.

Powell, Robert. 2004. "The Inefficient Use of Power: Costly Conflict with Complete Information." *American Political Science Review* 98(2):231–241.

Roessler, Philip. 2011. "The Enemy Within: Personal Rule, Coups, and Civil War in Africa." *World Politics* 63(2):300–346.

Roessler, Philip. 2016. *Ethnic Politics and State Power in Africa: The Logic of the Coup-Civil War Trap*. Cambridge University Press.

Roessler, Philip and David Ohls. 2018. "Self-Enforcing Power Sharing in Weak States." *International Organization, forthcoming* .

Slater, Dan. 2010. *Ordering Power: Contentious Politics and Authoritarian Leviathans in Southeast Asia*. Cambridge University Press.

Snyder, Richard. 1998. "Paths out of Sultanistic Regimes: Combining Structural and Voluntarist Perspectives." In *Sultanistic Regimes*, ed. H.E. Chehabi and Juan J. Linz. Baltimore, MD: Johns Hopkins University Press pp. 49–81.

Svolik, Milan W. 2009. "Power sharing and leadership dynamics in authoritarian regimes." *American Journal of Political Science* 53(2):477–494.

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

Tyson, Scott A. 2017. "The Agency Problem Underlying Repression.". Working paper, Department of Political Science, University of Michigan.

Online Appendix

A Proofs for Baseline Model

A.1 Baseline Model

Proof of Lemma 2. The restriction to stationary strategies implies in the bargaining phase that any strategy must take the form $x_t = x'$ and $\alpha_t(x_t) = \alpha'(x_t)$, for some $x' \in [0,1]$ and $\alpha'(\cdot) \in [0,1]$. It is without loss of generality to restrict D's offer to a pure strategy because, as shown below, D is never indifferent among its bargaining offers. With some abuse of notation, define $\alpha' \equiv \alpha'(x')$. Denote the future continuation values as $V^{D'}$ and $V^{E'}$. The following three steps demonstrate that the bargaining strategy profile stated in Lemma 2 is an equilibrium and is unique.

1. ***Necessary condition for C's best-response correspondence.*** In any equilibrium in which $\mu_t = \mu^i$ for all t, $\alpha'(x_t)$ must satisfy:

$$\alpha'(x_t) = \begin{cases} 0 & \text{if } E\big[U_E(\text{accept } |x_t, x', \alpha')\big] < E\big[U_E(\text{fight}|x', \alpha')\big] \\ [0,1] & \text{if } E\big[U_E(\text{accept } |x_t, x', \alpha')\big] = E\big[U_E(\text{fight}|x', \alpha')\big] \\ 1 & \text{if } E\big[U_E(\text{accept } |x_t, x', \alpha')\big] > E\big[U_E(\text{fight}|x', \alpha')\big], \end{cases} \tag{A.1}$$

where Equations 1 and 2 (after appropriate changes in notation) enable solving for:

$$E\big[U_E(\text{accept } |x_t, x', \alpha')\big] = \theta_i \cdot x_t + \delta \cdot V^{E'} \tag{A.2}$$

$$E\big[U_E(\text{fight}|x', \alpha')\big] = \delta \cdot \Big[p_i \cdot V^{D'} + (1-p_i) \cdot V^{E'}\Big] \tag{A.3}$$

$$V^{D'} = \frac{(1-\delta)\cdot(1-\theta_i \cdot x') + \delta \cdot p_i \cdot (1-\alpha')}{(1-\delta)\cdot\Big[1-\delta\cdot\big[1-2p_i\cdot(1-\alpha')\big]\Big]} \cdot \alpha' \tag{A.4}$$

$$V^{E'} = \frac{(1-\delta)\cdot\theta_i \cdot x' + \delta \cdot p_i \cdot (1-\alpha')}{(1-\delta)\cdot\Big[1-\delta\cdot\big[1-2p_i\cdot(1-\alpha')\big]\Big]} \cdot \alpha' \tag{A.5}$$

2. ***Necessity of indifference.*** Any equilibrium strategy profile requires $E\big[U_E(\text{accept } |x', x', \alpha')\big] = E\big[U_E(\text{fight}|x', \alpha')\big]$.

   - Suppose not and $E\big[U_E(\text{accept } |x', x', \alpha')\big] > E\big[U_E(\text{fight}|x', \alpha')\big]$. Then D has a profitable deviation to some $x_t < x'$. To show this, the continuity of $E\big[U_E(\text{accept } |x_t, x', \alpha')\big]$ in x_t and the completeness of $\mathbb{R}$ implies that there exists $\epsilon' > 0$ such that if $E\big[U_E(\text{accept } |x', x', \alpha')\big] > E\big[U_E(\text{fight}|x', \alpha')\big]$, then $E\big[U_E(\text{accept } |x' - \epsilon', x', \alpha')\big] > E\big[U_E(\text{fight}|x', \alpha')\big]$. D has a profitable deviation to $x_t = x' - \epsilon'$ if $1 - \theta_i \cdot (x' - \epsilon') + \delta \cdot V^{D'} > 1 - \theta_i \cdot x' + \delta \cdot V^{D'}$, which easily reduces to $\epsilon' > 0$.

   - Suppose not and $E\big[U_E(\text{accept } |x', x', \alpha')\big] < E\big[U_E(\text{fight}|x', \alpha')\big]$. Then E has a profitable deviation to $\alpha'(x_t) = 1$ for any $x_t > 0$. To see this, if $E\big[U_E(\text{accept } |x', x', \alpha')\big] < E\big[U_E(\text{fight}|x', \alpha')\big]$, then E's best-response correspondence implies $\alpha' = 0$. Substituting in equilibrium terms shows that $E\big[U_E(\text{accept }|x_t, x', 0)\big] = \theta_i \cdot x_t > 0 = E\big[U_E(\text{fight}|x', 0)\big]$. Intuitively, this result follows because if $\alpha' = 0$, then $V^{D'} = V^{E'} = 0$.

3. ***Relationship between offer and acceptance probability.*** The previous steps yield two facts about the relationship between x' and α' in any equilibrium.

   - $\max\{x', \alpha'\} = 1$. The assumed upper bounds of x' and α' imply $\max\{x', \alpha'\} \leq 1$. To prove $\max\{x', \alpha'\} \geq 1$, can show that if instead $x' < 1$ and $\alpha' < 1$, then D has a profitable deviation to some $x_t > x'$. To see this, because $E\big[U_E(\text{accept }|x_t, x', \alpha')\big]$ is continuous and strictly increases in x_t, for any $\epsilon > 0$, $E\big[U_E(\text{accept }|x' + \frac{\epsilon}{\theta_i}, x', \alpha')\big] > E\big[U_E(\text{accept }|x', x', \alpha')\big] = E\big[U_E(\text{fight}|x', \alpha')\big]$, where the latter equality follows from the necessity of indifference in any equilibrium. This in turn implies $\alpha'\big(x' + \frac{\epsilon}{\theta_i}\big) = 1$ because of the necessary condition for E's best-response correspondence. D can profitably deviate to $x_t = x' + \frac{\epsilon}{\theta_i}$ if:

     $$1 - \theta_i \cdot \left(x' + \frac{\epsilon}{\theta_i}\right) + \delta \cdot V^{D'} > \alpha' \cdot \left[1 - \theta_i \cdot x' + \delta \cdot V^{D'}\right] + (1 - \alpha') \cdot \delta \cdot \left[p_i \cdot V^{E'} + (1 - p_i) \cdot V^{D'}\right]$$

     This solves to:

     $$\epsilon < (1 - \alpha') \cdot \Bigg[1 - \underbrace{\Big(\theta_i \cdot x' - \delta \cdot p_i \cdot \big(V^{D'} - V^{E'}\big)\Big)}_{=E[U_E(\text{accept }|x', x', \alpha')] - E[U_E(\text{fight}|x', \alpha')]}\Bigg]$$

     The necessity of indifference implies $E\big[U_E(\text{accept }|x', x', \alpha')\big] - E\big[U_E(\text{fight}|x', \alpha')\big] = 0$, which enables further reducing the inequality to $\epsilon < 1 - \alpha'$. The completeness of $\mathbb{R}$ implies that for all $\alpha' \in (0, 1)$ there exists $\epsilon \in \big(0, 1 - \alpha'\big)$. Furthermore, because $x' < 1$, the completeness of $\mathbb{R}$ also implies that ϵ can be chosen such that $x' + \frac{\epsilon}{\theta_i} < 1$, which generates the contradiction.

   - Any equilibrium strategy profile requires:

     $$\theta_i \cdot x' = \delta \cdot p_i \cdot \underbrace{\underbrace{\frac{1 - \delta}{1 - \delta \cdot (1 - 2p_i)}}_{\in(0,1)} \cdot \underbrace{\frac{\alpha'}{1 - \delta}}_{=V^{D'} + V^{E'}}}_{=V^{D'} - V^{E'}}, \tag{A.6}$$

     which follows from combining Equations A.2 through A.5 and $E\big[U_E(\text{accept }|x', x', \alpha')\big] = E\big[U_E(\text{fight}|x', \alpha')\big]$. This can equivalently be stated as:

     $$(x', \alpha') = \left(\frac{1}{\theta_i} \cdot \frac{\delta \cdot p_i}{1 - \delta \cdot (1 - 2p_i)} \cdot \alpha', \alpha'\right) \tag{A.7}$$

     One immediate consequence of this result is that, in any equilibrium, there exists a 1-1 relationship between x' and α'.

Combining these two facts with Equation A.1 easily demonstrates that the bargaining strategy profile stated in Lemma 2 is a bargaining equilibrium and is unique. ■

If D deviates to exclude E, then E's constraint for accepting with positive probability is $\theta_t \cdot x_t \geq \delta \cdot p_e \cdot \left(V^D - V^E\right)$. This constraint follows from similar logic that yielded Equation 3, with two differences. First, θ_t may equal either θ_i or 0, capturing that, on average, D's commitment ability is lower if E is excluded from power. Second, E's probability of winning is p_e rather than $p_i > p_e$. However, the future continuation values are the same as above because we are evaluating a single deviation from a strategy profile in which D shares power with E in every period.

Lemma A.1 shows that there exists an offer that E will accept an offer with probability 1 if D deviates to exclusion.

Lemma A.1 (Acceptance following a deviation to exclusion)**.** *If D deviates to exclude E, then there exists $x_t < 1$ such that E will accept with probability 1.*

Proof. A sufficient inequality is:

$$\theta_i \geq \delta \cdot p_e \cdot (V^D - V^E) = \delta \cdot p_e \cdot \frac{\alpha^*}{1 - \delta \cdot (1 - 2p_i)}$$

There are two cases to consider. First, $\theta_i > \overline{\theta} = \frac{\delta \cdot p_i}{1 - \delta \cdot (1 - 2p_i)}$ and $\alpha^* = 1$. Second, $\theta_i < \overline{\theta}$ and $\alpha^* = \frac{1 - \delta \cdot (1 - 2p_i)}{\delta \cdot p_i} \cdot \theta_i$. In both cases, $p_i > p_e$ implies the inequality holds. ■

Lemma A.2 (Bargaining after a deviation to exclusion)**.** *Assume a strategy profile in which D shares power with E in every period, but D deviates by excluding E at time t. If D's commitment ability is the high value θ_i, then D chooses the unique x_t that satisfies with equality E's acceptance constraint. E accepts offer at least that large with probability 1, and all lower offers with probability 0. In equilibrium, E accepts with probability 1. If D does not have commitment ability, then D offers any $x_t \in [0, 1]$ and E accepts any offer with probability 0. Formally:*

- *If $\mu_t = \mu^e$ and $\theta_t = \theta_i$, then D offers $x_t = \tilde{x}_e^*$, for:*

$$\tilde{x}_e^* = \frac{1}{\theta_i} \cdot \frac{\delta \cdot p_e}{1 - \delta \cdot (1 - 2p_i)} \tag{A.8}$$

  E's equilibrium probability of acceptance function is:

$$\alpha_{e,i}^*(x_t) = \begin{cases} 0 & \text{if } x_t < \tilde{x}_e^* \\ 1 & \text{if } x_t \geq \tilde{x}_e^*, \end{cases}$$

  and the equilibrium probability of acceptance is $\alpha_{e,i}^* = 1$.

- *If $\mu_t = \mu^e$ and $\theta_t = 0$, then D offers $x_t \in [0, 1]$. E's equilibrium probability of acceptance function is $\alpha_{e,e}^*(x_t) = 0$ for all $x_t \in [0, 1]$, which also implies the equilibrium probability of acceptance is $\alpha_{e,e}^* = 0$.*

Proof of Lemma A.2. The following three steps demonstrate that the bargaining strategy profile stated in Lemma A.2 is an equilibrium and is unique.

1. ***Necessary condition for C's best-response correspondence.*** Suppose $\mu_t = \mu^e$ and $\mu_s = \mu^i$ for all $s > t$. Denote E's acceptance correspondence if $\theta_t = \theta_i$ as $\alpha'_{e,i}(\cdot)$ and if $\theta_t = 0$ as $\alpha'_{e,e}(\cdot)$. The respective acceptance correspondences must satisfy:

$$\alpha'_{e,i}(x_t) = \begin{cases} 0 & \text{if } E\big[U_E(\text{accept } |x_t, V^D, V^E; \theta_i)\big] < E\big[U_E(\text{fight}|V^D, V^E; p_e)\big] \\ [0,1] & \text{if } E\big[U_E(\text{accept } |x_t, V^D, V^E; \theta_i)\big] = E\big[U_E(\text{fight}|V^D, V^E; p_e)\big] \\ 1 & \text{if } E\big[U_E(\text{accept } |x_t, V^D, V^E; \theta_i)\big] > E\big[U_E(\text{fight}|V^D, V^E; p_e)\big], \end{cases} \tag{A.9}$$

$$\alpha'_{e,e}(x_t) = \begin{cases} 0 & \text{if } E\big[U_E(\text{accept } |x_t, V^D, V^E; 0)\big] < E\big[U_E(\text{fight}|V^D, V^E; p_e)\big] \\ [0,1] & \text{if } E\big[U_E(\text{accept } |x_t, V^D, V^E; 0)\big] = E\big[U_E(\text{fight}|V^D, V^E; p_e)\big] \\ 1 & \text{if } E\big[U_E(\text{accept } |x_t, V^D, V^E; 0)\big] > E\big[U_E(\text{fight}|V^D, V^E; p_e)\big], \end{cases} \tag{A.10}$$

for:

$$E\big[U_E(\text{accept } |x_t, V^D, V^E; \theta_i)\big] = \theta_i \cdot x_t + \delta \cdot V^E \tag{A.11}$$

$$E\big[U_E(\text{accept } |x_t, V^D, V^E; 0)\big] = 0 \tag{A.12}$$

$$E\big[U_E(\text{fight}|V^D, V^E; p_e)\big] = \delta \cdot \Big[p_e \cdot V^D + (1 - p_e) \cdot V^E\Big], \tag{A.13}$$

Denote a generic offer under exclusion if $\theta_t = \theta_i$ as $x'_{e,i}$ and if $\theta_t = 0$ as $x'_{e,e}$.

2. ***Necessity of indifference if $\boldsymbol{\theta_t = \theta_i}$.*** Any equilibrium strategy profile requires $E\big[U_E(\text{accept } |x'_{e,i}, V^D, V^E; \theta_i)\big] = E\big[U_E(\text{fight}|V^D, V^E; p_e)\big]$.

   - Suppose not and $E\big[U_E(\text{accept } |x'_{e,i}, V^D, V^E; \theta_i)\big] > E\big[U_E(\text{fight}|V^D, V^E; p_e)\big]$. Then D has a profitable deviation to some $x_t < x'_{e,i}$. To show this, the continuity of $E\big[U_E(\text{accept } |x_t, V^D, V^E; \theta_i)\big]$ in x_t and the completeness of $\mathbb{R}$ implies that there exists $\epsilon' > 0$ such that if $E\big[U_E(\text{accept } |x'_{e,i}, V^D, V^E)\big] > E\big[U_E(\text{fight}|V^D, V^E)\big]$, then $E\big[U_E(\text{accept } |x'_{e,i} - \epsilon', V^D, V^E)\big] > E\big[U_E(\text{fight}|V^D, V^E)\big]$. D has a profitable deviation to $x_t = x'_{e,i} - \epsilon'$ if $1 - \theta_i \cdot (x'_{e,i} - \epsilon') + \delta \cdot V^D > 1 - \theta_i \cdot x'_{e,i} + \delta \cdot V^D$, which easily reduces to $\epsilon' > 0$.

   - Suppose not and $E\big[U_E(\text{accept } |x'_{e,i}, V^D, V^E; \theta_i)\big] < E\big[U_E(\text{fight}|V^D, V^E; p_e)\big]$. Then there exists $\epsilon > 0$ such D has a profitable deviation to some $x_t = x'_{e,i} + \frac{\epsilon}{\theta_i}$ such that $E\big[U_E(\text{accept } |x'_{e,i} + \frac{\epsilon}{\theta_i}, V^D, V^E; \theta_i)\big] > E\big[U_E(\text{fight}|V^D, V^E; p_e)\big]$, which in turn implies $\alpha'_{e,i}(x'_{e,i} + \frac{\epsilon}{\theta_i}) = 1$:

   $$1 - \theta_i \cdot \left(x'_{e,i} + \frac{\epsilon}{\theta_i}\right) + \delta \cdot V^D > \delta \cdot \Big[p_e \cdot V^E + (1 - p_e) \cdot V^D\Big],$$

   which easily rearranges to:

   $$1 - \theta_i \cdot \left(x'_{e,i} + \frac{\epsilon}{\theta_i}\right) + \delta \cdot p_e \cdot \left(V^D - V^E\right) > 0,$$

   For all, $x'_{e,i}$, Lemma A.1 implies the existence of $\epsilon > 0$ such that $x'_{e,i} + \frac{\epsilon}{\theta_i} < 1$.

- This result easily establishes the uniqueness of $\tilde{x}_e^*$ stated in the lemma, as well as E's best-response correspondence.

3. ***Impossibility of peace if $\theta_t = 0$.*** It is trivial to establish that $E\big[U_E(\text{accept }|x'_{e,e}, V^D, V^E; 0)\big] < E\big[U_E(\text{fight}|V^D, V^E; p_e)\big]$ for all $x'_{e,e} \in [0,1]$. ■

Before proving Propositions 1 and 2, it will be useful to restate D's inclusion constraint from Equation 6 after substituting in for the equilibrium continuation values and rearranging:

$$\left[1 - \frac{\delta \cdot (p_i - p_e)}{1 - \delta \cdot (1 - 2p_i)}\right] \cdot \alpha_i^* - \sigma \geq 0 \tag{A.14}$$

Proof of Proposition 1. It is trivial to establish that $1 > \frac{\delta \cdot (p_i - p_e)}{1 - \delta \cdot (1 - 2p_i)}$. Therefore, the left-hand side of Equation A.14 achieves its upper bound if $\alpha_i^* = 1$. Substituting this into Equation A.14 shows that D has a profitable deviation if:

$$\sigma > \hat{\sigma} \equiv 1 - \frac{\delta \cdot (p_i - p_e)}{1 - \delta \cdot (1 - 2p_i)}$$

■

Proof of Proposition 2. Rearranging Equation A.14 shows that D will not deviate to exclusion only if E's probability of acceptance under inclusion is sufficiently high:

$$\alpha_i^* \geq \underline{\alpha} \equiv \frac{\sigma}{1 - \frac{\delta \cdot (p_i - p_e)}{1 - \delta \cdot (1 - 2p_i)}} \tag{A.15}$$

Substituting in for $\tilde{\alpha}_i^*$ from Equation 5 and rearranging enables stating:

$$\Omega(\theta_i) \equiv \frac{1 - \delta \cdot [1 - (p_i + p_e)]}{\delta \cdot p_i} \cdot \theta_i - \sigma \tag{A.16}$$

Applying the intermediate value theorem demonstrates the existence of at least one $\underline{\theta} \in \big(0, \overline{\theta}\big)$ that satisfies $\Omega(\underline{\theta}) = 0$.

- $\Omega(0) = -\sigma < 0$
- $\Omega\big(\overline{\theta}\big) = \frac{1 - \delta \cdot [1 - (p_i + p_e)]}{1 - \delta \cdot (1 - 2p_i)} - \sigma > 0$. Proposition 1 defines $\hat{\sigma}$, and the statement of the present proposition assumes $\sigma < \hat{\sigma}$.
- $\Omega(\theta_i)$ is continuous in θ_i.

The threshold claim follows because $\Omega(\theta_i)$ strictly increases in θ_i. The explicit characterization is:

$$\underline{\theta} \equiv \frac{\delta \cdot p_i \cdot \sigma}{1 - \delta \cdot \big[1 - (p_i + p_e)\big]}$$

■

B Model with an Outsider Threat

Notation-wise, the terms for actions and thresholds in the extension are identical to the baseline model except they are followed by q in parentheses. For example, the equilibrium continuation values are now $V^D(q)$ and $V^E(q)$.

B.1 Acceptance Constraint

If a power-sharing equilibrium exists, then the continuation values for the dictator and elite respectively equal:

$$V^D(q) = \alpha_i^*(q) \cdot \Big[1 - \theta_i \cdot x_i^*(q) + (1 - \Delta \cdot q) \cdot \delta \cdot V^D(q)\Big] + \big(1 - \alpha_i^*(q)\big) \cdot (1-q) \cdot \delta \cdot \Big[p_i \cdot V^E(q) + (1-p_i) \cdot V^D(q)\Big] \tag{B.1}$$

$$V^E(q) = \alpha_i^*(q) \cdot \Big[\theta_i \cdot x_i^*(q) + (1 - \Delta \cdot q) \cdot \delta \cdot V^E(q)\Big] + \big(1 - \alpha_i^*(q)\big) \cdot (1-q) \cdot \delta \cdot \Big[p_i \cdot V^D(q) + (1-p_i) \cdot V^E(q)\Big] \tag{B.2}$$

Although these terms resemble Equations 1 and 2, they each contain additional expressions that express the possibility of outsider takeover, after which D and E consume 0 in all future periods. Following similar logic as Equation 3,[20] these can be solved to show the optimal offer and probability of acceptance terms satisfy:

$$\Omega\big(x_i^*(q), \alpha_i^*(q); q\big) \equiv \underbrace{\frac{1}{1 - \delta \cdot (1 - \Delta \cdot q)}}_{\text{PDV if accept}} \cdot \underbrace{\theta_i \cdot x_i^*(q)}_{\text{Current opp. cost}}$$

$$- \underbrace{\frac{1}{1 - \delta \cdot (1-q)}}_{\text{PDV if coup}} \cdot \underbrace{(1-q)}_{\text{Pr(no outsider takeover)}} \cdot \underbrace{\delta \cdot p_i \cdot \overbrace{\overbrace{\frac{1 - \delta \cdot (1-q)}{1 - \delta \cdot (1-q) \cdot (1 - 2p_i)}}^{\in(0,1)} \cdot \overbrace{\frac{\alpha_i^*(q)}{1 - \delta \cdot \Big[1 - \big(1 - \alpha_i^*(q) \cdot (1 - \Delta)\big) \cdot q\Big]}}^{=V^D(q)+V^E(q)}}^{V^D(q)-V^E(q)}}_{\text{Expected future gains from winning conditional on no outsider takeover}} = 0 \tag{B.3}$$

As in Equation 3, E faces a current-period opportunity cost of fighting, $\theta_i \cdot x_i^*(q)$, and reaps expected future gains from winning, $\delta \cdot p_i \cdot \big[V^D(q) - V^E(q)\big]$. Unlike in the baseline model, the latter term is multiplied by the probability that no outsider takeover occurs following a coup attempt, $1 - q$.[21] By increasing the probability of outsider takeover, higher q decreases the expected utility of attempting a coup. Intuitively, E places less weight on the value of winning when the outside threat is present because simply defeating D is not sufficient to gain the prize of governing. The expected future gains from winning term itself is also somewhat different than in the baseline model for two reasons. First, total lifetime surplus, $V^D(q) + V^E(q)$, is lower in the

[20]Equations 3 and B.3 are identical if $q = 0$.

[21]Additionally, the current-period opportunity cost is multiplied by the present-discounted value of consumption if E accepts, and the future gains from winning term is multiplied by the present-discounted value of consumption if E attempts a coup. Similar terms do not appear in Equation 3 because both present-discounted value multipliers equal $\frac{1}{1-\delta}$ in the baseline model, and therefore cancel out. The present-discounted value multipliers differ in Equation B.3 because of the possibility of outsider takeover—which differs depending on whether E accepts or attempts a coup—and, consequently, of no future consumption.

baseline model because of the $\left[1-\alpha_i^*(q)\cdot(1-\Delta)\right]\cdot q$ probability of outsider takeover following each period. Second, the multiplier for $V^D(q)+V^E(q)$ that affects the distribution of consumption between D and E is higher. Positive q increases D's bargaining leverage because, as just mentioned, higher q decreases E's expected utility to attempting a coup, which enables D to induce acceptance with a lower offer or causes E to accept with higher probability.

Presenting these considerations in more detail, in equilibrium, $\theta_i\cdot x_i^*(q)=(1-q)\cdot\delta\cdot p_i\cdot\left[V^D(q)-V^E(q)\right]-q\cdot(1-\Delta)\cdot\delta\cdot V^E(q)$. Substituting this term into Equation B.2 enables solving for $V^E(q)$ as a function of $V^D(q)$:

$$V^E(q)=\frac{\delta\cdot(1-q)\cdot p_i}{1-\delta\cdot(1-q)\cdot(1-p_i)}\cdot V^D(q), \tag{B.4}$$

Separately, combining Equations B.1 and B.2 enables calculating:

$$V^D(q)+V^E(q)=\frac{\alpha_i^*(q)}{1-\delta\cdot\Big[1-q\cdot\big(1-\alpha_i^*(q)\cdot(1-\Delta)\big)\Big]} \tag{B.5}$$

Combining Equations B.4 and B.5 yields:

$$V^D(q)=\underbrace{\frac{1-\delta\cdot(1-q)\cdot(1-p_i)}{1-\delta\cdot(1-q)\cdot(1-2p_i)}}_{D\text{'s share of surplus}}\cdot\underbrace{\frac{\alpha_i^*(q)}{1-\delta\cdot\Big[1-q\cdot\big(1-\alpha_i^*(q)\cdot(1-\Delta)\big)\Big]}}_{=V^D(q)+V^E(q)} \tag{B.6}$$

$$V^E(q)=\underbrace{\frac{\delta\cdot(1-q)\cdot p_i}{1-\delta\cdot(1-q)\cdot(1-2p_i)}}_{E\text{'s share of surplus}}\cdot\underbrace{\frac{\alpha_i^*(q)}{1-\delta\cdot\Big[1-q\cdot\big(1-\alpha_i^*(q)\cdot(1-\Delta)\big)\Big]}}_{=V^D(q)+V^E(q)} \tag{B.7}$$

In the expressions for $V^D(q)$ and $V^E(q)$, the first term is each player's percentage share of the total lifetime surplus, and the second term is total lifetime surplus. This shows that as q increases, E gets a smaller percentage of a smaller pie. The $V^D(q)$ and $V^E(q)$ terms immediately enable calculating:

$$V^D(q)-V^E(q)=\underbrace{\frac{1-\delta\cdot(1-q)}{1-\delta\cdot(1-q)\cdot(1-2p_i)}}_{\in(0,1)}\cdot\underbrace{\frac{\alpha_i^*(q)}{1-\delta\cdot\Big[1-q\cdot\big(1-\alpha_i^*(q)\cdot(1-\Delta)\big)\Big]}}_{=V^D(q)+V^E(q)} \tag{B.8}$$

The first term in the previous expression tells us that q affects the distribution of consumption between D and E. The second term tells us that q affects joint consumption. Note that if $q>0$ and $\Delta>0$, then $V^D(q)+V^E(q)\neq V^D+V^E=\frac{1}{1-\delta}$ from the baseline model even if $\alpha_i^*(q)=1$, because there is a positive probability of outsider takeover even if E accepts D's offer in every period.

The analog for Equation 4 in the extension is:

$$\tilde{x}_i^*(q)=\frac{1}{\theta_i}\cdot\frac{(1-q)\cdot\delta\cdot p_i}{1-\delta\cdot(1-q)\cdot(1-2p_i)} \tag{B.9}$$

The analog for Lemma 1 in the extension is:

$$\overline{\theta}(q)\equiv\frac{\delta\cdot p_i\cdot(1-q)}{1-\delta\cdot(1-2p_i)\cdot(1-q)} \tag{B.10}$$

The analog for Equation 5 in the extension is:

$$\tilde{\alpha}_i^*(q) = \frac{\big[1-\delta\cdot(1-q)\big]\cdot\big[1-\delta\cdot(1-2p_i)\cdot(1-q)\big]}{\delta\cdot\Big[(1-\Delta)\cdot q\cdot\big(1-\delta\cdot(1-q)\big)\cdot\theta_i+(1-q)\cdot p_i\cdot\Big[1-\delta\cdot\big[1-q\cdot\big(2\theta_i-\Delta\cdot(1-2\theta_i)\big)\big]\Big]\Big]}\cdot\theta_i \tag{B.11}$$

Lemma B.1 provides the analog of Lemma 2 for the extension. The proof is omitted because it is identical in structure to that of Lemma 2.

Lemma B.1 (Extension: Bargaining with included elite)**.** *Assume a strategy profile in which D shares power with E in every period, and D follows the strategy profile in period t. If D's commitment ability is sufficiently high, then D chooses the unique x_t that satisfies with equality E's acceptance constraint. E accepts offer at least that large with probability 1, and all lower offers with probability 0. In equilibrium, E accepts with probability 1. If D's commitment ability is lower, then D offers $x_t = 1$. E accepts $x_t = 1$ with probability strictly bounded between 0 and 1, and accepts any $x_t < 1$ with probability 0. In equilibrium, E accepts with probability strictly bounded between 0 and 1. Formally, if $\mu_t = \mu^i$ for all t:*

- *If $\theta_i > \overline{\theta}(q)$, then D offers $x_t = \tilde{x}_i^*(q)$, for $\overline{\theta}(q)$ defined in Equation B.10 and $\tilde{x}_i^*(q)$ defined in Equation B.9. E's equilibrium probability of acceptance function is:*

$$\alpha_i^*(x_t;q) = \begin{cases} 0 & \textit{if } x_t < \tilde{x}_i^*(q) \\ 1 & \textit{if } x_t \geq \tilde{x}_i^*(q), \end{cases}$$

  and the equilibrium probability of acceptance is $\alpha_i^*(q) = 1$.

- *If $\theta_i < \overline{\theta}(q)$, then D offers $x_t = 1$. E's equilibrium probability of acceptance function is:*

$$\alpha_i^*(x_t;q) = \begin{cases} 0 & \textit{if } x_t < 1 \\ \tilde{\alpha}_i^*(q) & \textit{if } x_t = 1, \end{cases}$$

  for $\tilde{\alpha}_i^*(q) \in (0,1)$ defined in Equation B.11, and the equilibrium probability of acceptance is $\alpha_i^*(q) = \tilde{\alpha}_i^*(q)$.

Lemma B.2 (Extension: Bargaining after a deviation to exclusion)**.** *Assume a strategy profile in which D shares power with E in every period, but D deviates by excluding E at time t. If D's commitment ability is the high value θ_i, then D chooses the unique x_t that satisfies with equality E's acceptance constraint. E accepts offer at least that large with probability 1, and all lower offers with probability 0. In equilibrium, E accepts with probability 1. If D does not have commitment ability, then D offers any $x_t \in [0,1]$ and E accepts any offer with probability 0. Formally, if $\mu_t = \mu^e$ and $\mu_s = \mu^i$ for all $s > t$:*

- *If $\mu_t = \mu^e$ and $\theta_t = \theta_i$, then D offers $x_t = \tilde{x}_e^*(q)$, for:*

$$\tilde{x}_e^*(q) = \frac{1}{\theta_i}\cdot\frac{(1-q)\cdot\delta\cdot p_e}{1-\delta\cdot(1-q)\cdot(1-2p_i)} \tag{B.12}$$

E's equilibrium probability of acceptance function is:

$$\alpha_e^*(x_t) = \begin{cases} 0 & \textit{if } x_t < \tilde{x}_e^*(q) \\ 1 & \textit{if } x_t \geq \tilde{x}_e^*(q), \end{cases}$$

and the equilibrium probability of acceptance is $\alpha_e^* = 1$.

- *If $\mu_t = \mu^e$ and $\theta_t = 0$, then D offers $x_t \in [0,1]$. E's equilibrium probability of acceptance function is $\alpha_e(x_t) = 0$ for all $x_t \in [0,1]$, which also implies the equilibrium probability of acceptance is $\alpha_e^* = 0$.*

B.2 Power-Sharing Constraint

D's power-sharing constraint in Equation B.13 resembles that from the baseline model (see Equation 6),[22] but the outsider threat generates additional direct and indirect expected costs of exclusion for D. The direct expected cost, expressed by 1 in Equation B.13, arises from the possibility that D's future discounted consumption stream will decrease to 0. This possibility occurs with probability q if E is excluded and probability $\big[1-(1-\Delta)\cdot\alpha_i^*(q)\big]\cdot q$ if E is included. The difference equals $(1-\Delta)\cdot\alpha_i^*(q)\cdot q$, therefore creating a net expected cost of exclusion. One indirect cost, expressed by 2 in Equation B.13, arises because the possibility of outsider takeover weakens E's bargaining leverage under inclusion relative to exclusion. The discussion of E's acceptance constraint in Equation B.3 discussed how (a) attempting a coup raises the probability of outsider takeover and (b) if outsider takeover occurs, then E does not gain the prize of governing in period $t+1$ even if it defeats D in period t. These incentives decrease an included E's expected utility to attempting a coup relative to the case of no outsider threat. Therefore, if D excludes E, then it loses the net gains in bargaining leverage that arise from E's fears of attempting a coup.[23] Three additional indirect effects, expressed by 3, follow from the discussion of Equation B.3 because q affects $V^D(q)+V^E(q)$, $V^D(q)-V^E(q)$, and $\alpha_i^*(q)$.[24] Appendix Propositions B.1 and B.2 define the analogous opportunistic exclusion threshold $\hat{\sigma}(q)$ and strategic exclusion threshold $\underline{\theta}(q)$ as Propositions 1 and 2, respectively.

$$\underbrace{\overbrace{\alpha_i^*(q)}^{3}-\sigma}_{\text{Diff. in Pr(accept)}} + \underbrace{(1-\Delta)\cdot\overbrace{q}^{1}\cdot\overbrace{\alpha_i^*(q)}^{3}}_{\text{Diff. in Pr(outsider takeover)}}\cdot\underbrace{\delta\cdot\overbrace{\big[V^D(q)+V^E(q)\big]}^{3}}_{\text{Cost of outsider takeover}} \geq \underbrace{\overbrace{(1-q)}^{2}\cdot\delta\cdot(p_i-p_e)\cdot\overbrace{\big[V^D(q)-V^E(q)\big]}^{3}}_{\text{Diff. in } E\text{'s bargaining leverage}} \qquad \text{(B.13)}$$

Presenting these considerations in more detail, restating D's inclusion constraint from Equation B.13 by substituting in for the equilibrium continuation values and rearranging yields:

$$\Theta\big(\alpha_i^*(q);q\big) \equiv$$

[22]Equations 6 and B.13 are identical if $q=0$.

[23]There is also a possibility that E will not control the government in the future conditional on winning a fight if E is excluded, but E assigns greater weight to this expected cost of an outsider threat if included rather than excluded because $p_i > p_e$.

[24]Appendix Equations B.5, B.8, and B.11, respectively, present these terms.

$$\alpha_i^*(q)-\delta\cdot\left[(1-q)\cdot(p_i-p_e)\cdot\frac{1-\delta\cdot(1-q)}{1-\delta\cdot(1-q)\cdot(1-2p_i)}-(1-\Delta)\cdot q\cdot\alpha_i^*(q)\right]\cdot\frac{\alpha_i^*(q)}{1-\delta\cdot\left[1-\big(1-(1-\Delta)\cdot\alpha_i^*(q)\big)\cdot q\right]}-\sigma\geq 0 \tag{B.14}$$

A preliminary result shows that $\Theta\big(\alpha_i^*(q);q\big)$ strictly increases in $\alpha_i^*(q)$.

Lemma B.3.

$$\frac{d\Theta\big(\alpha_i^*(q);q\big)}{d\alpha_i^*(q)}>0$$

Proof of Lemma B.3. Computing the derivative and rearranging yields $\frac{d\Theta(\alpha_i^*(q);q)}{d\alpha_i^*(q)}=$

$$\frac{1-\delta\cdot(1-q)}{1-\delta\cdot\left[1-\big(1-(1-\Delta)\cdot\alpha_i^*(q)\big)\cdot q\right]}\cdot$$

$$\left[1-\frac{\delta}{1-\delta\cdot\left[1-\big(1-(1-\Delta)\cdot\alpha_i^*(q)\big)\cdot q\right]}\cdot\left((1-q)\cdot(p_i-p_e)\cdot\frac{1-\delta\cdot(1-q)}{1-\delta\cdot(1-q)\cdot(1-2p_i)}-(1-\Delta)\cdot q\cdot\alpha_i^*(q)\right)\right]$$

It suffices to demonstrate:

$$1>\frac{\delta}{1-\delta\cdot\left[1-\big(1-(1-\Delta)\cdot\alpha_i^*(q)\big)\cdot q\right]}\cdot\left[(1-q)\cdot(p_i-p_e)\cdot\frac{1-\delta\cdot(1-q)}{1-\delta\cdot(1-q)\cdot(1-2p_i)}-(1-\Delta)\cdot q\cdot\alpha_i^*(q)\right]$$

A series of straightforward algebra steps simplifies this to the following inequality, which follows immediately from assumptions about the constituent parameters:

$$1>\delta\cdot(1-q)\cdot\big[1-(p_i+p_e)\big]$$

■

Proposition B.1 (Predatory exclusion with outsider threat)**.** *If the likelihood of a rebellion under exclusion is low enough, then D will exclude E. Formally, if $\sigma>\hat{\sigma}(q)$, for $\hat{\sigma}(q)$ defined in the proof, then there does not exist an equilibrium in which $\mu^*(q)=\mu^i$.*

Proof. Lemma B.3 implies that the left-hand side of Equation B.14 achieves its upper bound if $\alpha_i^*=1$. Substituting $\alpha_i^*=1$ into Equation B.14 shows that D has a profitable deviation if:

$$\sigma>\hat{\sigma}(q)\equiv 1-\delta\cdot\left[(1-q)\cdot(p_i-p_e)\cdot\frac{1-\delta\cdot(1-q)}{1-\delta\cdot(1-q)\cdot(1-2p_i)}-(1-\Delta)\cdot q\right]\cdot\frac{1}{1-\delta\cdot(1-\Delta\cdot q)} \tag{B.15}$$

■

Proposition B.2 (Survival-based exclusion with outsider threat)**.** *Even if D will not opportunistically exclude E, if the probability of a coup under inclusion is high enough, then D will exclude E. Formally, suppose $\sigma<\hat{\sigma}(q)$, for $\hat{\sigma}(q)$ defined in Equation B.15.*

Part a. *For each value of* $q \in (0,1)$*, there exists a unique threshold* $\underline{\alpha}(q)$ *such that there does not exist an equilibrium in which* $\mu^*(q) = \mu^i$ *if* $\alpha_i^*(q) < \underline{\alpha}(q)$.

Part b. *For each value of* $q \in (0,1)$*, there exists a unique threshold* $\underline{\theta}(q) \in \big(0, \overline{\theta}(q)\big)$ *such that* $\alpha_i^* > \underline{\alpha}(q)$ *if* $\theta_i > \underline{\theta}(q)$ *and* $\alpha_i^* < \underline{\alpha}(q)$ *if* $\theta_i < \underline{\theta}(q)$.

Proof of part a. Applying the intermediate value theorem to $\Theta\big(\alpha_i^*(q); q\big)$ defined in Equation B.14 and invoking Lemma B.3 demonstrates that D will not deviate to exclusion if and only if E's probability of acceptance under inclusion is sufficiently high. Because $\Theta(0; q) = -\sigma < 0$, $\Theta(1; q) = \hat{\sigma}(q) - \sigma > 0$ $\big($the statement of the proposition assumes $\hat{\sigma}(q) - \sigma > 0\big)$, and $\Theta(\cdot; q)$ is continuous in $\alpha_i^*(q)$, the intermediate value theorem demonstrates the existence of at least one $\underline{\alpha}(q)$ such that $\Theta(\underline{\alpha}(q); q) = 0$. Lemma B.3 shows that $\Theta(\cdot; q)$ strictly increases in $\alpha_i^*(q)$, which yields the unique threshold claim.

Proof of part b. Substituting $\tilde{\alpha}_i^*(q)$ from Equation B.11 into Equation B.14 yields a somewhat cumbersome term $\Theta\big(\tilde{\alpha}_i^*(q); q\big)$. Define an equivalent term that states θ_i as an explicit parameter: $\tilde{\Theta}\big(\theta_i, q\big) \equiv \Theta\big(\tilde{\alpha}_i^*(q); q\big)$. Applying the intermediate value theorem demonstrates the existence of at least one $\underline{\theta}(q) \in \big(0, \overline{\theta}(q)\big)$ that satisfies $\tilde{\Theta}\big(\underline{\theta}(q), q\big) = 0$.

- $\tilde{\Theta}(0, q) = -\sigma < 0$
- $\tilde{\Theta}\big(\overline{\theta}(q), q\big) = \hat{\sigma}(q) - \sigma > 0$. Equation B.15 defines $\hat{\sigma}(q)$, and the statement of the proposition assumes $\hat{\sigma}(q) > \sigma$.
- $\tilde{\Theta}(\theta_i, q)$ is continuous in θ_i.

The threshold claim follows because $\tilde{\Theta}(\theta_i, q)$ strictly increases in θ_i:

$$\frac{d\hat{\Theta}(\theta_i, q)}{d\theta_i} = \frac{\big[1 - \delta \cdot (1-q)\big] \cdot \big[1 - \delta \cdot \big(1 - (p_i + p_e)\big) \cdot (1-q)\big]}{\delta \cdot p_i \cdot (1-q) \cdot \big[1 - \delta \cdot (1 - \Delta \cdot q)\big]} > 0$$

The explicit characterization of the threshold is:

$$\underline{\theta}(q) \equiv \frac{\delta \cdot p_i \cdot (1-q) \cdot \big[1 - \delta \cdot (1 - \Delta \cdot q)\big] \cdot \sigma}{\big[1 - \delta \cdot (1-q)\big] \cdot \big[1 - \delta \cdot \big(1 - (p_i + p_e)\big) \cdot (1-q)\big]} \tag{B.16}$$

■