Power Sharing with Weak Institutions*

Robert Powell[†]

June 19, 2021

Abstract

Democratic transitions, franchise extensions, and civil-war settlements can often be seen as power-sharing agreements in which opposing factions try to use institutional structures to "lock in" the terms of a settlement. But the commitment power inherent in institutions varies. When institutions are weak, it is difficult for a powerful elite to tie its hands and give up power. This paper studies a window-of-opportunity model in which an enfranchised elite faces a periodic threat. Institutional weakness is parameterized in terms of the elite's marginal return to trying to undermine a power-sharing agreement. The analysis shows that (i) bargaining breaks down if institutions are too weak and why it does; (ii) equilibrium agreements share more power with the opposition when institutions are weak; (iii) there is a non-monotonic relation between power-sharing and how often the opposition poses a threat; and (iv) power-sharing is path dependent.

*I thank Pradeep Chhibber, Ernesto Dal Bó, James Fearon, Mai Hassan, Shachar Kariv, Jack Paine, Daniel Sargent, Chris Shannon and participants at Conference on Armed Conflict at Berkeley and the Riker Seminar at the University of Rochester for very helpful comments, discussion, and criticism.

[†]Travers Department of Political Science, UC Berkeley. RPowell@Berkeley.edu.

Power-Sharing with Weak Institutions

Democratic transitions, franchise extensions, and civil-war settlements can often be seen as power-sharing agreements in which opposing factions try to use institutional structures to "lock in" the terms of a settlement. For example, Acemoglu and Robinson (2000, 2001, 2006) explain the extension of the franchise as resulting from the need of an enfranchised elite to commit to future redistribution in order to appease social unrest. "[E]xtending the franchise acted as a commitment to future redistribution and prevented social unrest. In contrast to democratization, the promise by the elite to redistribute in the future, while maintaining political power, would not have been credible" (2000, 1168).

But the commitment power of institutions varies. Strong institutions may enable an enfranchised elite to commit to intertemporal concessions and thereby avoid a revolution. Weak institutions, by contrast, may not provide the needed commitment power. Fearon and Laitin (2008), for instance, see the inability of the government and an opposition rebel group to commit to sharing power as one of the main reasons why civil wars tend to become all-or-nothing contests. "Parties to civil war settlement negotiations know there is a serious risk that agreements will be violated and power-sharing (whether over a central or regional government) will break down" (2008, 6).

Democratic transitions from authoritarian rule often pose similar commitment problems in conditions of varying institutional strength. Consider, for example, the 2019 uprising in Sudan. Protests over cuts to bread subsidies and other austerity measures broke out in December 2018.[1] The focus of the unrest quickly shifted from economic to political goals with the opposition calling for an end to President Omar al-Bashir's regime. Al-Bashir had survived waves of protest before, in 2011, 2012, and 2013. But he did not survive this one. After continuing protests and mounting pressure, the leaders of Sudan's security forces ousted al-Bashir in April and announced the formation of a Transitional Military Council (TMC) which would oversee a two-year transition to civilian rule. Doubting the credibility of this promise and seeing the leadership of the TMC as

[1] On the Sudan crisis, see BBC (Aug 16, 2019), CRS (2019), Burke and Salih (2019), *Economist* (2019a,b), Gallopin (2019), Hassan and Kodouda (2019), IGC (2019a,b), Lynch (2019), and Walsh (2019a, 2019b, 2019c).

being too close to al-Bashir, the opposition demanded civilian control over the transition. Protests continued and intensified until the TMC and the opposition Forces for Freedom and Choice agreed in August to a thirty-nine month transition to multiparty elections during which members of the TMC would retain significant influence.[2]

Despite the agreement, considerable doubt remained about whether the institutional lock-in would hold and whether there would ultimately be a transfer of power. As a member of the opposition summed it up, the agreement was "a very tough compromise. We just hope that we will achieve a civilian-led government at the end of the three years. And if we fail, we will go back to the street."[3] In short, gambling that the "lock-in" would hold was worth the risk, but it certainly was not seen as a sure thing.

This paper studies the effect of institutional weakness on the ability of an incumbent elite to use power-sharing agreements to buy the opposition off. The key idea motivating the model is that it is difficult for a powerful elite to give up power when institutions are weak. The elite can agree to take steps that on paper may look like a transfer of power. But the elite will have great difficulty tying its hands and actually transferring power when institutions are weak.

More specifically, this paper analyzes an infinite-horizon stochastic game in which an incumbent elite faces an intermittent threat of social unrest and revolution. When faced with unrest, the elite can attempt to buy the opposition off by making an offer combining policy concessions, like higher food subsides or lower taxes, along with institutional power-sharing concessions, like democratizing or expanding the existing franchise. The key distinction between the two is that the former entails no future commitment whereas the latter may. The opposition can accept this offer or rebel. If it accepts, the elite has an opportunity to "invest" in undermining the agreement, possibly during an implementation phase. The weaker the underlying institutions, the higher the marginal return to trying to undermine the agreement.

[2]See http://constitutionnet.org/vl/item/sudan-constitutional-declaration-august-2019 for the text of the agreement.

[3]Sara Abdelgalil of the Sudanese Professional Association quoted in Walsh (Aug 17, 2019). Outside observers also saw eventual civilian rule as uncertain (e.g., ICG 2019, Hassan and Kodouda 2019).

The study makes four main contributions. First, the analysis shows that bargaining breaks down if institutions are too weak and why it does. The weaker institutions are, the higher the elite's incentive to invest in undermining an agreement and the less likely the agreement is to hold. Weaker institutions thus lower the opposition's payoff to any proposed settlement. If they are too weak, the incumbent is unable to buy the opposition off. Even the promise of complete control is insufficient in expectation, and the opposition rebels.

Second, because weaker institutions mean that any particular power-sharing agreement is less likely to hold, the elite has to offer better terms if it is to induce the opposition to forego the option of rebelling. As a result, equilibrium agreements share more power with the opposition when institutions are weak. Conversely, the elite shares less power when institutions are strong. This is consistent with Albertus and Menaldo's (2018) empirical findings about democratic transitions from authoritarian rule. Although the mechanism studied here is quite different from theirs, they find a positive relationship between state capacity and "gamed constitutions" which protect the interests of the outgoing authoritarian elite.

This result also helps to explain a puzzling feature of many power-sharing agreements. As was the case in Sudan, the opposition often agrees to a settlement knowing full well that the incumbent may try to undermine it. Why does the opposition agree, especially if agreeing means foregoing the opportunity to bring the government down while the opposition is temporarily strong? As the analysis emphasizes and clarifies, weaker institutions make the gamble of a power-sharing agreement riskier. Nevertheless, better terms may still make this a gamble worth taking.

Third, seeing a power-sharing agreement as a risky gamble also highlights a non-monotonic relation between power-sharing and how often the opposition poses a threat. When the opposition poses a constant or frequent threat, there is no commitment problem and no power sharing. The elite buys the opposition off with policy concessions alone. When the threat of unrest is less frequent, the elite faces a commitment problem that it can only solve by sharing power with the opposition. This is the standard result from the intermittent-threat, window-of-opportunity models pioneered by Acemoglu and Robinson

(2000, 2001, 2006) in the context of democratization and Fearon (2004) in explaining why some civil wars last so long.[4] But there is more to this relationship when the elite can invest in undermining an agreement. As the threat becomes still less frequent, the elite's payoff to undermining a power-sharing agreement grows, the elite invests more, and the chances that an agreement will hold go down. Eventually the chances that an agreement will hold are so small that there is again no power-sharing, now because the opposition is unwilling to accept any power-sharing proposal.

Finally, extending the model to allow for a moderate threat in addition to a severe threat shows that power-sharing is path dependent. How much power the elite ultimately shares and whether the transition ends in an agreement or rebellion generally depends on the pattern of threats the elite faces over time. If the first threat the elite happens to face is severe, the elite will have to share a great deal of power with the opposition in order to buy it off. Indeed, if institutions are too weak, the elite may not be able to buy the opposition off and the threat will end in rebellion. If, by contrast, the elite happens to face a series of gradually intensifying threats, it will gradually share more and more power though ultimately less power than if the first threat had been severe.[5]

The key to this path dependence is that the elite's stakes in undermining an agreement when it faces a severe threat are higher if the elite is still in complete control of the state than the stakes would be if the elite had previously faced a moderate threat and had already shared some power with the opposition. These higher stakes induce the elite to invest more in undermining an agreement. The opposition anticipates this, and the elite must offer more when the stakes are higher. The effect of this is that the elite will generally share more power with the opposition if it is still in complete control when it faces a severe threat.

The next section discusses related work. The two subsequent sections present the model, characterize its equilibria, and establish necessary and sufficient conditions that

[4] Powell (2004) shows that the commitment problems in models of political transitions and civil- and inter-state war are fundamentally alike.

[5] Acemoglu, Egorov, and Sonin (2015) obtain related results where the steady-state distribution of political power depends on the pattern of shocks. This is discussed further below. On path dependence more generally, see Page (2006).

ensure that a power-sharing equilibrium exists. The fifth section presents the main results: establishing existence when institutions are strong, non-existence when they are weak, and describing the equilibrium comparative statics. A final section extends the model to allow for a moderate threat.

Related Work

Three elements define the specific focus of the present study: (i) an elite that faces periodic threats which may create commitment problems; (ii) an endogenous choice about the mix of policy and institutional (power-sharing) concessions to be offered to the opposition when it poses a threat, and (iii) the elite can try to undermine any agreement where the efficacy of any given level of effort depends on the strength of the institutions.[6] The research most directly related to these elements is that growing out of Acemoglu and Robinson's work on political transitions and extending the franchise. Indeed, Acemoglu and Robinson (2000, 2001, 2006) formalizes (i) and (ii) and provides a point of departure for this analysis.[7]

One of their key contributions is explaining the nature of the concessions that the elite makes and, more specifically, why the elite sometimes shares power by extending the franchise. As noted above, there is no credibility problem when the threat of unrest in any period is sufficiently high. The elite in these circumstances prefers to retain power, does

[6]Recent work on democratic transitions or franchise extensions with a broader or different focus includes Haggard and Kaufman (2016); Ziblatt (2017); Albertus and Menaldo (2018); Geddes, Wright, and Frantz (2018); Fearon and François (2020); Reidl et al. (2020); and Triesman (2020). On civil wars and, especially the role of commitment problems and institutional power-sharing settlements, see Wantchekon (2000), Walter (2002, 2009), and Hartzell and Hoddie (2003, 2007, 2020).

[7]There is now substantial empirical support linking *some* cases of democratization and franchise extensions to adverse economic shocks or unrest and the threat of revolution. See, for example, Przworski (2008), Burke and Leigh (2010), Bruckner and Ciccone (2011), Adit and Jensen (2014), Adit and Franck (2015), Aidt and Leon (2015), and Dower et al. (2018). There is however much less empirical work showing frequent shocks lead to policy concessions and infrequent shocks lead to institutional concessions intended to lock-in an agreement. An exception is de Figueiredo (2002). Lizzeri and Persico (2004), Haggard and Kaufman (2012, 2017), Ziblatt (2017), and Triesman (2020), among others, argue that the window-of-opportunity framework does not fit many cases. See also Acemoglu et al. (2013).

not extend the franchise, and buys the opposition of with policy concessions alone. By contrast, the elite shares power by extending the franchise when the threat of unrest in any period is sufficiently small. The payoff to getting large but infrequent policy concessions is now too small to buy the poor off. The only way that the elite can give the poor enough to avoid a revolution is to extend the franchise which locks in these concessions.

Acemoglu and Robinson's baseline model assumes that the elite is fully in control or that there is complete democracy and a poor person is the pivotal decisionmaker. But the scope of the franchise is at least partly a matter of choice. Dower et al. (2019, 2020) extend the basic window-of-opportunity model by endogenizing the degree of liberalization in their study of Czarist Russia.[8] The elite when first confronted with unrest can make a once-and-for-all decision to liberalize by choosing the probability that the poor will be pivotal and set policy in any future period. The elite sets future policies with the complementary probability. Liberalization thus lets the elite commit in expectation to any policy between what the poor would set if they were sure to be pivotal and what the elite would do.

Fearon and François (2020) also treat democratization as an endogenous choice in their analysis of some of the commitment problems inherent in elite-led transitions from authoritarian rule. These transitions frequently involve "gamed" constitutions which contain provisions that protect the elite's interests by, for example, granting the elite immunity, significant control over the military and the economy, disproportionate representation, or an effective veto of legislation and constitutional reform (Albertus and Menaldo 2018, 80-97). Why would a newly empowered democratic opposition honor this kind of a deal once the elite leaves office? How does the opposition commit to uphold the agreement?

Fearon and François develop the idea that commitment comes through power sharing. They argue that democratization at its core is often a power-sharing agreement in which the elite retains *de facto* control over enough of the state that the elite prefers this arrangement to a revolution. More formally, Fearon and François model the state as a divisible asset which throws off a flow of benefits. A power-sharing agreement divides the asset between the elite and democratic opposition so that neither the elite nor the

[8] Acemoglu and Robinson discuss endogenous extension informally (2006, 209-11).

opposition has an incentive to renege.

Jack and Lagunoff (2006); Acemoglu, Egorov, and Sonin (2012, 2015, 2020); and Gieczewski (2021) provide more microfounded models of franchise extension in which a currently empowered group chooses today's policy and who will be empowered tomorrow.[9] Acemoglu, Egorov, and Sonin (2015), in particular, study the effects of shocks to the distribution of power. The main focus of those models is on institutional evolution, and there is no explicit bargaining between a currently enfranchised group and a temporarily strong outside group.

Acemoglu and Robinson (2008) endogenize *de facto* power. In every period, citizens and members of the elite decide how much to invest in supporting or undermining the current institution. The more a group invests in its power in period t, the more likely it is to gain power and be able to set period t's economic policies and determine whether the regime will be democratic or nondemocratic at $t+1$. Democracy also provides a boost to the citizens' de facto power. If this advantage is not too large, then an "invariance" result may hold: The elite in equilibrium may invest so much more in democracy than it does in autocracy that the elite completely offsets the democratic advantage. As a result, the elite's probability of coming to power in the next period is the same regardless of whether the current regime is democratic or not.

The present study builds on and complements existing work. Taking a widow-of-opportunity setup as a point of departure, this analysis treats the division of power as a continuous choice with the elite choosing the optimal mix of policy and institutional concessions to offer. As in Fearon and François (2020), the elite and opposition are formally dividing a divisible asset and the flow of benefits off that asset. The emphasis here, however, is on the elite's inability to commit to a power-sharing agreement rather than on the opposition's inability to commit.[10] As in Acemoglu and Robinson (2008), the elite chooses how much to invest in shaping the future distribution of power. Their

[9] Jack and Lagunoff focus explicity on endogenous franchise extension. The others are more general but can be interpreted in terms of franchise extension.

[10] If the elite is in control of the state as modeled here, its offer holds the opposition down to its reservation value and the commitment issues Fearon and François study do not arise.

main focus is on the dynamics of regime switching between democracy and nondemocracy. There is no bargaining between the elite and opposition, and the elite never has the option of trying to buy the opposition off with a mix of policy and institutional concessions. By contrast, the emphasis here is on how variation in institutional strength affects the choice between policy and institutional concessions, the ability of the elite to buy the opposition off, and the ultimate division of power when it can buy the opposition off.

A Model of Power Sharing

Consider a situation in which an elite is initially in complete control of the state and periodically faces a revolutionary threat from an opposition group. The motivating idea that it is difficult for a powerful elite to give up power when institutions are weak is modeled as an infinite-horizon stochastic game played in discrete time in which the elite, E, and opposition, O, repeatedly bargain over a divisible asset. The asset throws off a flow of benefits equal to one in each period. An actor who controls a given percentage of the asset controls that percentage of the flow of benefits. Let $f_t \in [0, 1]$ be the share of the asset the opposition controls at the start of round $t = 0, 1, 2, \ldots$ where $f_0 = 0$ because the elite is assumed to control the state at the outset. We can think of f_t as the opposition's share of power or the extent of the franchise. (This interpretation and other aspects of the model are discussed below.)

The state of the stochastic game at time t is defined by the pair (f_t, ν_t) where $\nu_t \in \{s, n\}$ is the strength of the opposition's threat. The opposition is strong and poses a revolutionary threat when $\nu_t = s$. It is weak and poses no immediate threat when $\nu_t = n$. One reason why the opposition might be strong is that an adverse economic shock reduces the cost of rebelling. The probability that the opposition poses a revolutionary threat in a given period is r.

A round is "active" when the opposition poses an immediate threat and "inactive" when the threat is latent. Figure 1 depicts the sequence of moves in an active state when the opposition is strong and the existing franchise is f_t. The elite can try to buy the opposition off by making an offer (y_t, ϕ_t) to the opposition where $y_t \in [f_t, 1]$ is the policy component and $\phi_t \in [f_t, 1]$ is the institutional power-sharing component. The lower

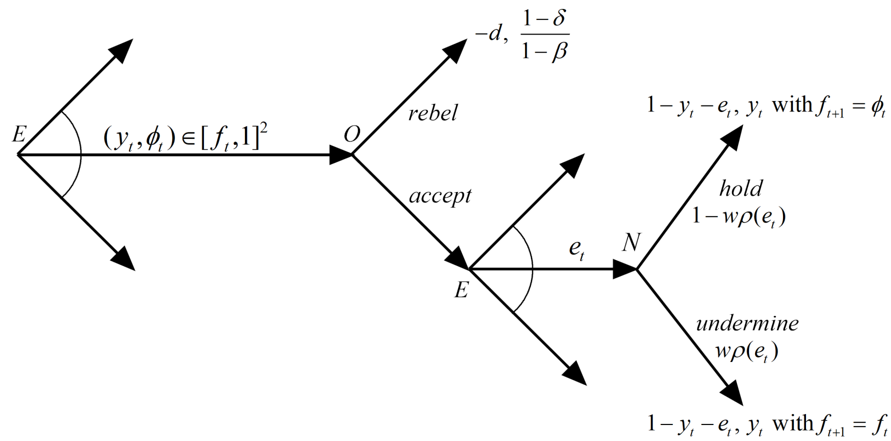

Figure 1: The stage game when the division of power is f_t and the opposition is strong.

bound on y_t and ϕ_t reflects the fact that the opposition already controls f_t of the asset, so that $y_t - f_t \geq 0$ is the additional amount of today's flow that the elite is offering to the opposition, and $\phi_t - f_t \geq 0$ is the share of the asset on offer. To emphasize the distinction between policy and institutional concessions, note that if the proposed division of power moves to ϕ_t, i.e., if $f_{t+1} = \phi_t$, then the opposition gets at least ϕ_t in all future periods. If, by contrast, the elite had bought the opposition off with policy concessions alone so that the distribution of power was unchanged ($f_{t+1} = f_t$), then the opposition's future payoffs would be at least f_t.

The opposition can accept the elite's offer or rebel. The latter ends the game with the opposition gaining complete control of the asset.[11] Fighting, however, destroys a faction δ of the flow off the asset, so the opposition's payoff to rebelling is $1 - \delta$ in every round thereafter or a total payoff $(1 - \delta)/(1 - \beta)$ where β is the actors' common discount factor. The elite's payoff to being deposed is $-d < 0$.

If the opposition accepts, the elite can to try to undermine the institutional concessions promised in the proposal. If the elite invests effort e_t in undoing the power-sharing agreement, it suffers a disutility of $-e_t$ and succeeds with probability $\min\{w\rho(e_t), 1\}$.[12] Success prevents the transfer of power and the opposition's share remains unchanged, i.e.,

[11] A straightforward generalization of the model is to assume that the opposition succeeds with probability $\overline{\pi}$ when strong and probability $\underline{\pi} < \overline{\pi}$ when weak.

[12] Because effort is costly, the elite never invests more than $\rho^{-1}(1/w)$, and we will write the probability of undermining an agreement as $w\rho(e)$ to ease the exposition.

$f_{t+1} = f_t$. If the elite's efforts to prevent the transfer of power fail and the institutional concessions hold, then the opposition's share in the next round will be ϕ_t, i.e., $f_{t+1} = \phi_t$.[13]

The function ρ is taken to be strictly increasing and strictly concave with $\rho(0) = 0$. We also assume $\lim_{e \to 0} \rho'(e) = \infty$ and $[\rho'(e)]^3/\rho''(e)$ is bounded at zero. If, for example, $\rho(e) = e^\lambda$, then these assumptions imply $\lambda \in [1/2, 1)$. The substantive import of the first assumption is that the elite always exerts positive effort to undo an agreement whenever it prefers the status quo to the agreement. The second assumption ensures that the marginal increase in the probability of undoing an agreement induced by sharing slightly more power is bounded at zero (see footnote 29).

The parameter $w \geq 0$ formalizes the underlying institutional strength, i.e., the extent to which the elite and opposition can use institutions to make credible commitments. When $w = 0$, the probability that the elite can undermine an agreement is zero regardless of how much effort the elite puts into doing so. A promise to give up power is certain to hold. This is the baseline assumption in many models of democratization and power-sharing where an offer to extend the franchise is sure to hold if the opposition accepts. Institutions are weaker and provide less commitment power when w is larger. More specifically, the larger w, the higher the marginal return to investing in undermining a power-sharing agreement, $w\rho'(e)$, and the more likely it is that the agreed transfer of power will not be realized, i.e., the higher $w\rho(e)$. When w is very large, institutions provide little commitment power.

The round ends after the elite decides on e_t. Payoffs are realized, and play moves to the next round. The elite and opposition respectively get $1 - y_t - e_t$ and y_t regardless of whether the elite's efforts to undermine the institutional concessions were successful or

[13] This is a kind of putty-clay assumption about institutional change: Once the division of power f_t takes hold it cannot be undone. At the expense of much greater complexity, we could assume that if the elite successfully undoes an agreement when the existing division of power is f_t, then the resulting division of power at $t + 1$ is distributed over $[0, f_t]$ according to the cumulative distribution function $G(f|f_t)$ where $G(f|f'_t)$ first-order stochastically dominates $G(f|f_t)$ if $f'_t > f_t$.

The key feature of this formulation, as with the simpler putty-clay assumption, is that the elite's expected gain to successfully undoing an agreement on ϕ is higher when the stakes are higher, i.e., at f_t rather than f'_t. That different stakes induce different levels of effort leads to path dependence in the extended game described below.

not. This formalizes the assumption that the actors can commit to the division of today's pie but not to the division of the future flow of benefits benefits.

If the opposition is weak and poses no immediate threat at t, the round is inactive and there are no strategic actions. The elite gets the status-quo payoff $1 - f_t$, the opposition gets f_t, and the distribution of power is unchanged ($f_{t+1} = f_t$).[14]

It is useful to elaborate on two aspects of the model. The first is the timing of moves in the active stage game. If the elite succeeds in undermining an agreement, the opposition cannot respond until the next period when there is some chance that it will be weak and pose no threat. This is a reduced-form way of modeling the idea that agreements are fundamentally gambles when institutions are weak. If, alternatively, the model assumed that the opposition was sure to be strong when it learned an agreement had been undone, the opposition could simply rebel then. Anticipating this, the elite would have no incentive to invest in undermining an agreement. The threat of revolution would effectively enforce implementation of the agreement.

More generally, long transition periods, like the one in Sudan, present a subtle challenge to window-of-opportunity commitment problems if the opposition is assumed to remain strong throughout the transition. If the opposition remains strong for too long, there is no commitment problem. The elite can buy the opposition off with policy concessions alone and does not have to share any power. Assuming the opposition is strong throughout a long transition thus necessitates a delicate balance: The opposition must be strong (in expectation) for long enough to enforce the agreement but not long enough to eliminate the commitment problem. By contrast, assuming that the opposition may be weak when it learns that the agreement has been undone finesses the need for this delicate balance.

Second, it is also useful to elaborate on modeling power sharing in terms of dividing an asset. This approach provides a simple, albeit highly reduced-form, way of distinguishing policy concessions from institutional concessions. Each actor controls the flow off its share of the asset and can costlessly allocate it as it pleases. The only way that another actor

[14]Assuming the round is inactive when the opposition is weak simplifies the characterization of the equilibrium strategies.

can directly affect this allocation is through some costly action that affects the distribution of the asset. This reduced-form assumption provides a tractable point of departure for endogenizing the elite's choice between policy and institutional concessions.

The Equilibrium

An equilibrium is "power-sharing" if the elite and opposition agree on a division of power and avoid a revolution. This section identifies necessary and sufficient conditions needed to ensure that a Markov perfect power-sharing equilibrium exists. The emphasis here is on the main ideas and intuitions. Proofs are in online Appendix A.

The equilibrium path turns out to be very simple. There is a threshold $\tilde{f}$ such that the elite can and does buy the opposition off solely with policy concessions whenever it is facing an immediate threat and $f_t \geq \tilde{f}$. The elite faces a commitment problem when $f_t < \tilde{f}$ and cannot buy the opposition off with policy concessions alone. Rather, the elite offers the maximal policy concession ($y_t = 1$) and shares as little power as possible, say $\phi^*(f_t)$. We show that $\phi^*(f_t) \geq \tilde{f}$ which means that sharing $\phi^*(f_t)$ eliminates the commitment problem. Once the division of power moves to $\phi^*(f_t)$, the elite buys the opposition off with policy concessions alone and the division of power remains at $\phi^*(f_t)$.

The analysis focuses on Markov perfect equilibria (MPEs). A strategy is Markov if an actor always takes the same action whenever it is in same state regardless of what has happened in the past. A Markov strategy for the elite specifies the policy offer $y(f)$ and the proposed division of power $\phi(f)$ as functions of the existing distribution of power. The elite's strategy also specifies how much the elite invests in undermining any agreement (y, ϕ) as a function of the existing division of power and the agreement. Call this effort $e(y, \phi, f)$. A Markov strategy for the opposition stipulates whether the opposition would accept or reject an offer as a function of the current state. A strategy profile is Markov if the elite's and opposition's strategies are Markov. An MPE is a Markov profile that is subgame perfect.

A key property of Markov profiles is that what happens at time t only affects future actions to the extent that actions at t affect the state at $t + 1$. We exploit this fact to solve for the actions in the active state (f, s) by "backwards induction" from the end of the stage game. We first determine the elite's optimal effort e^* given the existing division of power f and the accepted proposal. Anticipating that the elite will invest e^* in undermining the proposal and succeed with probability $w\rho(e^*)$, the opposition can calculate its expected payoff to agreeing. It accepts if this payoff is at least as good as the payoff to fighting. Finally, the elite determines its optimal offer in light of what it will take to buy the opposition off.

More formally, let σ be any Markovian profile, and take $V_E(f,\nu|\sigma)$ and $V_O(f,\nu|\sigma)$ to be the elite's and opposition's continuation payoffs to following σ starting from the beginning of the stage game in state (f,ν). Define $V_E(f|\sigma)$ to be the elite's expected payoff at division of power f before knowing the opposition's strength, i.e., $V_E(f|\sigma) \equiv (1-r)V_E(f,n|\sigma) + rV_E(f,s|\sigma)$ and similarly for $V_O(f|\sigma)$. To simplify the notation, we drop the argument σ when the meaning is clear.

Let σ^* be a power-sharing, pure-strategy MPE. We work backwards from the elite's investment decision to determine what the elite and opposition actually do in σ^*. The elite's payoff to investing e in undermining agreement (y,ϕ) given that play subsequently follows σ^* is

$$\gamma(e|(y,\phi),(f,\nu),\sigma^*) = 1 - y - e + \beta w\rho(e)V_E(f|\sigma^*) + \beta[1 - w\rho(e)]V_E(\phi|\sigma^*) \tag{1}$$

where, recall, y and ϕ must be at least as large as f. In words, the elite gets $1 - y - e$ regardless of whether the agreement holds. The elite succeeds in undermining the power-sharing agreement with probability $w\rho(e)$ and gets the discounted payoff to the division remaining at f. The proposal is implemented with probability $1 - w\rho(e)$, and the elite gets the discounted payoff to power-sharing division ϕ.

In equilibrium, the elite must invest the level of effort that maximizes γ. Differentiation gives $\partial\gamma/\partial e = -1 + \beta w\rho'(e)[V_E(f) - V_E(\phi)]$ where, recall, future play and hence the payoff $V_E(\phi)$ is unaffected by the elite's effort in the current period. The concavity of ρ ensures that the optimal effort is unique. If there is nothing to be gained by undermining the agreement, i.e., if $V_E(f) - V_E(\phi) \leq 0$, then the marginal gain of exerting effort is negative ($\partial\gamma/\partial e < 0$) and the optimal effort is $e^*(f,\phi) = 0$. If there is something to be gained, the optimal effort satisfies the first-order condition

$$1 = \beta w \rho'(e^*(f, \phi))[V_E(f) - V_E(\phi)] \tag{2}$$

where we assume an interior solution to ease the exposition.[15]

As (1) and (2) make clear, the optimal level of effort depends on the current division of power f and the proposed division of power ϕ but not on the policy component y. The current and proposed divisions of power set the stakes for the elite who will either get $V_E(f)$ or $V_E(\phi)$. The elite decides how much effort to invest in undermining the agreement in light of these stakes. The assumption that $\lim_{e \to 0} \rho'(e) = \infty$ ensures that the elite exerts at least some effort to undermine an agreement whenever it has something to lose. Lemma 1 summarizes these results:

LEMMA 1. *Let σ^* be a power-sharing MPE. Then the elite invests effort $e^*(f, \phi)$ following offer (y, ϕ) where $e^*(f, \phi) = 0$ if $V_E(f) - V_E(\phi) \leq 0$ and $e^*(f, \phi)$ uniquely solves $1 = \beta w \rho'(e^*)[V_E(f) - V_E(\phi)]$ if $V_E(f) - V_E(\phi) > 0$ and there is an interior solution.*

Turning to the opposition's decision, it anticipates that the elite will invest effort $e^*(f, \phi)$ in undermining agreement (y, ϕ) at division of power f. Consequently, the opposition is only willing to accept offers satisfying the no-revolution constraint

$$\frac{1 - \delta}{1 - \beta} \leq f + y + \beta w \rho(e^*) V_O(f) + \beta[1 - w\rho(e^*)] V_O(\phi) \tag{3}$$

where the expression on the right is the opposition's expected payoff to agreeing.[16]

Now consider the elite's proposal. There are two cases to consider depending on whether there actually is a commitment problem. Suppose that the opposition controls f and that the elite makes the maximal policy concession of $y = 1$ whenever the opposition is strong. The opposition weakly prefers this to rebelling when

[15] Result 1 in Appendix A discusses corner solutions and boundary conditions.

[16] To be more precise, the opposition is sure to accept offers it strictly prefers to rebelling. As elaborated below (see footnote 17), it may not accept some out-of-equilibrium offers it weakly prefers to fighting.

$$\frac{1-\delta}{1-\beta} \leq 1 + \frac{\beta[(1-r)f + r]}{1-\beta}$$

$$f \geq \tilde{f} \equiv \frac{\beta(1-r) - \delta}{\beta(1-r)}. \tag{4}$$

It follows that there is no commitment problem when $f \geq \tilde{f}$. The elite can buy the opposition off with policy concessions alone and does not need to share any more power. To focus on the substantive problem of interest, assume that there is a commitment problem when the elite is in complete control of the state, i.e., when $f = 0$.

Assumption 1. *There is a commitment problem when the elite is fully in control of the state:* $\tilde{f} > 0$ *or, equivalently,* $r < (\beta - \delta)/\beta$.

When $f \geq \tilde{f}$, the elite does not have to share any additional power. There are, however, two subcases to consider. Suppose that $f > 1 - \delta$ where the definition of $\tilde{f}$ implies $1 - \delta > \tilde{f}$. Then, the opposition's payoff to getting f in every round is already greater than its payoff to fighting. The opposition cannot credibly threaten to fight in these circumstances, and the elite makes no additional concessions when the opposition is strong.

If $\tilde{f} \leq f \leq 1 - \delta$, the opposition poses a credible threat when strong. In response, the elite makes the smallest possible policy concession needed to buy the opposition off without offering to share any additional power ($\phi = f$). This avoids the deadweight loss of exerting effort to undermine an agreement. To specify this offer, let $x(f)$ be the offer that solves

$$\frac{1-\delta}{1-\beta} = x(f) + \frac{\beta[(1-r)f + rx(f)]}{1-\beta}$$

$$x(f) = \frac{1 - \beta(1-r)f - \delta}{1 - \beta(1-r)}.$$

The expression on the right of the first equality is the opposition's payoff to getting $x(f)$ in the current period plus the expected payoff to getting f when weak and $x(f)$ when strong in all future periods. It follows that the elite offers $y(f) = \max\{f, x(f)\}$ and $\phi = f$

for $f \in [\tilde{f}, 1]$. Lemma 2 summarizes these results.

LEMMA 2. *When $f \geq \tilde{f}$, there is no commitment problem. The elite can buy the opposition off with policy concessions alone and does so. The elite proposes $y^*(f) = \max\{f, x(f)\}$ and $\phi^*(f) = f$ along the unique equilibrium path, the opposition accepts, and the elite does not try to undermine the agreement, $e^*(f|f) = 0$. The continuation payoffs when the opposition is strong are $V_O(f, s) = \max\{f, 1 - \delta\}/(1 - \beta)$ and $V_E(f, s) = 1/(1 - \beta) - V_O(f, s)$.*[17]

Proof: See the appendix.

Now consider the elite's offer when there is a commitment problem ($f < \tilde{f}$) and the opposition poses an immediate threat. The elite cannot buy the opposition off with policy concessions alone and shares as little power as possible. That is, the elite makes the maximal policy concession of $y = 1$ and proposes the smallest division of power ϕ needed to satisfy the no-revolution constraint.

To establish this, let $v_E(y, \phi|f, \sigma^*)$ be the elite's payoff to offering (y, ϕ) at (f, s) given that this offer is accepted and that play subsequently reverts to σ^* where (y, ϕ) need not be an equilibrium offer. Then

$$v_E(y, \phi|f, \sigma^*) = 1 - y - e^*(f, \phi) + \beta w\rho(e^*(f, \phi))V_E(f) + \beta[1 - w\rho(e^*(f, \phi))]V_E(\phi). \quad (5)$$

The elite's equilibrium proposal in a power-sharing equilibrium must satisfy two conditions. First, it must maximize v_E subject to satisfying the no-revolution constraint. Second, the elite must weakly prefer making this offer to triggering a revolution by making an offer the opposition is sure to reject.

Maximizing v_E means buying a "yes" at the cheapest possible price. That is, the elite holds the opposition down to its reservation value so that the no-revolution constraint binds.[18] Using (3) to eliminate y, the elite's payoff along the binding no-revolution

[17] The opposition also accepts any offers that it strictly prefers to fighting. There are, however, many off-the-equilibrium-path offers (y, ϕ) at which the no-revolution constraint binds. The opposition is indifferent between fighting and accepting these offers, and this generates multiple equilibria with the same equilibrium path.

[18] More formally, v_E is decreasing in y and ϕ, so the elite could profitably deviate by lowering y or ϕ if the no-revolution constraint is slack. See Result 4 in Appendix A.

constraint is

$$
\begin{aligned}
b_s(\phi) = 1 - \frac{1-\delta}{1-\beta} &- e^*(f,\phi) + \beta w \rho(e^*(f,\phi))[V_E(f) + V_O(f)] \\
&+ \beta[1 - w\rho(e^*(f,\phi))][V_E(\phi) + V_O(\phi)]
\end{aligned} \tag{6}
$$

Differentiating and using the first-order condition $1 = \beta w \rho'(e^*)[V_E(f) - V_E(\phi)]$ to simply the derivative gives

$$
\frac{\partial b_s(\phi)}{\partial \phi} = -\beta r w \rho'(e^*) \left[V_O(\phi) - V_O(f)\right] \frac{\partial e^*}{\partial \phi} + \beta \left[1 - w\rho(e^*)\right] \frac{\partial (V_E(\phi) + V_O(\phi))}{\partial \phi}. \tag{7}
$$

This expression is negative and, hence, the elite shares as little power as possible. To establish that $\partial b_s(\phi)/\partial \phi < 0$, observe first that the proposed division of power ϕ must also be at least as large as $\tilde{f}$ for the proposal (y, ϕ) to satisfy the no-revolution constraint (see Result 5 in the appendix). More intuitively, recall that $\tilde{f}$ is the smallest f such that the elite can buy the opposition off by making the maximal policy concession and not conceding any power. If $\phi < \tilde{f}$, then (y, ϕ) for any $y \in [f, 1]$ will not satisfy the no-revolution constraint. Thus, ϕ must be at least as large as $\tilde{f}$.

Lemma 2 and $\phi \geq \tilde{f}$ imply that the elite buys the opposition off with policy concessions alone and no longer invests in undermining agreements once the division of power moves to ϕ. From this point on, there is no deadweight loss from effort, and the sum of the actors' continuation values is the total flow of benefits, i.e., $V_E(\phi, s) + V_O(\phi, s) = 1/(1-\beta)$. Hence, the second term in (7) is zero.

To sign the first term, note that the opposition's continuation payoff in any inactive period is $V_O(f, n) = f + \beta[rV_O(f, s) + (1-r)V_O(f, n)]$ for any $f \in [0, 1]$. Algebra gives

$$
V_O(f, n) = \frac{f + r\beta V_O(f, s)}{1 - \beta(1-r)}. \tag{8}
$$

$$
V_O(f) = \frac{(1-r)f + rV_O(f, s)}{1 - \beta(1-r)}. \tag{9}
$$

Similarly, the elite's continuation values are

$$V_E(f,n) = \frac{1 - f + r\beta V_E(f,s)}{1 - \beta(1-r)} \tag{10}$$

$$V_E(f) = \frac{(1-r)(1-f) + rV_E(f,s)}{1 - \beta(1-r)}. \tag{11}$$

The sign of the first term in (7) now follows. The elite holds the opposition down to its reservation value (see Result 4), so $V_O(f,s) = (1-\delta)/(1-\beta)$. Moreover, $V_O(\phi,s) = \max\{\phi, 1-\delta\}/(1-\beta)$ by Lemma 2, so $V_O(\phi) - V_O(f) > 0$. The appendix also shows that the effort the elite invests in undermining an agreement increases with the stakes, i.e., $\partial e^*(f,\phi)/\partial\phi > 0$. (See Result 6.) Thus, the first term is negative, and the elite's payoff along the binding no-revolution constraint is decreases in the amount of power the elite shares.

A direct consequence of sharing as little power as possible is that the elite makes the maximal policy concession of $y^*(f) = 1$ when there is a commitment problem. The last step in specifying the equilibrium is to pin down the equilibrium institutional concession $\phi^*(f)$. When $y = 1$, ϕ must satisfy $1 + \beta V_O(f) + \beta[1 - w\rho(e^*(f,\phi))][V_O(\phi) - V_O(f)] = (1-\delta)/(1-\beta)$. Using $V_O(f,s) = (1-\delta)/(1-\beta)$, $V_O(\phi,s) = \max\{\phi, 1-\delta\}/(1-\beta)$, and (9), we can rewrite the binding no-revolution constraint as

$$\phi = \begin{cases} \tilde{f} + \left(\frac{w\rho(e^*(f,\phi))}{1-w\rho(e^*(f,\phi))}\right)(\tilde{f} - f) & \text{if } w\rho(e^*(f,\phi)) \leq \frac{1-\delta-\tilde{f}}{1-\delta-f} \\ \\ \frac{\beta-\delta}{\beta} + \left(\frac{(1-\beta)(1-r)}{1-\beta(1-r)}\right)\left(\frac{w\rho(e^*(f,\phi))}{1-w\rho(e^*(f,\phi))}\right)(\tilde{f} - f) & \text{if } w\rho(e^*(f,\phi)) > \frac{1-\delta-\tilde{f}}{1-\delta-f}. \end{cases} \tag{12}$$

The dependence of ϕ on the existing division of power f highlights a key point. The more severe the commitment problem, i.e., the larger $\tilde{f} - f$, the better the terms ϕ needed to buy the opposition off. To gain some intuition, suppose the elite only had to offer $\tilde{f}$ to resolve the commitment problem as Lemma 2 might be thought to suggest. To see that the elite cannot just offer $\tilde{f}$, observe that the better the status quo is for the elite (the smaller f), the higher the elite's stakes in undermining a proposal to give the opposition $\tilde{f}$. As f decreases and the stakes grow, the elite invests more in undermining the agreement on $\tilde{f}$. This in turn reduces the expected payoff to agreeing to $\tilde{f}$. As a result, the elite has to offer better terms ($\phi > \tilde{f}$) in order to get the opposition to agree. In the case of Sudan's transition, the greater the doubt that the transition phase would end in civilian control, the more the opposition would have to be offered to get them to accept.

Finally, recall that the elite's payoff to buying the opposition off must be at least as good as its payoff to being deposed. That is, the total cost of the elite's efforts to sustain f cannot exceed the cost of being deposed. Formally, $V_E(f, s) \geq -d$ where the binding no-revolution constraint implies

$$V_E(f, s) = \frac{\delta}{1-\beta} - \frac{[1-\beta(1-r)]e^*(f,\phi)}{1-\beta+\beta r[1-w\rho(e^*(f,\phi))]}. \tag{13}$$

This expression is derived in the appendix (see Result 7). For some intuition, observe that the elite gives $(1-\delta)/(1-\beta)$ to the opposition when faced with an immediate threat. Consequently, the elite's payoff is bounded above by $\delta/(1-\beta)$ which is the first term in (13). As for the second, the elite invests e^* whenever there is an immediate threat for as long as the division of power remains at f. Eventually the elite's efforts to undermine the proposal fail, and the division of power moves to ϕ where it remains. Once at ϕ, the elite no longer invests in undermining a proposed change. The second term in (13) is the expected cost of paying e^* until the division of power moves to ϕ.

Proposition 1 brings all of the results together. To state them most simply, use $V_E(\phi, s) = \min\{\delta, 1-\phi\}/(1-\beta)$, (11), and (13) to write the first-order condition as

$$\begin{aligned} 1 = &\beta w\rho'(e^*(f,\phi)) \qquad\qquad (14)\\ &\times \left[\frac{(1-\beta)(1-r)(\phi-f)+r\max\{0,\phi+\delta-1\}}{(1-\beta)[1-\beta(1-r)]} - \frac{re^*(f,\phi)}{1-\beta+\beta r[1-w\rho(e^*(f,\phi))]}\right]. \end{aligned}$$

This leaves

Proposition 1. *A power-sharing MPE exists if and only if for every* $f \in [0, \tilde{f})$, *there exists an offer-effort pair* $\phi^*(f)$ *and* $e^*(f, \phi^*(f))$ *that satisfies (12) and (14) as well as the feasibility constraints* $\phi^*(f) \in [f, 1]$, $w\rho(e^*(f, \phi^*(f))) \in [0, 1]$, *and* $V_E(f, s) \geq -d$.

Condition (14) guarantees that the elite is exerting the optimal level of effort given the current division of power f and the proposed division ϕ^*. Condition (12) ensures that the proposal of $(1, \phi^*)$ satisfies the binding no-revolution constraint

Weak Institutions

This section presents the main results. The first centers on how the feasibility of power sharing varies with institutional strength and commitment power. When institutions are strong enough, power-sharing equilibria exist. When institutions are too weak, the commitment problem is too severe and there are no power-sharing equilibria.

PROPOSITION 2. *There exist thresholds* $0 < w' \leq w''$ *such that a unique power-sharing Markov perfect equilibrium path exists if institutions are sufficiently strong, i.e., if* $0 \leq w < w'$. *No power-sharing equilibria exist if institutions are too weak, i.e., if* $w > w''$.

The details of the proof are tedious, but the intuition is straightforward. When institutions are strong (w is small), the marginal return on trying to undermine an agreement is low. As a result, the elite invests less and the probability that an agreement will actually hold is high. Given that an agreement is likely to hold, the elite can commit to enough future redistribution (in expectation) to buy the opposition off.

When institutions are weaker (w is larger), the marginal return to trying to undermine an agreement is higher. This induces the elite to invest more in undercutting an agreement and the probability that an agreement will hold decreases (see Proposition 3(i)). As the opposition's expected return to any agreement goes down, the elite must offer ever better terms to the opposition to offset the higher chances that the agreement will not hold ($\partial\phi^*(f)/\partial w > 0$). As institutions continue to weaken, the elite is eventually unable to offer enough to buy the opposition off. Even if the elite promises full democratization ($\phi = 1$), the chances that this will actually happen are too low and the promise to democratize too hollow to buy the opposition off.

Proposition 3 summarizes the comparative statics.

PROPOSITION 3 (Comparative Statics). *Let* $\phi^*(f)$ *be the equilibrium power-sharing proposal at* f *and* $e^*(f, \phi^*(f))$ *be the equilibrium effort the elite exerts in undermining the agreement. Then:*

(i) The elite invests more in undermining agreements and power-sharing agreements are less likely to hold when institutions are weak: $\partial e^*/\partial w > 0$ *and* $\partial w\rho(e^*)/\partial w > 0$.

(ii) The more power the opposition already has, the less the elite offers when facing an immediate threat: $\partial \phi^*(f)/\partial f < 0$.

(iii) The less frequent the threat of revolution, the more the elite has to offer to buy the opposition off: $\partial \phi^*/\partial r < 0$. *Assume* $w\rho(\overline{e}) = 1$ *for some effort* $\overline{e}$.[19] *Then there exists a* $\underline{\beta} < 1$ *such that for any* $\beta \in (\underline{\beta}, 1)$, *no power-sharing equilibrium exists if* r *is sufficiently small.*

(iv) The weaker institutions are, the more power the elite ultimately shares: $\partial \phi^*/\partial w > 0$.

We discuss each of the comparative-static results in turn. When institutions are weak, the marginal return to trying to undermine an agreement is higher and this tends to induce more effort. There is however a second, countervailing equilibrium effect. Higher effort and a higher probability of undoing an agreement reduce the elite's continuation value $V_E(f, s)$ (see (13)). This reduces the payoff to undermining the agreement and the incentive to invest. Nevertheless, the net effect of weaker institutions is that they induce more effort and make agreements less likely to hold.

As for the effect of changes in f on ϕ^*, the less favorable the division of power is for the elite when it faces an immediate threat, i.e., the larger f, the lower the elite's stakes in undermining the agreement and preserving the existing distribution of power. Lower stakes in turn reduce the elite's incentive to invest in undoing the agreement. As a result, agreements are more likely to hold, and the less the elite has to offer to the opposition. This result plays an important role in the discussion below about path dependence.

Now consider the effects of the frequency of the threat of revolution. The expected duration to the next threat is $1/r$. The smaller r, the longer the expected interval until the elite next has to deal with a direct threat. This longer interval increases the stakes in undermining a proposed division of power ϕ^*. If the elite succeeds in undoing an

[19]This assumption ensures that the elite can generate any probability of undoing an agreement if it is willing to exert enough effort. This assumption simplifies the analysis by ruling out the possibility that $\lim_{e\to\infty} w\rho(e) < 1$, i.e., that the probability of undoing an agreement is bounded above by something less than one.

agreement on ϕ^*, it expects to get $1-f$ instead of $1-\phi^*$ for $1/r$ periods. Because the stakes are higher, the elite invests more. This lowers the opposition's payoff to accepting and forces the elite to offer better terms.

These dynamics lead to a non-monotonic relationship between power sharing and how often the opposition poses a threat when the elite and opposition care enough about the future. If the threat of revolution is constant or sufficiently frequent, i.e., if $r \geq (\beta-\delta)/\beta$, there is no commitment problem. The elite in theses circumstances does not share any power and buys the opposition off with policy concessions alone.[20] As r decreases and threats become less frequent, the elite starts to face a commitment problem that it can only solve by sharing at least some power with the opposition. This is the standard result in window-of-opportunity models.

But as r continues to decrease in the present model, the stakes continue to grow, and the elite invests more in undermining the agreement. This reduces the opposition's payoff to accepting and forces the elite to offer better and better terms. If the elite and opposition care enough about the future, then the stakes will eventually be so high that even the offer of complete democratization ($\phi = 1$) will not be enough to stop a revolution.

In sum, there is a non-monotonic relation between power-sharing and the frequency of the opposition's threat. When r is high enough, there is no power sharing. The elite buys the opposition off with policy concessions when needed. For intermediate values of r, the elite and opposition agree to share power with better terms associated with less frequent threats. When r is too small, the stakes are too high. The elite will work too hard to undermine an agreement, and the opposition fights. There are no power-sharing agreements.

Finally, weaker institutions (larger w) increase the incentive to invest in undermining an agreement. This lowers the opposition's payoff to accepting. The elite, therefore, has to offer more in order to satisfy the no-revolution constraint.

This result resonates with Albertus and Menaldo's (2018) findings about gamed constitutions. They emphasize the distinction between transitions to elite-biased democracies and to popular democracies.[21] In the former, the incumbent elite is able to negotiate an

[20]If $r \geq (\beta-\delta)/\beta$, then $\tilde{f} \leq 0$ and Lemma 2 describes the equilibrium for all $f \in [0,1]$.
[21]Of 214 transitions from authoritarian rule between 1946 and 2010, only 48 percent

agreement that protects many of its interests. These agreements often take the form of "gamed" constitutions drafted under authoritarian rule which give disproportionate power and influence to the elite through, for example, "favorable electoral rules or malapportionment *along with* obstacles to constitutional change, such as requiring large supermajorities to scrap the constitution" (2018, 63). Gamed constitutions often also contain provisions prohibiting retroactive criminal prosecution and banning left-wing parties (2018, 63, 81-97). By contrast, a transition to popular democracy occurs when the elite is unable to cut a deal protecting its interests and elected officials and institutions more faithfully represent the median voter (2018, 27, 42-43). Of 122 democratic transitions between 1800 and 2006, eighty were elite-biased (2018, 110).[22]

Albertus and Menaldo argue that elites will be better able to coordinate on an agreement and commit to it "when they posses tailor-made administrative infrastructures and political forums." As a result, "there should be a strong, positive relationship between state capacity and a transition to elite-biased democracy" as well as a "strong, positive relationship between the presence of a legislature under dictatorship and a transition to an elite-biased democracy" (2018, 48-49). These hypotheses find significant support in the data (2018, 112-16).

The present analysis emphasizes the limited ability of the elite to commit to an agreement whereas Albertus and Menaldo focus on the problem of multiple bargainers coordinating on an agreement and on the limited ability of the opposition to commit to an agreement.[23] Nevertheless, the present analysis yields comparative-static predictions consistent with Albertus and Menaldo's statistical findings. Suppose we think of greater state capacity or the presence of a legislature under dictatorship as indicative of stronger institutions (smaller w) and gamed constitutions as agreements that are better for the elite (smaller ϕ^*). Then the model predicts a positive association between stronger institutions and better terms for the elite. Formally, $\partial\phi^*/\partial w > 0$ by Proposition 3(iv).

(102) transitioned to democracy whereas 52 percent (112) transitioned to a different autocratic regime (Geddes, Wright, Frantz 2014).

[22]Albertus and Menaldo code a transition as elite biased if it occurs under a constitution drafted and adopted while the state was still under authoritarian rule (2018, 72-73).

[23]See Fearon and François (2020) for an effort to formalize this commitment problem.

In the baseline model developed above, the opposition either poses a direct threat and is sure to depose the elite if it rebels, or there is no immediate threat and no chance of a successful rebellion. The corresponding equilibrium path is very simple. When the elite begins in complete control ($f = 0$), it proposes $\phi^*(0) \geq \tilde{f}$ whenever the opposition is strong. Even if the elite initially succeeds in undermining this agreement, it eventually holds. The distribution of power moves to $\phi^*(0)$ where it remains.

When the severity of the threat can vary, power-sharing is path dependent. This section adds a moderate threat to the baseline game and describes a path dependent equilibrium. Online Appendix B adds technical detail and constructs a numerical example demonstrating existence.

If the elite happens to face a series of threats that are increasing in intensity, then the elite gradually shares more and more power with the opposition. If, by contrast, the first threat the elite faces is severe, the elite has to make a very large initial proposal in order to satisfy the no-revolution constraint. These two paths generally lead to different endpoints with the elite retaining more power if the transition is gradual. Indeed, if institutions are too weak, the elite may be deposed because it is unable to buy the opposition off if the first threat is severe whereas a series of gradually increasing threats would have led to a peaceful division of power.[24]

The key intuition underlying this path dependence is that the stakes inherent in the gamble of an agreement depend on the pattern of threats. The more power the opposition already has when the elite faces a severe threat, the smaller the elite's stakes in buying the opposition off and the less the elite invests in undermining an agreement. As a result, it is easier to satisfy the no-revolution constraint and the equilibrium agreement shares less power with opposition. That is, $\partial\phi^*(f)/\partial f < 0$ as Proposition 3(ii) showed for the baseline game.

In addition to path dependence, there is a second difference between the equilibrium when the opposition can pose a moderate or severe threat and the equilibrium when the

[24]Acemoglu, Egorov, and Sonin (2015) also find that the steady state distribution of power depends on the pattern of shocks.

opposition can only pose a severe threat. The elite in both shares as little power as possible when facing a severe threat. But there is a countervailing consideration when the elite faces a moderate threat. The more power the elite shares when facing a moderate threat, the lower the stakes when the elite ultimately does face a severe threat. This consideration induces the elite to share some power even when it could buy the opposition off with policy concessions alone.

To describe the extended game more formally, suppose that there are now three levels of threat $\nu_t \in \{n, m, s\}$ where m denotes a moderate threat which occurs with probability μ. The state is active when the treat is moderate or severe: the elite can propose an agreement, the opposition can accept or fight, and the elite can try to undermine an agreement. As in the baseline game, fighting is a game-ending move. The probability that the opposition prevails when posing a moderate threat is $\pi < 1$, and the opposition's payoff to fighting is $\pi(1-\delta)/(1-\beta)$. The elite gets $-\pi d + (1-\pi)(1-\delta)/(1-\beta)$ where the cost of being deposed, d, is high enough that the elite prefers to buy the opposition off whenever possible. A Markov strategy for the elite specifies its offer when facing a moderate threat, $(y_m^*(f), \phi_m^*(f))$; its offer when facing a severe threat, $(y^*(f), \phi^*(f))$; and the effort it exerts to undo an agreement on $\phi \in [f, 1]$, $e^*(f, \phi)$.[25]

There are two no-revolution constraints in the extended game. When the elite is facing a severe threat, the no-revolution constraint (s-NRC) is still defined by (3). The moderate-threat, no-revolution constraint (m-NRC) is

$$\frac{\pi(1-\delta)}{1-\beta} \leq y_m + \beta w \rho(e^*(f, \phi_m)) V_O(f) + \beta[1 - w\rho(e^*(f, \phi_m))] V_O(\phi_m).$$

In equilibrium, the elite holds the opposition down to its reservation value and the relevant no-revolution constraint binds whenever the opposition poses a credible threat.

Three thresholds help characterize the equilibrium path. Recall that the opposition at least weakly prefers the no-concession offer $(y, \phi) = (f, f)$ to rebeling in the baseline game when $f \geq 1 - \delta$. As a result, a threat to fight is incredible, and the elite makes no

[25] As in the baseline game, effort depends on the stakes defined by f and ϕ and not on the policy proposal. It is also clear that effort does not depend on the severity of the threat except insofar as the severity affects the stakes through ϕ.

additional concessions. When $f \in [\tilde{f}, 1-\delta)$, the threat is credible, but f is sufficiently large that the elite can buy the opposition off with policy concessions alone. The threshold $\tilde{f}$ is the cutpoint at which the opposition's payoff to accepting $(1, f)$ equals its payoff to fighting, i.e., $1 + \beta V_O(\tilde{f}) = (1-\delta)/(1-\beta)$. These thresholds hold in the extended game where it is convenient to denote the latter by $\tilde{f}_s$.[26] As for the third, the opposition weakly prefers the status quo f to fighting when it poses a moderate threat as long as $f + \beta V_O(f) \geq \pi(1-\delta)/(1-\beta)$ or, equivalently, $f \geq f_c$ where f_c is satisfies $f_c + \beta V_O(f_c) = \pi(1-\delta)/(1-\beta)$. If π is not too large: $1 - \delta > \tilde{f}_s \geq f_c$ (see Assumption 2 in Appendix B.)

To describe the equilibrium path, suppose that elite is in complete control ($f = 0$) when it faces a moderate threat. (It may be that the elite has already faced threats in the passed but had that no previous agreement has held.) Faced with a moderate threat, the elite proposes $\phi^*_m(0)$. If this agreement holds, the division of power moves to $\phi^*_m(0)$. If $\phi^*_m(0) \geq f_c$, the opposition cannot credibly threaten to rebel when it subsequently poses a moderate threat, and the division of power remains at $\phi^*_m(0)$ until the opposition poses a severe threat. If $\phi^*_m(0) < f_c$, the elite must buy the opposition off when it next faces a moderate threat, and the division of power moves to $\phi^*_m(\phi^*_m(0)) \geq \phi^*_m(0)$ if this agreement holds. Let $q \geq \phi^*_m(0)$ be the division of power when the elite faces a severe threat and cuts a deal that holds. Then the division of power moves to $\phi^*(q)$ where it remains because $\phi^*(q) \geq \tilde{f}_s$. If, by contrast, the elite is still in complete control when it faces a severe threat, it offers $\phi^*(0) > \tilde{f}_s$ where the division of power remains once it moves there.

Appendix B constructs a peaceful MPE at which $\phi^*(0) = 1$ for a specific institutional strength $\tilde{w}$.[27] The offer ϕ^* is decreasing, so $\phi^*(0) > \phi^*(\phi_m(0)) \geq \phi^*(q)$ and there is path dependence. All transitions end peacefully but at different divisions of power depending on the pattern of threats. If, however, institutions are slightly weaker, i.e., $w = \tilde{w} - \epsilon$ for a small ϵ, then $\phi^*(0) > 1 > \phi^*(\phi_m(0))$. Now a severe threat to an elite in complete control

[26]The actors' continuation values starting from f in the extended game may differ from the continuation values starting from f in the baseline game because the continuation games differ with the extended game allowing for a moderate threat. As a result, $\tilde{f}_s$ may not be given by (4). See the online appendix for a discussion of the relation between $\tilde{f}$ and $\tilde{f}_s$.

[27]Rather than taking w as a parameter and solving for the equilibrium, one can solve for w subject to the constraint that $\phi^*(0) = 1$.

ends in revolution whereas a gradual transition would have led to a peaceful transition ending at $\phi^*(q)$.

We now show that the elite shares as little power as possible when facing a severe threat as in the baseline game, but it shares at least some power when facing a credible moderate threat. Paralleling the derivation of (7), the elite's payoff along the binding m-NRC is

$$b_m(\phi_m) = 1 - \frac{\pi(1-\delta)}{1-\beta} - e^*(f,\phi_m) + \beta w\rho(e^*(f,\phi_m))[V_E(f) + V_O(f)]$$
$$+ \beta[1 - w\rho(e^*(f,\phi_m))][V_E(\phi_m) + V_O(\phi_m)]$$

where we have used the binding m-NRC to eliminate y_m. Differentiation and the first order condition $1 = w\rho'(e^*)[V_E(f) - V_E(\phi_m)]$ yield

$$\frac{\partial b_m}{\partial \phi_m} = -\beta w\rho'(e^*)\left[V_O(\phi_m) - V_O(f)\right]\frac{\partial e^*}{\partial \phi_m} + \beta\left[1 - w\rho(e^*)\right]\frac{\partial(V_E(\phi_m) + V_O(\phi_m))}{\partial \phi_m}. \quad (15)$$

The key difference between (7) and (15) is that ϕ must be at least as large as $\tilde{f}_s$ in order to satisfy the s-NRC, but ϕ_m need not be this large. The former implies $V_E(\phi) + V_O(\phi) = 1/(1-\beta)$ and, consequently, $\partial(V_E(\phi) + V_O(\phi))/\partial\phi = 0$. It follows that the sign of $\partial b_s/\partial\phi$ depends solely on the sign of the first term of (7) which is negative. In words, the elite's payoff is decreasing in the amount of power it shares, so it shares as little as possible when facing a severe threat. By contrast, the sign of $\partial b_m/\partial\phi_m$ is ambiguous as $V_E(\phi_m) + V_O(\phi_m)$ is increasing when $\phi_m < \tilde{f}_s$.[28] The second term in (15) formalizes the "hedge" of sharing power when facing a moderate threat. The larger ϕ_m, the smaller the stakes and thus the smaller the deadweight loss when the elite ultimately does face a severe threat.

To see that the elite shares some power when facing a credible moderate threat, i.e., when $f < f_c$, observe that the elite cannot buy the opposition off solely with policy concessions alone if $1 + \beta V_O(f) < \pi(1-\delta)/(1-\beta)$ or, equivalently, if $f < \tilde{f}_m$ where

[28]The sum of the players' continuation values along a peaceful path is the total flow of benefits, $1/(1-\beta)$, less the deadweight loss due to the elite's efforts to undermine agreements. The higher ϕ_m, the smaller the stakes when the elite subsequently faces a credible threat and the smaller the loss due to trying to undermine the agreement. As a result, $V_E(\phi_m) + V_O(\phi_m)$ is increasing.

$1 + \beta V_O(\tilde{f}_m) = \pi(1-\delta)/(1-\beta)$. It follows that the elite must share some power at $f < \tilde{f}_m$. For $f \in [\tilde{f}_m, f_c)$, it suffices to show that the marginal gain to offering at least a little power is positive, i.e, $\lim_{\phi_m \to f} \partial b_m/\partial \phi_m > 0$. The first term in (15) clearly goes to zero as ϕ goes to f as long as the marginal effect on the probability of war of sharing more power, $w\rho'(e^*)\partial e^*/\partial \phi_m$, is bounded at zero. The assumption that $[\rho'(e)]^3/\rho''(e)$ is bounded at zero ensures that this is the case.[29] By contrast, the second term in (15) is positive because $V_E(\phi_m) + V_O(\phi_m)$ is increasing as noted above. Thus, $\lim_{\phi_m \to f} \partial b_m(\phi_m)/\partial \phi_m > 0$ and the elite is sure to share some power when facing a moderate threat.

Conclusion

Democratic transitions, franchise extensions, and civil-war settlements can often be seen as power-sharing agreements in which opposing factions try to use institutional structures to "lock in" the terms of a settlement. But the commitment power inherent in institutions varies. When institutions are weak, it is difficult for a powerful elite to tie its hands and give up power. If institutions are too weak, the elite is unable to buy the opposition off. Even the expected payoff of a power-sharing agreement giving the opposition complete control is too little to induce the opposition to forego rebellion. When institutions are strong, the elite can more readily give up power if it chooses to and can use institutions to commit to future redistribution.

Power-sharing agreements are effectively gambles when institutions lack complete commitment power. Weaker institutions make the gambles riskier. This induces an inverse relation between the strength of the institutions and the terms of an equilibrium agreement. Stronger institutions mean that less power is shared and there is more "gaming" favoring the elite.

[29] Differentiating the first-order condition gives $\rho''(e^*)[V_E(f) - V_E(\phi_m)]\partial e^*/\partial \phi_m = \rho'(e^*)\partial V_E(\phi_m)/\partial \phi_m$. Using the first-order condition again to eliminate $V_E(f) - V_E(\phi_m)$ from the previous equation yields

$$w\rho'(e^*)\frac{\partial e^*}{\partial \phi_m} = \frac{\beta w^2[\rho'(e^*)]^3 \partial V_E(\phi_m)/\partial \phi_m}{\rho''(e^*)} \tag{16}$$

which is bounded at zero if $[\rho'(e)]^3/\rho''(e)$ is.

Finally, power sharing is path dependent. How much power the elite ultimately shares and whether the transition ends in an agreement or rebellion generally depends on the pattern of threats the elite faces over time. If the elite is still in complete control when it faces a severe threat, it will have to share a lot of power in order to buy the opposition off. Indeed, it may not be able to buy the opposition off if institutions are too weak. If the elite has already faced a moderate threat and has shared some power, the elite's stakes will be smaller when it confronts a severe threat. As a result, the elite will ultimately have to share less power than it would have had to share had it still been in complete control when it faced a severe threat.

References

Acemoglu, Daron, and James A. Robinson. 2000. "Why Did the West Extend the Franchise," *Quarterly Journal of Economics*, 115(November): 1167-99.

Acemoglu, Daron, and James Robinson. 2001. "A Theory of Political Transitions," *American Economic Review*, 91(September): 938-63.

Acemoglu, Daron, and James Robinson. 2006. *The Economic Origins of Democracy.* New York: Cambridge University Press.

Acemoglu, Daron, and James Robinson. 2008. "Persistence of Power, Elites, and Institutions," *American Economic Review*, 918(1): 267-93.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2012. "Dynamic Stability of Constitutions, Coalitions, and Clubs," *American Economic Review*, 102(4): 1446-76.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2015. "Political Economy in a Changing World," *Journal of Political Economy*, 123(5): 1038-86.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2020. ???

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James Robinson. 2013. "Democracy, Public Policy and Inequality," *Comparative Democratization*, 11(3):2,16-20.

Aidt, Toke S. and Peter S. Jensen. 2014. "Workers of the World, Unite!" *European Economic Review*,72(C): 52-75.

Aidt, Toke S. and Gabriel Leon. 2015. "The Democratic Window of Opportunity," *Journal of Conflict Resolution*, 60(4): 694-717.

Aidt, Toke S. and Raphael Frank. 2015. "Democratization under the Threat of Revolution, *Econometrica*, 83(2): 505-47.

Albertus, Michael and Victor Menaldo. 2018. *Authoritarianism and the Elite Origins of Democracy.* New York: Cambridge University Press.

BBC. 2016. "Sudan Crisis: What you need to know," August 16. Available at https://www.bbc.com/news/world-africa-48511226.

Burke, Jason and Zeinab Mohammed Salih. 2019. "Sudanese Military and Protesters Sign Power-Sharing accord," *The Guardian*, July 17. Available at www.theguardian.com/world/2019/jul/17/sudan-military-and-protesters-sign-power-sharing-accord. Accessed on July 31, 2020.

Burke, Paul J. and Andrew Leigh. 2010. "Do Output Contractions Trigger Democratic Change?" *American Economic Journal: Macroeconomics* 2(4): 124-57.

Bruckner, Markus and Antionio Ciccone. 2011. "Rain and the Democratic Window of Opportunity," *Econometrica*, 7(3): 923-47.

CRS. 2019. "Sudan's Uncertain Transition." Washington, D.C.: Congressional Research Service.

de Figueiredo, Rui. 2002. "Electoral Competition, Political Uncertainty, and Policy Insulation," *American Political Science Review*, 96(2): 32135.

Dower, Paul C., Evgeny Finkel, Scott Gehlbach, and Steven Nafzinger. 2018. "Collective Action and Representation in Autocracies," *American Political Science Review*, 112(1): 125-47.

Dower, Paul C., Evgeny Finkel, Scott Gehlbach, and Steven Nafzinger. 2020. "Democratization as a Continuous Choice," *Journal of Politics*, 82(2): 776-80.

Economist. 2019a. "A Power-sharing deal in Sudan paves the way of civilian rule," August 22. Available at: https://www.economist.com/middle-east-and-africa/2019/08/22/a-power-sharing-deal-in-sudan-paves-the-way-for-civilian-rule.

Economist. 2019b. "The Sudan peace deal is greeted with little excitement," July 8.

Fearon, James D. 2004. "Why Do Some Civil Wars Last So Long?" *Journal of Peace Research*, 41(3):275-302.

Fearon, James D. and David Laitin. 2008. "Civil War Termination," Manuscript, Department of Political Science, Stanford University. 2008. Available at: http://www.stanford.edu/jfearon/papers/termination.pdf.

Fearon, James D. and Patrick François. 2020. "A Theory of Elite-Initiated Democratization, Illustrated With the Case of Myanmar." Manuscript, Department of Political Science, Stanford University.

Gallopin, Jean-Baptiste. 2019. "Sudan's Chance for Democracy," *European Council on Foreign Affairs*. Available at https://www.ecfr.eu/article/commentary_sudans_chance_for_democracy. Accessed July 31, 2020.

Geddes, Barbara., Wright, Joseph, and Erica Frantz. 2014. "Autocratic Breakdown and Regime Transitions: A New Data Set," *Perspectives on Politics*, 12(2): 313-31. Available at doi.org/10.1017/S1537592714000851.

Gieczewski, Germán. 2021. "Policy Persistence and Drift in Organizations," *Econometrica*, 89(1):251-79.

Haggard, Stephan and Robert Kaufman. 2012. "Inequality and Regime Change," *American Political Science Review*, 106(3): 495-516.

Haggard, Stephan and Robert Kaufman. 2016. *Dictators and Democrats*. Princeton, NJ: Princeton University Press.

Hartzell, Caroline and Matthew Hoddie. 2003. "Institutionalizing Peace," *American Journal of Political Science*, 47(2):318-32.

Hartzell, Caroline and Matthew Hoddie. 2007. *Crafting Peace: Power Sharing Institutions and the Negotiated Resolution of Civil Wars*. University Park, PA: Penn State University Press.

Hartzell, Caroline and Matthew Hoddie. 2020. *Power Sharing and Democracy in Post-Civil War States*. New York: Cambridge University Press.

Hassan, Mai and Ahmed Kodouda. 2019. "Sudan's Uprising: The Fall of a Dictator," *Journal of Democracy*, 30(4): 89-103.

ICG. 2019. *Safe Guarding Sudan's Revolution*. Brussels: International Crisis Group.

Jack, William and Roger Lagunoff. 2006. "Dynamic Enfranchisement," *Journal of Public Economics*, 90: 551-72.

Lizzeri, Alessandro and Nicola Persico. 2004. "Why Did the Elites Extend the Suffrage?" *Quarterly Journal of Economics*, 119(2):707-765.

Lynch, Justin. 2019. "How Sudan's Military Overcame the Revolution," *Foreign Policy Dispatch*. Available at foreignpolicy.com/2019/08/05/how-sudan-military-overcame-the-revolution-constitution-protests/. Accessed on July 30, 2020.

Page, Scott. 2006. "Path Dependence," *Quarterly Journal of Political Science*, 1: 87115

Powell, Robert. 2004. "The Inefficient Use of Power: Costly Conflict with Complete Information," *American Political Science Review*, 98(2): 231-41.

Przworski, Adam. 2008. "Conquered or Granted? A History of Suffrage Extensions," *British Journal of Political Science*, 39: 291321.

Riedl, Rachel, Dan Slater, Joseph Wong, and Daniel Ziblatt. 2020. "Authoritarian-Led Democratization," *Annual Review of Political Science*, 23(1):315-32. Available at https://doi.org/10.1146/annurev-polisci-052318-025732.

Triesman, Daniel. 2020. "Democracy by Mistake," *American Political Science Review*, 114(3):792-810. Available at: https://doi.org/10.1017/S0003055420000180.

Walsh, Declan. 2019a. "Bouncing Back from Crackdown, Protestors Surge Through Sudan's Street," *New York Times*, June 30. Available at https://www.nytimes.com/2019/06/30/world/africa/sudan-protest-hamdan.html. Accessed July 30, 2020.

Walsh, Declan. 2019b. "Sudan Erupts in Celebration After Army and Civilians Agree to Share Power," *New York Times*, August 17. Available at https://www.nytimes.com/2019/08/17/world/africa/sudan-agreement-signed.html. Accessed March 24, 2020.

Walsh, Declan. 2019c. "Sudan's Protestors Reject Military Plan After Crackdown Kills Dozens," *New York Times*, June 4. Available at https://www.nytimes.com/2019/06/04/world/africa/sudan-protest-military.html. Accessed July 30, 2020.

Walter, Barbara. 2002. *Committing to Peace: The Successful Settlement of Civil Wars*. Princeton, NJ: Princeton University Press.

Walter, Barbara. 2009. "Bargaining Failures and Civil War," *Annual Review of Political Science*, 12:24361.

Wantchekon, Leonard. 2000. "Credible Power-Sharing Agreements," *Constitutional Political Economy* 11(4):339352.

Ziblatt, Daniel. 2017. *Conservative Parties and the Birth of Democracy*. New York: Cambridge University Press.

Appendix A (online)

The appendix begins by establishing a series of results needed to identify necessary conditions that a peaceful power-sharing MPE must satisfy. The section then proves Propositions 1-3. To define some useful notation, let σ be a Markov profile and take $\alpha(y, \phi|f, \sigma)$ be the opposition's payoff to accepting (y, ϕ) given that play subsequently follows σ. That is, $\alpha(y, \phi|f, \sigma) \equiv y + \beta w \rho(e(f, \phi))V_O(f) + \beta[1 - w\rho(e(f, \phi))]V_O(\phi)$ where we will often write $\alpha(y, \phi)$ to ease the notation. Take $\mathcal{A}(f|\sigma) \equiv \{(y, \phi) : \alpha(y, \phi) \geq (1 - \delta)/(1 - \beta)\}$ to be the set of offers that satisfy the no-revolution constraint at (f, s) where, again, we generally write $\mathcal{A}(f)$.

The first result restates Lemma 1 taking the boundary conditions and a possible corner solution into account. Define $\underline{\rho}' \equiv \rho'(\rho^{-1}(1/w))$ if $\rho^{-1}(1/w)$ is well defined and $\underline{\rho}' \equiv \lim_{e \to \infty} \rho'(e)$ otherwise.

Result 1 (Optimal Effort with boundary conditions). *Let σ^* be an MPE and $\Delta_E \equiv V_E(f|\sigma^*) - V_E(\phi|\sigma^*)$. Then $e^*(f, \phi|\sigma^*) = 0$ if $\Delta_E \leq 0$. If $\Delta_E > 0$, then $e^*(f, \phi|\sigma^*)$ uniquely satisfies $1 = \beta w \rho'(e^*(f, \phi|\sigma^*))\Delta_E$ if $1 > \beta w \underline{\rho}' \Delta_E$. If $\Delta_E > 0$ and $\beta w \underline{\rho}' \Delta_E \geq 1$, then $e^*(f, \phi|\sigma^*) = \rho^{-1}(1/w)$ if this is well defined and $e^*(f, \phi) = \infty$ otherwise.*

Proof of Lemma 2: Let σ^* be any MPE and assume $f \in [\tilde{f}, 1 - \delta]$.[30] The first step is to show that

$$V_O(f) \geq \frac{(1 - \beta)(1 - r)f + r(1 - \delta)}{(1 - \beta)[1 - \beta(1 - r)]}.$$

Consider any possible path along σ^* starting from (f, n). This will be a sequence of weak states that ultimately leads to a state in which the opposition is strong. The opposition's payoff in the weak state is f, and the opposition's continuation payoff at the last state is bounded below by its payoff to fighting which it will get by fighting or because the elite satisfies the opposition's no-revolution constraint. It follows that $V_O(f, n)$ is bounded below by $H_O(f, n)$ where $H_O(f, n)$ satisfies the recursive relation

$$H_O(f, n) = f + \beta \left[(1 - r)H_O(f, n) + \frac{r(1 - \delta)}{1 - \beta}\right]$$

[30] If $f > 1 - \delta$, then the opposition's payoff to getting f in every period exceeds its payoff to fighting. Lacking a credible threat to fight at (f, s), it is straightforward to show that the elite always offers (f, f) and exerts $e^*(f, f) = 0$.

$$H_O(f,n) = \frac{(1-\beta)f + \beta r(1-\delta)}{(1-\beta)[1-\beta(1-r)]}. \tag{17}$$

This implies $V_O(f) \geq (1-r)H_O(f,n) + r(1-\delta)/(1-\beta)$ or

$$V_O(f) \geq \frac{(1-\beta)(1-r)f + r(1-\delta)}{(1-\beta)[1-\beta(1-r)]}.$$

This lower bound means that the opposition weakly prefers accepting $(x(f), f)$ to fighting when $f \geq \tilde{f}$. To verify this, note that the elite invests $e^*(f,f) = 0$, so the division of power f is sure to continue into the next round. The opposition's payoff to accepting is $x(f) + \beta V_O(f)$ which is bounded below by $(1-\delta)/(1-\beta)$

We now establish that $V_E(f,s)$ is bounded below by $\delta/(1-\beta)$. To avoid tedious limit arguments, assume the opposition accepts if indifferent.[31] Then, the elite's payoff to the strategy of offering $(x(f), f)$ is

$$1 - x(f) + \beta\left[\frac{(1-r)(1-f) + r(1-x(f))}{1-\beta}\right] = \frac{\delta}{1-\beta}.$$

This strategy cannot be a profitable deviation, so $V_E(f,s) \geq \delta/(1-\beta)$.

Continuation values now follow. In addition to the lower bound on $V_E(f,s)$, we also have $V_O(f,s) \geq (1-\delta)/(1-\beta)$ and $V_E(f,s) + V_O(f,s) \leq 1/(1-\beta)$. Hence, $V_E(f,s) = \delta/(1-\beta)$ and $V_O(f,s) = (1-\delta)/(1-\beta)$ in any MPE. Using $V_E(f,n) = 1 - f + \beta(1-r)V_E(f,n) + \beta r V_E(f,s)$ and $V_O(f,n) = 1 - f + \beta(1-r)V_O(f,n) + \beta r V_O(f,s)$ gives

$$V_E(f) = \frac{(1-\beta)(1-r)(1-f) + r\delta}{(1-\beta)[1-\beta(1-r)]}. \tag{18}$$

$$V_O(f) = \frac{(1-\beta)(1-r)f + r(1-\delta)}{(1-\beta)[1-\beta(1-r)]}. \tag{19}$$

[31] If the elite offers slightly more than $x(f)$ or slightly more than f if $x(f) = 1$, the opposition strictly prefers accepting. This leaves the elite with an arbitrarily small amount less than $\delta/(1-\beta)$ and, consequently, the argument showing that $V_E(f,s)$ is bounded below by $\delta/(1-\beta)$ still goes through.

Clearly, $V_O(f)$ is increasing in f and $V_E(f)$ is decreasing. The latter implies $V_E(f) - V_E(\phi) > 0$ when $\phi > f$. Result 1 and the assumption that $\lim_{e\to 0} \rho'(e) = \infty$ imply $e^*(f, \phi) > 0$ if $\phi > f$.

Arguing by contradiction to show that the elite actually does propose $(x(f), f)$ when the opposition is strong, suppose the elite's equilibrium strategy were to offer some (z, ϕ) with $\phi > f$. We demonstrate that the payoff to this strategy is strictly less than $\delta/(1-\beta)$ which was shown above to be the elite's continuation payoff at (f, s).

The elite clearly does worse by offering (z, ϕ) if this is too little to dissuade the opposition from rebelling. Assume then that (z, ϕ) satisfies the no-revolution constraint. This leaves $z + \beta w\rho(e^*(f, \phi))V_O(f) + \beta[1 - w\rho(e^*(f, \phi))]V_O(\phi) \geq (1 - \delta)/(1 - \beta)$. The elite gets $v_E(z, \phi) = 1 - z - e^*(f, \phi) + \beta w\rho(e^*(f, \phi))V_E(f) + \beta[1 - w\rho(e^*(f, \phi))]V_E(\phi)$. It follows that satisfying the no-revolution constraint puts an upper bound on the elite's payoff: $v_E(z, \phi) \leq [(\beta - \delta)/(1 - \beta)] - e^*(f, \phi) + \beta w\rho(e^*(f, \phi))[V_E(f) + V_O(f)] + \beta[1 - w\rho(e^*(f, \phi))][V_E(\phi) + V_O(\phi)]$. The expressions determined above for $V_E(f, s)$ and $V_O(f, s)$ imply $V_E(f) + V_O(f) = V_E(\phi) + V_O(\phi) = 1/(1 - \beta)$. Substitution then gives $v_E(z, \phi) \leq \delta/(1 - \beta) - e^*(f, \phi)$. But $e^*(f, \phi) > 0$ when $\phi > f$, so the elite's payoff is less than $\delta/(1 - \beta)$.

This contradiction means that the elite strictly prefers not to share any power when $f \geq \tilde{f}$. Given that $\phi = f$, the opposition's payoff to accepting (y, ϕ) is $y + \beta V_O(f)$ which in turn must satisfy the no-revolution constraint. That is, $y + \beta V_O(f) \geq (1 - \delta)/(1 - \beta)$. The elite will offer the smallest $y \geq f$ satisfying this constraint which, as shown above, is $\max\{x(f), f\}$. Hence, the elite offers $(\max\{x(f), f\}, f)$ when faced with unrest. □

The second and third results verify the intuition that both the elite and opposition prefer more power to less.

Result 2. *Let σ^* be an MPE with $f < \tilde{f}$. If the no-revolution constraint is slack at (y, ϕ) or if it binds with $y < 1$, then $V_E(f|\sigma^*) > V_E(\phi|\sigma^*)$.*

Proof: Arguing by contradiction, assume that $V_E(\phi|\sigma^*) \geq V_E(f|\sigma^*)$ and note that $e^*(f, \phi) = 0$ by Result 1.

The payoff to accepting is $\alpha(y, \phi) = y + \beta V_O(\phi)$ because $e^*(f, \phi) = 0$. If the no-revolution constraint is slack or if $y < 1$, then $(1 - \delta)/(1 - \beta) < 1 + \beta V_O(\phi)$ or $V_O(\phi) >$

$(\beta-\delta)/[\beta(1-\beta)]$. We also have $1/(1-\beta) \geq V_E(\phi)+V_O(\phi)$, so $V_E(\phi) < \delta/[\beta(1-\beta)]$. Algebra and (18) then give $V_E(\tilde{f}) > V_E(\phi)$. Continuity now ensures that $V_E(\tilde{f}+\epsilon) > V_E(\phi)$ for sufficiently small $\epsilon > 0$. This in turn yields $V_E(\tilde{f}+\epsilon) - V_E(f) > V_E(\phi) - V_E(f) \geq 0$ with $e^*(f, \tilde{f}+\epsilon) = 0$ by Result 1.

The opposition is sure to accept $(1, \tilde{f}+\epsilon)$ as the opposition's payoff is strictly greater than its payoff to fighting: $\alpha(1, \tilde{f}+\epsilon) = 1+\beta V_O(\tilde{f}+\epsilon) > 1+\beta V_O(\tilde{f}) = (1-\delta)/(1-\beta)$. Given this acceptance, the elite's payoff to proposing $(1, \tilde{f}+\epsilon)$ and then reverting to σ^* is $\beta V_E(\tilde{f}+\epsilon)$ given $e^*(f, \tilde{f}+\epsilon) = 0$. This cannot be a profitable deviation so $V_E(f,s) \geq \beta V_E(\tilde{f}+\epsilon)$ for arbitrarily small $\epsilon > 0$. Consequently, $V_E(f,s) \geq \beta V_E(\tilde{f})$.

This bound and (11) yield $V_E(f) \geq [(1-r)(1-f) + \beta r V_E(\tilde{f})]/[1-\beta(1-r)]$. But $V_E(\tilde{f}) = (1-r)(1-\tilde{f})/(1-\beta)$ by Lemma 2 which in turn gives $V_E(f) > V_E(\tilde{f})$. This is a contradiction. As shown above, $V_E(\tilde{f}) > V_E(\phi)$ and, by assumption, $V_E(\phi) \geq V_E(f)$. □

Result 3. *Let σ^* be an MPE with $f < \tilde{f}$. If the no-revolution constraint is slack at (y, ϕ) or if it binds with $y < 1$, then $V_O(\phi) > V_O(f)$.*

Proof: Arguing by contradiction, assume $V_O(f) \geq V_O(\phi)$. That (y, ϕ) satisfies the no-revolution constraint means

$$\frac{1-\delta}{1-\beta} \leq y + \beta V_O(f) + \beta[1 - w\rho(e^*(f,\phi))][V_O(\phi) - V_O(f)] \leq 1 + \beta V_O(f).$$

If the no-revolution constraint is slack at (y, ϕ), then the first inequality is strict. If $y < 1$, the second inequality is strict. Either way, $(1-\delta)/(1-\beta) < 1 + \beta V_O(f)$.

Result 1 also implies $e^*(f, f) = 0$. Thus, the opposition's payoff to accepting $(1, f)$ at (f, s) is strictly greater than its payoff to fighting, so it is sure to accept. It follows that if the elite offers $(1, f)$ whenever the opposition is strong it gets $\beta(1-r)(1-f)/(1-\beta)$. This cannot be a profitable deviation and $\tilde{f} > f$, so

$$V_E(f,s) \geq \frac{\beta(1-r)(1-f)}{1-\beta} = \frac{\beta(1-r)(1-\tilde{f})}{1-\beta} + \frac{\beta(1-r)(\tilde{f}-f)}{1-\beta} > \frac{\delta}{1-\beta}.$$

This is a contradiction as $V_O(f,s) \geq (1-\delta)/(1-\beta)$ and $V_E(f,s) + V_O(f,s) \leq 1/(1-\beta)$. □

The next result establishes that the no-revolution constraint binds at the equilibrium offer when $f < \tilde{f}$.

Result 4. *Let σ^* be a power-sharing MPE and let (y^*, ϕ^*) denote the equilibrium offer at (f, s) where $f < \tilde{f}$. Then the no-revolution constraint binds at (y^*, ϕ^*).*

Proof: If $y^* > f$ and the no-revolution constraint is slack, then the elite can profitably deviate by offering a slightly smaller policy concession $y^* - \epsilon$. Arguing by contradiction, assume $y^* = f$ and the no-revolution constraint is slack at (f, ϕ^*). There are two cases to consider.

Case i: $\phi^* > \tilde{f}$. We show the elite has a profitable deviation. The elite's payoff to proposing any (f, ϕ) in the interior of $\mathcal{A}(f)$ is $v_E(f, \phi) = 1 - f - e^*(f, \phi) + \beta w \rho(e^*(f, \phi)) V_E(f) + \beta[1 - w\rho(e^*(f, \phi))] V_E(\phi)$ where $V_E(\phi)$ is given by (18) when $\phi \geq \tilde{f}$. Differentiation then gives

$$\frac{\partial v_E}{\partial \phi} = \Big[-1 + \beta r w \rho'(e^*(f, \phi))[V_E(f) - V_E(\phi)] \Big] \frac{\partial e^*(f, \phi)}{\partial \phi} - \frac{\beta[1 - w\rho(e^*(f, \phi))](1 - r)}{1 - \beta(1 - r)}.$$

If e^* solves the first-order condition (2), then $1 = \beta w \rho'(e^*(f, \phi))[V_E(f) - V_E(\phi)]$ and, consequently, $\partial v_E / \partial \phi < 0$. This, however, leads to a contradiction. Because the no-revolution constraint is slack at (f, ϕ^*), the elite can profitably deviate by offering $(f, \phi^* - \epsilon)$ which the opposition is sure to accept for ϵ small enough.

If $e^*(f, \phi)$ is not an interior solution, then Result 1 implies that the marginal return to effort is so high that $e^* = \rho^{-1}(1/w)$ if this is well-defined or $e^* = \infty$ otherwise. If the former, then $w\rho(e^*) = 1$ and the elite is sure to undermine the agreement in σ^*. It follows that the division of power will remain at f forever. This implies that the opposition's payoff to accepting (f, ϕ^*) is $f/(1 - \beta)$. This contradicts the assumptions that $f < \tilde{f}$ and (f, ϕ^*) satisfies the no-revolution constraint. If the elite's effort is unboundedly large because $\lim_{e \to \infty} \rho^{-1}(e) < 1/w$, then the elite's payoff $v_E(f, \phi^*)$ unboundedly negative. The elite now can profitably deviate to making an offer the opposition is sure to reject

Case ii: $\phi^* \leq \tilde{f}$. The first step is to show that the opposition is sure to accept the offer (ϕ^*, ϕ^*). That the no-revolution constraint is slack at (f, ϕ^*) implies $(1 - \delta)/(1 - \beta) < f + \beta w \rho(e^*) V_O(f) + \beta[1 - w\rho(e^*)] V_O(\phi^*) \leq f + \beta V_O(\phi^*) \leq \phi^* + \beta V_O(\phi^*)$ where we use $V_O(\phi^*) > V_O(f)$ from Result 3. We also have $e^*(\phi^*, \phi^*) = 0$ from Result 1. Hence, the opposition's payoff to accepting (ϕ^*, ϕ^*) is $\phi^* + \beta V_O(\phi^*)$. This is strictly more than the payoff to fighting, so the opposition accepts.

It follows that $V_E(\phi^*, s) \geq (1-\phi^*)/(1-\beta)$ as the elite can always get the latter payoff by offering (ϕ^*, ϕ^*) in every period. This lower bound on $V_E(\phi^*, s)$ means that $V_O(\phi^*, s) \leq \phi^*/(1-\beta)$. But $V_O(\phi^*, s) \geq (1-\delta)/(1-\beta)$, so $\delta \geq 1-\phi^*$. Using $\delta = \beta(1-r)(1-\tilde{f})$ yields the contradiction that $\phi^* > \tilde{f}$. □

An immediate consequence of Result 4 is that $V_O(f)$ is given by (19) for all $f \in [0, 1-\tilde{f})$. (Lemma 2 establishes it for $f \in [\tilde{f}, 1-\delta]$.) The binding no-revolution constraint for $f < \tilde{f}$ implies $V_O(f, s) = (1-\delta)/(1-\beta)$. Substituting this into (9) gives (19).

We next demonstrate that the proposed division of power ϕ in any acceptable offer, i.e, any $(y, \phi) \in \mathcal{A}(f)$, must be at least as large as $\tilde{f}$. One implication of this is that the equilibrium path of the game is very simple. Starting out with the elite in complete control ($f = 0$), the elite proposes some new division of power $\phi^*(0) \geq \tilde{f}$ when the opposition is strong. Either the division of power moves to $\phi^*(0)$ or stays at $f = 0$ until the opposition is again strong and the elite proposes $\phi^*(0)$. Eventually, the division of power moves to $\phi^*(0)$ where Lemma 2 ensures that it remains.

Result 5. *If σ^* is a power-sharing MPE and $(y, \phi) \in \mathcal{A}(f)$, then $\phi \geq \tilde{f}$.*

Proof: There is nothing to show if $f \geq \tilde{f}$. Assume then that $f < \tilde{f}$. The fact that $\phi \geq f$ along with (19) imply $V_O(\phi) \geq V_O(f)$. We also have

$$\frac{1-\delta}{1-\beta} \leq y + \beta V_O(f) + \beta[1 - w\rho(e^*(f, \phi))][V_O(\phi) - V_O(f)] \leq 1 + \beta V_O(\phi)$$

because $(y, \phi) \in \mathcal{A}(f)$. Hence, $V_O(\phi) \geq (\beta - \delta)/[\beta(1-\beta)]$. This lower bound combined with (19) yields $\phi \geq \tilde{f}$. □

We now establish that the elite's effort to undo an agreement on (y, ϕ) is increasing in the stakes, i.e., $\partial e^*(f, \phi)/\partial \phi > 0$.

Result 6. *Let σ^* be a power-sharing MPE with $(y, \phi) \in \mathcal{A}(f)$. If $w\rho(e^*(f, \phi)) \in (0, 1)$, then $\partial e^*(f, \phi)/\partial \phi > 0$.*

Proof: Observe first that $\phi \geq \tilde{f}$ by Result 5. Lemma 2 then gives

$$V_E(\phi) = \frac{(1-\beta)(1-r)(1-\phi) + r\min\{\delta, 1-\phi\}}{(1-\beta)[1-\beta(1-r)]}$$

which is clearly decreasing in ϕ. Moreover, $e^*(f,\phi)$ must satisfy the first order condition because it is an interior solution. Implicit differentiation of that gives

$$0 = \beta w \rho''(e^*(f,\phi))[V_E(f) - V_E(\phi)]\frac{\partial e^*(f,\phi)}{\partial \phi} - \beta w \rho'(e^*(f,\phi))\frac{\partial V_E(\phi)}{\partial \phi}.$$

The concavity of ρ and $\partial V_E(\phi)/\partial\phi < 0$ imply $\partial e^*(f,\phi)/\partial\phi > 0$. □

We now establish that $V_E(f,s)$ is given by (13).

RESULT 7. *Let σ^* be a power-sharing MPE. Then $V_E(f,s)$ is given by (13) for any $f < \tilde{f}$.*

Proof: We have $V_E(f,s) = 1-y-e^*(f,\phi)+\beta w\rho(e^*(f,\phi))V_E(f)+\beta[1-w\rho(e^*(f,\phi))]V_E(\phi)$. Using the binding no-revolution constraint to eliminate y gives

$$V_E(f,s) = \frac{\delta-\beta}{1-\beta} - e^*(f,\phi) + \beta w\rho(e^*(f,\phi))[V_E(f) + V_O(f)]$$
$$+ \beta[1 - w\rho(e^*(f,\phi))][V_E(\phi) + V_O(\phi)]$$

where $\phi \geq \tilde{f}$ by Result 5. Hence, $V_E(\phi) + V_O(\phi) = 1/(1-\beta)$ by Lemma 2. Using (11) and (19) to solve for $V_E(f,s)$ yields (13). □

The next result establishes that the elite makes the maximal policy concession.

RESULT 8. *Let σ^* be a power-sharing MPE, then $y^*(f) = 1$ for any $f < \tilde{f}$.*

Proof: We begin with two observations. First, let (y,ϕ) be any offer at which the no-revolution constraint binds and that the opposition accepts. The elite's payoff to making the offer is given in (5). Using the binding no-revolution constraint to eliminate y, the elite's payoff is

$$b_s(\phi) = \frac{\beta-\delta}{1-\beta} - e^*(f,\phi) + \beta[V_E(f) + V_O(f)]$$
$$+ \beta[1 - w\rho(e^*(f,\phi))]\left(\frac{1}{1-\beta} - [V_E(f) + V_O(f)]\right)$$

where we use $\phi \geq \tilde{f}$ and $V_E(\phi)+V_O(\phi) = 1/(1-\beta)$ from Lemma 2. (See the derivation of (6).) This payoff is decreasing in ϕ as long as $e^*(f,\phi)$ is increasing which Result 6 ensures is the case as long as $w\rho(e^*(f,\phi)) \in (0,1)$.

Observe second that $w\rho(e^*(f,\phi)) < 1$ for all $(y,\phi) \in \mathcal{A}(f)$. Suppose instead that there were a $(y,\phi) \in \mathcal{A}(f)$ for which $w\rho(e^*(f,\phi)) = 1$. Then $(1,\phi) \in \mathcal{A}(f)$ and the no-revolution constraint with $w\rho(e^*(f,\phi)) = 1$ is $1+\beta V_O(f) \geq (1-\delta)/(1-\beta)$. This and (19) yield the contradiction that $f \geq \tilde{f}$.

Arguing by contradiction to show that $y^* = 1$, let (y^*,ϕ^*) denote the equilibrium offer at (f,s) and assume $y^* < 1$. Then, the no-revolution constraint is slack at (y',ϕ^*) for any $y' \in (y^*,1]$. This implies that there exists a $\phi' < \phi^*$ such the no-revolution constraint binds at (y',ϕ').

To verify that such a ϕ' exists, suppose the no-revolution constraint is slack at (y',ϕ) for all $\phi \in [f,\phi^*]$. This means that the opposition is sure to accept (y',f). The elite's payoff to the strategy of offering (y',f) and exerting zero effort whenever the opposition is strong is $1-y'+\beta[(1-r)(1-f)+r(1-y')]/(1-\beta)$. This, however, is strictly larger than $V_E(\tilde{f},s)$ when $y' < 1$. That this strategy cannot be a profitable deviation implies $V_E(f,s) > V_E(\tilde{f},s) = \delta/(1-\beta)$. But $V_E(f,s)+V_O(f,s) \leq 1/(1-\beta)$ and leaves the contradiction $V_O(f,s) < (1-\delta)/(1-\beta)$. Hence, for any $y' \in (y^*,1]$, there is a $\phi' < \phi^*$ such that the no-revolution constraint binds at a (y',ϕ').

Result 2 and $y^* < 1$ imply $V_E(f) - V_E(\phi^*) > 0$. This and the assumption that $\lim_{e\to 0}\rho'(e) = \infty$ ensure $e^*(f,\phi^*) > 0$. Hence, $e^*(f,\phi^*) \in (0,1)$ with Result 6 ensuring that $b_s(\phi)$ is strictly decreasing at ϕ^* and that $\phi^* > \tilde{f}$. Taking y' to be arbitrarily close to y^* yields $b(\phi^*) < b(\phi')$ though there is no guarantee that the opposition accepts since it is indifferent to fighting at (y',ϕ'). However, the opposition is sure to accept $(y'+\epsilon,\phi')$ for any $\epsilon > 0$. It follows that there exists an ϵ small enough that $v_E(y'+\epsilon,\phi') > b(\phi^*)$ and is, therefore, a profitable deviation. This contradiction implies $y^* = 1$. □

We restate Proposition 1 here to take note of possible boundary conditions.

Proposition 1. *A power-sharing MPE exists if and only for every $f \in [0,\tilde{f})$, there exists a $\phi^*(f)$ and $e^*(f,\phi^*(f)))$ that satisfy the optimal effort conditions in Result 1, equation*

(12), and the feasibility conditions $\phi^*(f) \in [f, 1]$ *and* $V_E(f,s) \geq -d$.[32]

Proof: A clear consequence of Results 1-8 is that satifying Result 1 and equation (12) are necessary conditions. To see that these conditions along with the two feasibility conditions are sufficient, suppose that there exists an $e^*(f,\phi)$ and $\phi^*(f)$ that satisfies them. (If more than one pair exists, take the pair with the smallest ϕ^*.) Then consider the profile $\hat{\sigma}$ in which the elite and opposition play according to Lemma 2 when $f \geq \tilde{f}$. For $f < \tilde{f}$, the elite offers $(1, \phi^*(f))$ and exerts effort $e^*(f, \phi^*(f))$. The opposition accepts any offer it weakly prefers to fighting. To complete the specification of $\hat{\sigma}$, we need to define the effort function $e^*(f,\phi)$ for $\phi \neq \phi^*(f)$ and the acceptance set $\mathcal{A}(f)$.

Observe first that (12) implies that $\phi^* \geq \tilde{f}$ when $f < \tilde{f}$. Play therefore follows Lemma 2 once the division of power moves to ϕ^*. We now verify that the opposition accepts the elite's offer when $f < \tilde{f}$. The opposition's payoff to accepting $(1, \phi^*)$ at (f,s) is bounded below by its payoff to getting 1 in the current period, fighting at the next opportunity if the agreement comes undone, and playing according to Lemma 2 if the agreement holds. That is,

$$\alpha(1,\phi^*|\hat{\sigma}) \geq 1 + \beta w \rho(e^*) \left(\frac{(1-\beta)(1-r)f + r(1-\delta)}{(1-\beta)[1-\beta(1-r)]} \right) + \beta[1 - w\rho(e^*)]V_O(\phi^*)$$

where $V_O(\phi^*) = [(1-\beta)(1-r)\phi^* + r\min\{\phi^*, (1-\delta)\}]/[(1-\beta)[1-\beta(1-r)]]$. Equation (12) implies that the expression on the right of the previous inequality equals $(1-\delta)/(1-\beta)$. Hence, the opposition's payoff to accepting is at least as large as its payoff to fighting, so it accepts in $\hat{\sigma}$.

In fact, the opposition's payoff to accepting equals its reservation value: $V_O(f,s) = (1-\delta)/(1-\beta)$. Acceptance implies $V_O(f,s) = 1 + \beta w\rho(e^*)V_O(f) + \beta[1 - w\rho(e^*)]V_O(\phi^*)$ where $V_O(f) = (1-r)[1 + \beta V_O(f)] + rV_O(f,s)$. Solving for $V_O(f,s)$ in terms of $V_O(\phi^*)$ and then using (12) to eliminate ϕ^* yields $V_O(f,s) = (1-\delta)/(1-\beta)$. This in turn implies that the opposition's continuation values $V_O(f|\hat{\sigma})$ are now well defined for all $f \in [0,1]$.

As for the elite's continuation values, the fact that it holds the opposition down to its reservation value means that the argument in the proof of Result 7 can be used to show

[32] Satisfying Result 1 ensures that $e^*(f, \phi^*(f))$ also satisfies $w\rho(e^*(f, \phi^*(f))) \in [0,1]$.

that $V_E(f,s)$ is given by (13) for any $f < \tilde{f}$. $V_E(f)$ follows from (11).

With $V_E(f)$ well defined for $f \in [0,1]$, take $e^*(f,\phi)$ for any $f \in [0,1]$ and $\phi \in [f,1]$ to be the effort function that satisfies Result 1. The acceptance set $\mathcal{A}(f)$ now follows and completes the specification of the profile $\hat{\sigma}$.

The profile $\hat{\sigma}$ is clearly peaceful as the opposition always accepts the elite's offer. We use the one-stage deviation principle to verify that it is an MPE. Note trivially that the opposition has no profitable one-stage deviation as it accepts whenever it weakly prefers that to fighting. It is also that clear that Result 1 ensures that the elite has no profitable one-stage deviation when deciding on its effort. As for the elite's offer, let $(\hat{y},\hat{\phi}) \in \arg\max\{v_E(y,\phi|f,\hat{\sigma}) : (y,\phi) \in \mathcal{A}(f)\}$ where v_E is defined in (5) and $(\hat{y},\hat{\phi})$ is sure to exist because v_E is continuous and $\mathcal{A}(f)$ is compact. We show first that the no-revolution constraint binds at $(\hat{y},\hat{\phi})$.

Arguing by contradiction, suppose the no-revolution constraint is slack at $(\hat{y},\hat{\phi})$. Clearly, $\hat{\phi} > f$. If not, $(1-\delta)/(1-\beta) < \alpha(\hat{y},\hat{\phi}) = \alpha(\hat{y},f) = \hat{y} + \beta V_O(f)$. This leaves $1 + \beta V_O(f) > (1-\delta)/(1-\beta) = 1 + V_O(\tilde{f})$ and the contradiction that $f > \tilde{f}$. Given that $\hat{\phi} > f$, it suffices to show that $v_E(y,\phi)$ is decreasing in ϕ at $(\hat{y},\hat{\phi})$ as $\hat{\phi} > f$ would mean that $(\hat{y},\hat{\phi})$ cannot maximize v_E.

Observe that $\hat{\phi} > \tilde{f}$ and $w\rho(e^*(f,\hat{\phi})) < 1$. This follows from $1 + \beta V_O(\tilde{f}) = (1-\delta)/(1-\beta) < \alpha(\hat{y},\hat{\phi}) \leq \alpha(1,\hat{\phi}) = 1 + \beta w\rho(e^*(f,\hat{\phi}))V_O(f) + \beta[1 - w\rho(e^*(f,\hat{\phi}))]V_O(\hat{\phi})$. This, along with the fact that V_O is increasing, leaves $0 < w\rho(e^*(f,\hat{\phi}))[V_O(\tilde{f}) - V_O(f)] < [1 - w\rho(e^*(f,\hat{\phi}))][V_O(\hat{\phi}) - V_O(\tilde{f})]$ which implies $\hat{\phi} > \tilde{f}$ and $w\rho(e^*(f,\hat{\phi})) < 1$.

We also have $e^*(f,\hat{\phi}) > 0$. Arguing by contradiction, assume $e^*(f,\hat{\phi}) = 0$. Then, $\hat{\phi} > \tilde{f}$ and Result 1 imply $0 \geq V_E(f) - V_E(\hat{\phi}) > V_E(f) - V_E(\tilde{f})$. As a result, $e^*(f,\tilde{f}) = 0$ which in turn implies $v_E(\hat{y},\hat{\phi}) - v_E(\hat{y},\tilde{f}) = \beta V_E(\hat{\phi}) - \beta V_E(\tilde{f}) < 0$. This however contradicts the assumption that $(\hat{y},\hat{\phi})$ maximizes v_E.

Given that $e^*(f,\hat{\phi}) > 0$ and $w\rho(e^*(f,\hat{\phi})) < 1$, $e^*(f,\hat{\phi})$ must satisfy the first order condition $1 = \beta w\rho'(e^*(f,\hat{\phi}))[V_E(f) - V_E(\hat{\phi})]$. This implies that v_E is decreasing in ϕ at

$(\hat{y}, \hat{\phi})$:

$$\frac{\partial v_E(y,\hat{\phi})}{\partial \phi} = \left[-1 + \beta w \rho'(e^*(f,\hat{\phi}))[V_E(f) - V_E(\hat{\phi})]\right] \frac{\partial e^*(f,\hat{\phi})}{\partial \phi} + \beta[1 - w\rho(e^*(f,\hat{\phi}))] \frac{\partial V_E(\hat{\phi})}{\partial \phi}$$

$$= \beta[1 - w\rho(e^*(f,\hat{\phi}))] \frac{\partial V_E(\hat{\phi})}{\partial \phi}.$$

where $\hat{\phi} > \tilde{f}$ and Lemma 2 ensure $\partial V_E(\hat{\phi})/\partial \phi < 0$. It follows that there exists an $\epsilon > 0$ such that $(\hat{y}, \hat{\phi} - \epsilon) \in \mathcal{A}$ and $v_E(\hat{y}, \hat{\phi} - \epsilon) > v_E(\hat{y}, \hat{\phi})$. This contradiction ensures that the no-revolution constraint binds at $(\hat{y}, \hat{\phi})$.

We also have $\hat{y} = 1$. If not, then $\alpha(1, \hat{\phi}) > (1-\delta)/(1-\beta)$. Suppose $\alpha(1, \phi') > (1-\delta)/(1-\beta)$ for all $\phi' \in [f, \hat{\phi}]$. Then $1 + \beta V_O(\tilde{f}) = (1-\delta)/(1-\beta) < \alpha(1, f) = 1 + \beta V_O(f)$ which leads to the contradiction that $f > \tilde{f}$. It follows that there exists a $\phi' < \hat{\phi}$ such that $\alpha(1, \phi') = (1-\delta)/(1-\beta)$. This too leads to a contradiction. The elite's payoff along the binding no-revolution constraint is decreasing in ϕ, so $v_E(1, \phi') = b_s(\phi') > b_s(\hat{\phi}) = v_E(\hat{y}, \hat{\phi})$, contradicting the assumption that $(\hat{y}, \hat{\phi})$ maximizes v_E. Hence, $\hat{y} = 1$. An immediate implication of the facts that $\hat{y} = 1$ and that the no-revolution constraint binds at $(\hat{y}, \hat{\phi})$ is that $\hat{\phi}$ and $e^*(f, \hat{\phi})$ satisfy (12).

Finally, $\hat{\phi} = \phi^*$. Suppose not. Then $\hat{\phi} < \phi^*$ as both $(1, \hat{\phi})$ and $(1, \phi^*)$ lie on the binding no-revolution constraint and b_s is decreasing in ϕ. The pair $\hat{\phi}$ and $e^*(f, \hat{\phi})$ also satisfy Result 1 and (12) as well as the two feasiblity conditions. That is, $\hat{\phi} \in [f, 1]$ since $f < \tilde{f} < \hat{\phi} < \phi^* \leq 1$, and $v_E(1, \hat{\phi}) = b_s(\hat{\phi}) > b_s(\phi^*) = v_E(1, \phi^*) = V_E(f, s) \geq -d$. However, ϕ^* is defined to be the minimum ϕ satisfying these conditions which yields the contradiction $\phi^* \leq \hat{\phi}$. Hence, $(\hat{y}, \hat{\phi}) = (1, \phi^*)$, and there is no profitable one-stage deviation. The profile $\hat{\sigma}$ is subgame perfect. □

We now verify that a unique power-sharing MPE always exists if the institutional commitment power is high enough and that no power-sharing MPEs exist if institutions are too weak.

Proof of Proposition 2. We begin by parameterizing the equilibrium conditions in terms of the elite's effort e and the underlying institutional strength w. Define the division of power $\hat{\phi}(e, w)$ to be the function defined by the right side of (12). Define the elite's payoff

$v_E(e, w)$ to be the right side of (13). Finally, the right side of (14) is the marginal gain from effort. Using (12) when $w\rho(e) \leq (1 - \delta - \tilde{f})/(1 - \delta - f)$ to eliminate ϕ from the right side of (14) gives the elite's marginal gain from effort:

$$\gamma(e, w) \equiv \beta w \rho'(e) \left[\frac{(1-r)(\tilde{f} - f)}{[1 - \beta(1-r)][1 - w\rho(e)]} - \frac{re}{1 - \beta(1-r) - \beta r w \rho(e))} \right].$$

Proposition 1 ensures that there is a power-sharing equilibrium associated with $(\hat{e}, \hat{w})$ if the marginal gain equals the marginal cost, i.e., if $\gamma(\hat{e}, \hat{w}) = 1$, and if the feasibility conditions $\hat{w}\rho(\hat{e}) \leq (1 - \delta - \tilde{f})/(1 - \delta - f)$, $\hat{\phi}(\hat{e}, \hat{w}) \leq 1$, and $v_E(\hat{e}, \hat{w}) \geq -d$ are satisfied. The elite in this equilibrium proposes $y^* = 1$ and $\phi^* = \hat{\phi}(\hat{e}, \hat{w})$ at (f, s) for $f \in [0, \tilde{f})$ and exerts effort $\hat{e}$ where $1 = \gamma(\hat{e}, \hat{w})$.

To establish existence, it suffices to show that such an $(\hat{e}, \hat{w})$ exists when w is small enough. A pair (e, w) satisfies the first two feasibility conditions whenever $e \leq \rho^{-1}((1 - \delta - \tilde{f})/[w(1 - \delta - f)])$ and $e \leq \rho^{-1}((1 - \tilde{f})/[w(1 - f)])$. The third feasibility condition is sure to hold if $e[1 - \beta(1-r)] \leq \delta$. There clearly exists a small enough w' such the first and second conditions hold whenever $w \in [0, w']$ and $e \leq \delta/[1 - \beta(1-r)] \equiv e'$.

Now fix an $e_1 \in (0, e')$ and observe $\lim_{w \to 0} \gamma(e_1, w) = 0$. This implies there is a $\hat{w} \in (0, w')$ such that $\gamma(e_1, \hat{w}) < 1$. The fact that ρ' is unboundedly large at zero also means $\lim_{e \to 0} \gamma(e, \hat{w}) = \infty$. Continuity then ensures that there is an $\hat{e}$ such that $\gamma(\hat{e}, \hat{w}) = 1$.

We now show that no power-sharing MPE exist when institutions are sufficiently weak by demonstrating that the marginal gain from effort always exceeds the marginal cost of one when institutions are very weak. The elite in these circumstances exerts so much effort that the probability that an agreement will hold is so low that the division of power needed to buy the opposition off exceeds one.

Let $g(e, w)$ be the marginal gain from effort which is given by the right side of (14). Using $\phi \geq \tilde{f}$ from Result 5,

$$g(e, w) \geq \beta w \rho'(e) \left(\frac{(1-r)(\tilde{f} - f)}{1 - \beta(1-r)} - \frac{re}{1 - \beta} \right) \equiv \underline{g}(e, w)$$

for all $e \in [0, \rho^{-1}(1/w)]$ where w is assumed to be large enough that $\rho^{-1}(1/w)$ is well defined. The lower bound $\underline{g}$ is also decreasing in e, so $g(e,w) \geq \underline{g}(\rho^{-1}(1/w), w)$ for all $e \in [0, \rho^{-1}(1/w)]$.

When institutions are very weak, the marginal gain in trying to undermine them is very high: $\lim_{w\to\infty} g(e,w) \geq \lim_{w\to\infty} \underline{g}(\rho^{-1}(1/w), w) = \infty$. Thus, there exists a $\overline{w}$ such that $g(e,w) > 1$ for all $e \in [0, \rho^{-1}(1/w)]$ whenever $w > \overline{w}$. Result 1 then implies $e^*(f,\phi) = \rho^{-1}(1/w)$. This however yields a contradiction as the division of power needed to buy the opposition off given by (12) exceeds one.

We now demonstrate that if an equilibrium exists, the subgame perfect equilibrium path starting from any f is unique. Proposition 1 guarantees that the path satisfies (12). The feasibility condition $\phi^*(f) \leq 1$ and (12) imply $w\rho(e^*(f, \phi^*(f)) < 1$. It follows that $\phi^*(f)$ and $e^*(f, \phi^*(f))$ satisfy (14).

It now suffices to show that (12) and (14) only intersect once. Rewrite (12) by defining

$$h_1(e,w) \equiv \tilde{f} + \left(\frac{w\rho(e)}{1 - w\rho(e)}\right)(\tilde{f} - f)$$

$$h_2(e,w) \equiv \frac{\beta - \delta}{\beta} + \left(\frac{w\rho(e)}{1 - w\rho(e)}\right)\left(\frac{(1-\beta)(1-r)}{1-\beta(1-r)}\right)(\tilde{f} - f)$$

where e need not be an equilibrium level of effort. Then $\phi(f)$ as defined by (12) can be expressed as $\phi(f) = \phi_h(e,w) \equiv \min\{h_1(e,w), h_2(e,w)\}$ where we are abusing notation by explicitly including different arguments. Observe further that $0 < \partial h_2/\partial e < \partial h_1/\partial e$ so $\partial h_1/\partial e \geq \partial \phi_h/\partial e$.

Similarly, to solve (14) for ϕ define

$$\eta_1(e,w) \equiv f + \frac{1-\beta(1-r)}{1-r}\left(\frac{1}{\beta w\rho'(e)} + \frac{re}{1-\beta+\beta r[1 - w\rho(e)]}\right)$$

$$\eta_2(e,w) \equiv \frac{(1-\beta)(1-r)f + r(1-\delta)}{1-\beta(1-r)} + \left(\frac{1-\beta}{\beta w\rho'(e)} + \frac{(1-\beta)re}{1-\beta+\beta r[1 - w\rho(e)]}\right)$$

Then $\phi(f)$ as defined by (14) can be expressed as $\phi(f) = \phi_\eta(e,w) \equiv \min\{\eta_1(e,w), \eta_2(e,w)\}$ where $0 < \partial\eta_2/\partial e < \partial\eta_1/\partial e$ with $\partial\phi_\eta/\partial e \geq \partial\eta_2/\partial e$.

Existence when institutions are strong enough insures that ϕ_h and ϕ_η intersect at least once for w small enough. Note that $\lim_{e\to 0} \phi_\eta(e,w) = f < \tilde{f} = \lim_{e\to 0} \phi_h(e,w)$. It follows that ϕ_η and ϕ_h can only intersect once and, hence, that there is a unique equilibrium path if ϕ_η cuts ϕ_h from below whenever they intersect. Differentiation gives

$$\frac{\partial h_1}{\partial e} = \frac{w\rho'(e)(\tilde{f} - f)}{[1 - w\rho(e)]^2} \quad \text{and} \quad \frac{\partial \eta_2}{\partial e} > \frac{-(1-\beta)\rho''(e)}{\beta w[\rho'(e)]^2}.$$

Then ϕ_η is sure to cut ϕ_h from below if $\partial\eta_2/\partial e > \partial h_1/\partial e$ as this will give $\partial\phi_\eta/\partial e \geq \partial\eta_2/\partial e > \partial h_1/\partial e \geq \partial\phi_h/\partial e$.[33] And, $\partial\eta_2/\partial e > \partial h_1/\partial e$ is sure to hold if

$$\frac{-(1-\beta)\rho''(e)}{\beta w[\rho'(e)]^2} > \frac{w\rho'(e)(\tilde{f} - f)}{[1 - w\rho(e)]^2}.$$

Now use $\phi_n \leq 1$ to bound $1 - w\rho(e)$ where, recall, $\phi_n = n_2$ when $\phi_n \in (1-\delta, 1]$. Then the previous inequality is satisfied if

$$\frac{(1-\beta)(1-r)(\tilde{f} - f) + \beta(1-r)}{(1-\beta)(1-r)(\tilde{f} - f) + \delta} > \frac{-\beta w^2[\rho'(e)]^3(\tilde{f} - f)}{(1-\beta)\rho''(e)}.$$

By assumption, $[\rho'(e)]^3/\rho''(e)$ is bounded as $e \to 0$, so the last inequality is certain to hold as $w \to 0$. □

Proposition 3 establishes the comparative static results.

Proof of Proposition 3(i). To see that the elite invests more in undermining agreements when institutions are weak, i.e., $\partial e^*(f, \phi^*))/\partial w > 0$, observe that $\phi_h(e^*(w), w) = \phi_\eta(e^*(w), w)$ at the equilibrium effort $e^*(w)$. To avoid working with one-sided derivatives, assume $\phi_n(e^*(w), w) = \phi_\eta(e^*(w), w) \neq 1 - \delta$. Implicit differentiation gives

$$0 = \frac{\partial\phi_\eta}{\partial e}\frac{de^*}{dw} + \frac{\partial\phi_\eta}{\partial w} - \frac{\partial\phi_h}{\partial e}\frac{de^*}{dw} - \frac{\partial\phi_h}{\partial w}$$

$$\frac{de^*}{dw} = \left[\frac{\partial\phi_h}{\partial w} - \frac{\partial\phi_\eta}{\partial w}\right]\left[\frac{\partial\phi_\eta}{\partial e} - \frac{\partial\phi_h}{\partial e}\right]^{-1}.$$

The second factor in brackets is negative because ϕ_η cuts ϕ_h from below at the unique intersection. Inspection shows that $\partial\phi_h/\partial w > 0$ and $\partial\phi_\eta/\partial w < 0$. Hence, $de^*/dw > 0$.

[33]It straightforward to show that $\partial\eta_2/\partial e > \partial h_1/\partial e$ also ensures that ϕ_η cuts ϕ_h from below if they intersect at the kink where $\phi_\eta(e, w) = \phi_n(e, w) = 1 - \delta$.

The probability of undermining an agreement, $w\rho(e^*(w))$ is clearly increasing too.

Proof of Proposition 3(ii). The claim follows trivially from the facts that ϕ_η cuts ϕ_h from below and, by inspection, ϕ_η is increasing in f and ϕ_h is decreasing in f.

Proof of Proposition 3(iv). The claim that $\partial\phi^*/\partial w < 0$ also follows trivially from analogous observations. By inspection, ϕ_η is decreasing in w and ϕ_h is increasing in w.

Proof of Proposition 3(iii). Similarly, ϕ^* is decreasing in r because ϕ_h is decreasing in r and ϕ_η is increasing. We now demonstrate that there exists a $\underline{\beta} < 1$ such that for any $\beta \in (\underline{\beta}, 1)$, there exists an $\overline{r}(\beta) > 0$ such that no peaceful MPE exists for any $r \in (0, \overline{r}(\beta))$. Observe first that the equilibrium effort is bounded above by $e \in [0, \overline{e}]$ where $w\rho(\overline{e}) = 1$. It suffices to show that $\phi_\eta(e, \beta, r) \neq \phi_h(e, \beta, r)$ for any $e \in [0, \overline{e}]$ when the elite is in complete control ($f = 0$) where it is useful to make the dependence of ϕ_η and ϕ_h on β and r explicit. It is also useful to write $\tilde{f}(\beta, r) = [\beta(1-r) - \delta]/[\beta(1-r)]$.

Clearly, there exists a $\underline{\beta} < 0$ such that

$$\phi_\eta(\overline{e}, \beta, 0) = \frac{1-\beta}{w\rho'(\overline{e})} < \frac{\beta - \delta}{\beta} = \tilde{f}(\beta, 0) \leq \phi_h(0, \beta, 0)$$

for all $\beta \in (\underline{\beta}, 1)$. Continuity ensures that there is an $\overline{r}(\beta) > 0$ such that $\phi_\eta(\overline{e}, \beta, r) < \phi_h(0, \beta, r)$ for all $r \in [0, \overline{r}(\beta)]$. The fact that ϕ_η and ϕ_h are increasing in e implies $\phi_\eta(e, \beta, r) \leq \phi_\eta(\overline{e}, \beta, r) < \phi_h(0, \beta, r) \leq \phi_h(\overline{e}, \beta, r)$. Hence, ϕ_η and ϕ_h do not intersect. $\square$

Appendix B (online)

This appendix demonstrates existence by sketching the construction of a numerical example of a power-sharing MPE which exhibits path dependence and in which some patterns of shocks lead to revolution while others lead to a graduated extension of the franchise. It is useful to distinguish offers and continuation values in the baseline game from those in the extended game. Let $W_j(f,\nu|\sigma)$ for $j \in \{E, O\}$ denote j's continuation value in the extended game starting from (f,ν) given that play follows the Markov profile σ. Similarly, $W_j(f) \equiv (1-\mu-r)W_j(f,n)+\mu W_j(f,m)+rW_j(f,s)$. Take $(z^*_m(f), \theta^*_m(f)) \in [f,1]^2$ and $(z^*(f), \theta^*(f)) \in [f,1]^2$ to be, respectively, the elite's equilibrium offer when facing a moderate or severe threat where we use z and z_m to denote policy offers and θ and θ_m to denote proposed divisions of power in the extended game. Finally, let $u^*(f,\theta) : [f,1]^2 \to \mathbb{R}_+$ be the elite's effort where it is clear from the analysis of the baseline game that effort depends on the stakes but not directly on the intensity of the threat. A Markov strategy for the elite is a five-tuple $\{z^*_m(f), \theta^*_m(f), z^*(f), \theta^*(f),\ u^*(f,\theta)\}$.

In the numerical example, the elite's equilibrium offer when in complete control and facing a severe threat is full democratization ($\theta^*(0) = 1$) at institutional strength $\tilde{w}$. We also show that the equilibrium offer when facing a severe threat, $\theta^*(f)$, is decreasing in f as ϕ^* is in the baseline game (see Proposition 3(ii)). Continuity ensures that there is an $\epsilon > 0$ such that if institutions are slightly weaker than $\tilde{w}$, say $\tilde{w}+\epsilon$, then $\theta^*(0) > 1 > \theta^*(\theta^*_m(0))$. In words, a severe threat to an elite in complete control would end in revolution whereas a gradual transition would have led to a peaceful transition. By contrast, $1 = \theta^*(0) > \theta^*(\theta^*_m(0))$ at $\tilde{w}$. In this case, the elite always averts a revolution but shares more power if it is still in complete control when it faces a severe threat.

Recall that there are three thresholds in the extended game. When $f \geq 1-\delta$, f is so large that the threat to rebel is incredible and no concessions are forthcoming. When $f \in [\tilde{f}_s, 1-\delta)$, the threat is credible, but f is sufficiently large that the elite can buy the opposition off with policy concessions alone. This threshold is the cutpoint at which the opposition's payoff to accepting $(1,f)$ is just equal to its payoff to fighting, i.e., $1 + \beta W_O(\tilde{f}_s) = (1-\delta)/(1-\beta)$. Finally, the opposition weakly prefers the status quo f to fighting and lacks a credible threat as long as $f + \beta W_O(f) \geq \pi(1-\delta)/(1-\beta)$ or, equivalently, $f \geq f_c$ where f_c is satisfies $f_c + \beta W_O(f_c) = \pi(1-\delta)/(1-\beta)$.

To reduce the number of cases and simplify the analysis, assume further that a threat to rebel is incredible when $f \geq \tilde{f}_s$ and the opposition poses a moderate threat. That is, $f + \beta[(1-r)f + r]/(1-\beta) \geq \pi(1-\delta)/(1-\beta)$ for $f \geq \tilde{f}_s$ where the expression on the left of the inequality is the opposition's payoff starting from (f, m) and getting $z = 1$ whenever the opposition is strong and f otherwise.

Assumption 2. *The moderate threat is not too severe:* $\pi \in (0, \overline{\pi})$ *where* $\overline{\pi} = [(1-\beta r)\tilde{f}_s + \beta r]/(1-\delta)$.

Three implications follow from Assumption 2. First, $\tilde{f}_s \geq f_c$ which means: $1-\delta > \tilde{f}_s \geq f_c$. Second, the cutpoints at which the elite can buy the opposition off with concessions alone when facing a severe threat are the same in the baseline and extended games. That is, $\tilde{f}_s = \tilde{f}$.[34] Finally, if σ^* is an MPE of the baseline game and $f \geq f_c$, the fact the opposition only poses a credible threat when the threat is severe means that σ^* is also an MPE in any subgame of the extended game starting from f.

To fix the numerical example, let $\rho(u) = u^{1/2}$, $\delta = .05$, $\beta = .9$, $r = .04$, $\mu = .1$, and $\pi = 1/2$. Severe threats therefore occur in expectation every $1/r = 25$ "years," moderate threats every $1/\mu = 10$ years. The cutpoints are $\tilde{f}_s = 407/432 \approx .9421$ and $f_c = 38/125 = .304$. The institutional strength is $\tilde{w} \approx 0.1797$ and is derived below.

We now construct a Markov profile σ_N satisfying four conditions and then show that it is subgame perfect. The conditions are:

1. The opposition accepts whenever it weakly prefers the proposal on offer to fighting and rebels otherwise.
2. The elite's effort is zero when the stakes are zero ($\theta = f$) and satisfies the first-order condition $1 = w\rho'(u)[W_E(f) - W_E(\theta)]$ otherwise.

[34] When a severe threat is credible ($f < 1-\delta$), $W_O(f)$ is defined recursively by

$$W_O(f) = (1-\mu-r)[f + \beta W_O(f)] + \mu \max\left\{\frac{\pi(1-\delta)}{1-\beta}, f + \beta W_O(f)\right\} + \frac{r(1-\delta)}{1-\beta}.$$

The value of the middle term depends on whether a threat to rebel when the opposition poses a moderate threat is credible. Assumption 2 implies that the maximum is $f + \beta W_O(f)$ when $f \in [\tilde{f}_s, 1-\delta]$. Solving for $W_O(f)$ and then using $1 + W_O(\tilde{f}_s) = (1-\delta)/(1-\beta)$ to derive $\tilde{f}_s$ gives $\tilde{f}_s = \tilde{f}$.

3. When the opposition is strong and poses a credible threat, the elite holds the opposition down to its reservation value while sharing as little power as possible.

4. When the opposition is moderately strong and poses a credible threat, the elite buys the opposition off with the offer $(z_m, \theta_m) \in [f, 1]^2$ that maximizes its payoff along the binding m-NRC.

The opposition's continuation values $W_O(f)$ for $f \in [0, 1]$ follow from the fact that the elite holds the opposition down to its reservation value whenever it poses a credible threat. The treat to rebel is incredible regardless of its intensity when $f \geq 1-\delta$. The elite therefore makes no concessions and the opposition gets $W_O(f) = f/(1-\beta)$. If $f \in [f_c, \tilde{f})$, Assumption 2 implies that a threat to rebel is credible only if the opposition is strong. The elite holds the opposition down to its reservation value in (f, s) and otherwise makes no concessions. This leaves $W_O(f) = (1-\mu-r)[f+\beta W_O(f)]+\mu[f+W_O(f)]+r(1-\delta)/(1-\beta)$. When $f < f_c$, the opposition poses a credible threat to rebel in both (f, s) and (f, m): $W_O(f) = (1-\mu-r)[f+\beta W_O(f)] + \mu[\pi(1-\delta)/(1-\beta)] + r[(1-\delta)/(1-\beta)$. In sum,

$$W_O(f) = \begin{cases} \frac{f}{1-\beta} & \text{if } f \geq 1-\delta, \\ \frac{(1-\beta)(1-r)f+r(1-\delta)}{(1-\beta)[1-\beta(1-r)} & \text{if } f \in [\tilde{f}_c, 1-\delta), \\ \frac{(1-\beta)(1-\mu-r)f+(\pi\mu+r)(1-\delta)}{(1-\beta)[1-\beta(1-\mu-r)]} & \text{if } f < \tilde{f}_c. \end{cases} \tag{20}$$

We specify the elite's continuation values and actions by working "backwards" from high to low values of f. If $f \geq 1-\delta$, a threat to rebel is incredible regardless of its severity. The elite makes no concessions, so $(z^*(f), \theta^*(f)) = (z^*_m(f), \theta^*_m(f)) = (f, f)$ and $W_E(f) = 1/(1-\beta) - W_O(f)$. If $f \in [\tilde{f}_s, 1-\delta)$, the elite buys the opposition off with policy concessions alone when it poses a severe threat as per Lemma 2.

If $f \in [f_c, \tilde{f}_s)$, the opposition's threat to rebel is credible only if it is strong. Conditions (3) and (4) imply $(z^*_m(f), \theta^*_m(f)) = (f, f)$ and $z^*(f) = 1$. The division of power $\theta^*(f)$ and effort $u^*(f, \theta^*(f))$ satisfy the following:

$$\frac{1-\delta}{1-\beta} = 1 + \beta w\rho(u^*)W_O(f) + \beta[1 - w\rho(u^*)]W_O(\theta^*) \tag{21}$$

$$1 = \beta w\rho'(u^*)[W_E(f) - W_E(\theta^*)] \tag{22}$$

$$\begin{aligned} W_E(f) = (1-\mu-r)[1-f+\beta W_E(f)] + \mu[1-f+\beta W_E(f)] \\ + r\big[-u^* + \beta[w\rho(u^*)W_E(f) + (1-w\rho(u^*))W_E(\theta^*)]\big] \end{aligned} \tag{23}$$

where we drop the arguments to simplify the notation. Equation (21) is the binding s-NRC, and (22) is the first-order condition defining the optimal effort. Equation (23) recursively defines W_E when the elite only makes a concession when facing a severe threat. Call this Problem 1.

To solve it, observe that θ^* must be at least as large as $\tilde{f}_s$ in order to satisfy the s-NRC, so $W_E(\theta^*) = 1/(1-\beta) - W_O(\theta^*)$. Solving (22) for $W_E(f)$ and eliminating $W_E(f)$ from (23) yields two equations in the two variables $u^*(f, \theta^*(f))$ and $\theta^*(f)$ which can be solved numerically.[35] Substituting these results in (23) gives $W_E(f)$ for any $f \in [f_c, \tilde{f})$. Finally, these continuation values can be used to solve (22) for $u^*(f, \theta)$ for any $\theta > f$ where θ need not be an equilibrium offer.

If $f \in [0, f_c)$, the opposition's threat to rebel is always credible. Condition (3) implies $z^*(f) = 1$ with $z^*_m(f)$, $\theta^*_m(f)$, $\theta^*(f)$, $u^*(f, \theta)$, and $W_E(f)$ to be determined. Two cutpoints $0 < k_2 < k_1 < f_c$ are relevant. As shown below, when f is "close" to f_c, i.e., when $f \in [k_1, f_c)$, the constraint $z^*_m(f) \geq f$ binds. Call this Problem 2. When $f \in [k_2, k_1]$, there is an interior solution, and $\phi^*_m(f)$ satisfies the first-order condition (15) where $\theta^*_m(f) \geq f_c$ and, as a result, $W_E(\theta^*_m(f))$ is given by the solution to Problem 1. Call this Problem 3. Problem 4 arises when $\theta^*_m(f)$ drops below f_c. There continues to be an interior solution with $\phi^*_m(f)$ satisfying the first-order condition (15). However, $W_E(f)$ is now given by the solution to Problem 2 rather than Problem 1.

To solve Problem 2 and determine the elite's actions for $f \in [k_1, f_c)$, observe first that the division of power $\theta^*(f)$ and effort $u^*(f, \theta^*(f))$ continue to satisfy (21) and (22). Assuming $z^*_m(f) = f$, then $\theta^*_m(f)$ and $u^*_m(f, \theta^*_m(f))$ satisfy

[35] Mathematica notebook detailing the numerical analysis available on request.

$$\frac{\pi(1-\delta)}{1-\beta} = f + \beta w\rho(u_m^*)W_O(f) + \beta[1-w\rho(u_m^*)]W_O(\theta_m^*) \tag{24}$$

$$1 = \beta w\rho'(u_m^*)[W_E(f) - W_E(\theta_m^*)] \tag{25}$$

$$\begin{aligned} W_E(f) = &(1-\mu-r)[1-f+\beta W_E(f)] \\ &+ \mu\big[1-f-u_m^* + \beta[w\rho(u_m^*)W_E(f) + (1-w\rho(u^*))W_E(\theta_m^*)]\big] \\ &+ r\big[-u^* + \beta[w\rho(u^*)W_E(f) + (1-w\rho(u^*))W_E(\theta^*)]\big]. \end{aligned} \tag{26}$$

This system reduces to two equations in θ_m^* and θ^* which are solved numerically.[36]

Turing to Problem 3, which assumes the constraint $z_m^*(f) \geq f$ no longer binds, we have $z^*(f) = 1$ which leaves $z_m^*(f)$, $\theta_m^*(f)$, $\theta^*(f)$, $u^*(f, \theta)$, and $W_E(f)$ to be determined. They satisfy (21), (22), and (25). The binding m-NRC (24) becomes

$$\frac{\pi(1-\delta)}{1-\beta} = z_m^* + \beta w\rho(u_m^*)W_O(f) + \beta[1-w\rho(u_m^*)]W_O(\theta_m^*). \tag{27}$$

The optimal offer must maximize the elite's payoff along the binding m-NRC and satisfy the first-order condition $0 = \partial b_m/\partial\theta_m$ where $\partial b_m/\partial\theta_m$ is defined by (15). Using (16) to simplify this condition gives

$$\begin{aligned} 0 = &-\frac{\beta^2 w^2[\rho'(u_m^*)]^3}{\rho''(u_m^*)}\left(\frac{\partial W_E(\theta_m^*)}{\partial\theta_m}\right)[W_O(\phi_m) - W_O(f)] \\ &+ \beta\left[1 - w\rho(u_m^*)\right)]\frac{\partial(W_E(\theta_m^*) + W_O(\theta_m^*))}{\partial\theta_m}. \end{aligned} \tag{28}$$

Finally, the continuation value $W_E(f)$ is

[36] Solve (21) for u^*, (22) for $W_E(f)$, and (25) for u_m^*. Using these results, the recursive definition of $W_E(f)$ in (26) reduces to an equation involving only θ_m^* and θ^*. Similarly, the binding m-NRC (24) can be reduced to an equation involving only θ_m^* and θ^*. This solution assumes $\theta_m^*(f) \geq f_c$ which means that $W_E(\theta_m^*(f))$ is given by the solution to Problem 1.

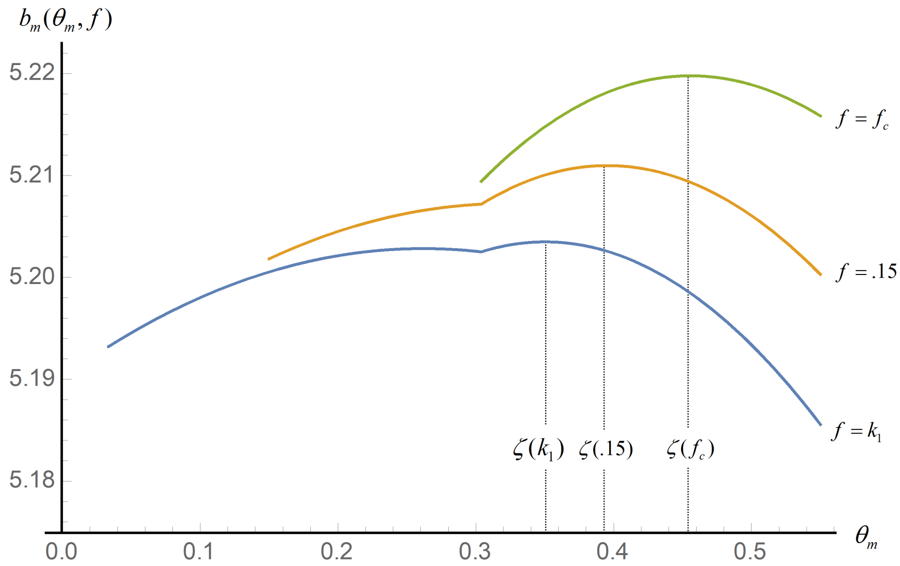

Figure 2: The elite's payoff along the binding m-NRC.

$$W_E(f) = (1-\mu-r)[1-f+\beta W_E(f)] \\ +\mu\big[1-z_m^*-u_m^*+\beta[w\rho(u_m^*)W_E(f)+(1-w\rho(u^*))W_E(\theta_m^*)]\big] \\ +r\big[-u^*+\beta[w\rho(u^*)W_E(f)+(1-w\rho(u^*))W_E(\theta^*)]\big]. \tag{29}$$

Equations (21), (22), (25), and (27)-(29) can be reduced to two equations in θ_m^* and θ^* and solved numerically.[37]

Figure 2 provides some intuition. It plots the elite's payoff to offering $\theta_m \geq f$ along the binding m-NRC, $b_m(\theta_m, f)$, where θ_m need not be the equilibrium offer and we disregard the constraint that $z_m \geq f$.[38] Suppose the existing distribution of power is f_c. The payoff $b_m(\theta_m, f_c)$ is clearly increasing at $\theta_m = f_c \approx .304$, and the elite's unconstrained optimal is $\theta_m = \zeta(f_c)$. That is, the elite would like to slide along the binding m-NRC

[37]Paralleling the procedure outlined in footnote (36), solve (21) for u^*, (22) for $W_E(f)$, (25) for u_m^*, and (27) for z_m^*. Using these results, $W_E(f)$ defined in (29) reduces to an equation involving only θ_m^* and θ^*. Similarly, the first-order condition (28) reduces to an equation involving only θ_m^* and θ^*. This solution also assumes $\theta_m^*(f) \geq f_c$ which means that $W_E(\theta_m^*(f))$ and given by the solution to Problem 1.

[38]For a partial intuition about the kink at $f = f_c$, note from (20) that the derivative of $W_O(f)$ at $f > f_c$ is $(1-r)/[1-\beta(1-r)]$ and $(1-\mu-r)/[1-\beta(1-\mu-r)]$ at $f < f_c$.

from f_c to $\zeta(f_c)$, giving up greater institutional concessions (higher θ_m) in return for smaller policy concessions (lower z_m). The elite, however, cannot offer $\zeta(f_c)$ or any other $\theta_m > f_c$ in return for a smaller z_m because the constraint $z_m \geq f_c$ binds when the distribution of power is f_c. To put the point somewhat differently, recall that f_c solves $f_c + \beta W_O(f_c) = \pi(1-\delta)/(1-\beta)$ which is to say that $(z_m, \theta_m) = (f_c, f_c)$ lies on the binding m-NRC. Further, the set of feasible offers at f_c is the upper-right quadrant $[f_c, 1]^2$ which only intersects the binding m-NRC at (f_c, f_c). Hence, the elite cannot move from (f_c, f_c) to any other point on the binding m-NRC given the assumption that $(z_m, \theta_m) \in [f_c, 1]^2$.[39]

The constraint $z_m^*(f) \geq f$ continues to bind when f is close to f_c. To see where it ceases to bind, solve the binding m-NRC for the policy concession implied by the optimal offer $\zeta(f)$:

$$z_m(\zeta(f)) = \frac{\pi(1-\delta)}{1-\beta} - \beta w \rho(u_m^*) W_O(f) - \beta[1 - w\rho(u_m^*)] W_O(\zeta(f))$$

The difference $z_m(\zeta(f)) - f$ increases and eventually equals zero as f drops farther and farther below f_c. This defines the cutpoint $k_1 \approx .0336$. The constraint $z_m(f) \geq f$ no longer binds when $f < k_1$, and the solution to Problem 3 specifies the elite's optimal actions.

Note, however, that as f decreases, the height of the right "hump," $b_m(f, \zeta(f))$, decreases while the height of the left "hump" increases. The cutpoint $k_2 \approx .006$ is the value of f a which the heights of the two local maxima are also global maxima (see Figure 3). When $f > k_2$, the global maximum is the local maximum on the right. When $f < k_2$, the global maximum is the local maximum on the left.

When f drops below k_2, the optimal offer $\zeta(f)$ drops below f_c. The assumption in Problem 3 that $\theta_m^*(f) \geq f_c$ and, as a result, that $W_E(\theta_m^*(f))$ is given by the solution to Problem 1 no longer holds (see footnote 37). Rather, $\theta_m^*(f) \in [k_1, f_c)$ and $W_E(f)$ is given

[39] Clearly, the reduced-form assumption that the policy and institutional offers at f are bounded below by f precludes what would otherwise be Pareto improving trades when the opposition poses a moderate threat. This issue does not arise in the baseline game or in the extended game when the opposition poses a severe threat. If $f \in [\tilde{f}_s, 1-\delta)$, then $W_E(f) = 1/(1-\beta) - W_O(f)$ and there are no Pareto improving trades along the binding s-NRC. If $f < \tilde{f}_s$, the elite's payoff along the binding s-NRC is decreasing in θ and the constraint $z \leq 1$ binds at the optimal offer.

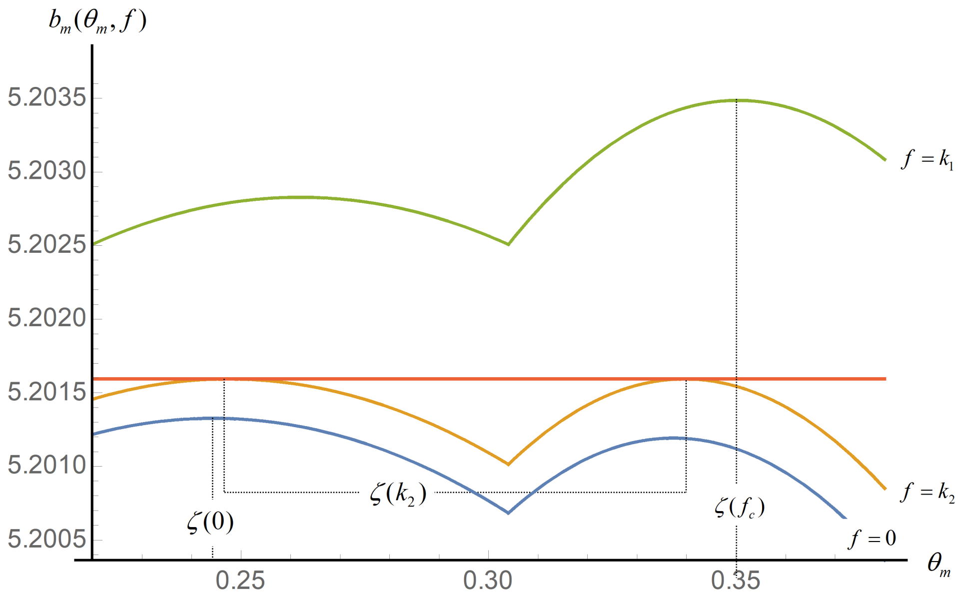

Figure 3: $\theta^*_m(f)$ dropping below f_c.

by the solution to Problem 2. This is Problem 4. Its solution defines the elite's actions and continuation payoff for $f \in [0, k_2)$, and this completes the specification of σ_N.[40]

Note that this example has three absorbing states. If the elite is in complete control when if faces a severe threat and cuts a deal that holds, the resulting division of power is $\theta^*(0)$. Suppose instead that the elite faced a moderate threat when it cut a deal that held. The division of power moves to $\theta^*_m(0) \approx 0.2444 < f_c$. If the elite now faces a severe threat and cuts a deal that holds, the distribution of power moves to $\theta^*(\theta^*(0))$ where it remains. If, by contrast, the elites next faces a moderate threat when it cuts a deal that holds, the distribution of power moves to $\theta^*_m(\theta^*_m(0)) > f_c$ and then ultimately to $\theta^*(\theta^*_m(\theta^*_m(0)))$.

We establish that σ_N is subgame perfect by showing that there is no profitable one-stage deviation. This is trivially true for the opposition whose strategy is to accept any offer it weakly prefers to fighting. As for the elite, there is nothing to show for $f \geq \tilde{f}_s$ as the actions specified by σ_N are in accord with Lemma 2. For $f < \tilde{f}_s$, observe first that $W_E(f)$ is decreasing as illustrated in Figure 4. This ensures that the elite's best reply when facing a credible threat is to hold the opposition down to its reservation value so

[40] This problem is solved in the same way as Problem 3 except that $W_E(f)$ is given by the solution to Problem 2 instead of Problem 1.

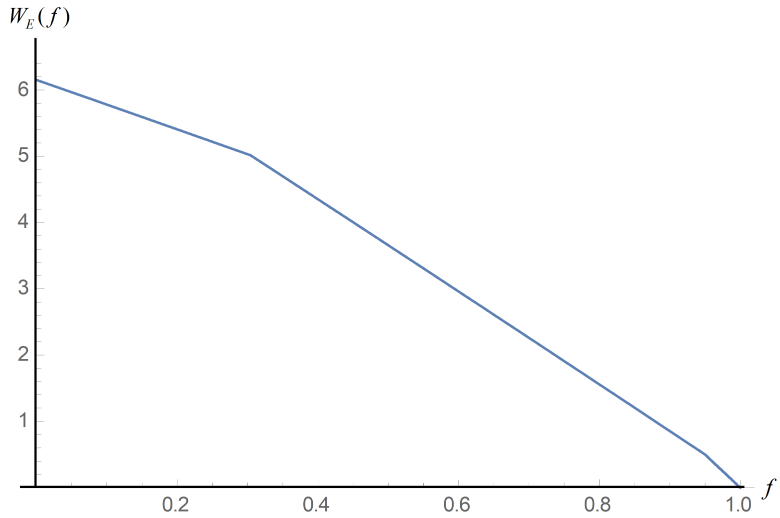

Figure 4: The elite's continuation payoff $W_E(f)$.

that the relevant no-revolution constraint binds. If it did not bind and $z > f$, then the elite could profitably deviate by offering a slightly smaller policy concession. If $z = f$, the elite could profitably deviate by offering slightly less power. Hence, the elite's optimal offer lies along the binding no-revolution constraints as is the case in σ_N.

As shown above, $b_s(\theta)/\partial\theta < 0$, so the optimal offer when facing a severe threat is to make the maximal policy concession ($z = 1$) and share as little power as possible. This too is what σ_N says to do via Condition (3). The construction described above also has the elite making the making the offer that maximizes its payoff along the binding m-NRC. Hence, there are no profitable one-stage deviations when the elite is making an offer.

As for the elite's effort decisions, they satisfy the first-order conditions (22) and (25) and are optimal as long as they are feasible, i.e., as long as probabilities $w\rho(u^*)$ and $w\rho(u^*_m)$ are between zero and one, which they are.

The construction of the profile and the solutions to Problems 1-4 assume a particular value of w, namely, $\tilde{w} = 0.1797$ at which $\theta^*(0) = 1$ when the elite is in complete control ($f = 0$). To determine the value of $\tilde{w}$, solve Problem 4 with the value of w left unspecified. The procedure outlined above yields two equations in f, w, θ^*, and θ^*_m. These were solved above assuming $\tilde{w} = 0.1797$ to obtain the equilibrium proposals θ^*_m and θ^* as functions of f. To determine $\tilde{w}$, assume $f = 0$ and $\theta^* = 1$ and solve the two equations for w and θ^*_m.

The solution yields $\tilde{w}$.

Finally, $\theta^*(f)$ is decreasing in f. It follows that there exists an $\epsilon > 0$ such that $\theta^*_m(0) > 1 > \theta^*(\theta^*_m(0))$ for $w = \tilde{w} + \epsilon$.