Stronger Challengers can Cause More (or Less) Conflict and Institutional Reform*

Andrew Little† Jack Paine‡

March 22, 2022

Abstract

Prominent theories propose that phenomena such as war and democratization occur when rulers cannot commit to future promises toward challengers. Different variants of these theories give divergent answers to a key question: how does the strength of a challenger affect prospects for bargaining breakdown and/or institutional reform? We provide a new answer by analyzing a model with a general distribution of the probability that the challenger would win a conflict in a given period ("threat"). The effect of the challenger's underlying coercive strength depends on the relationship between their average and maximum threat. When the maximum threat is fixed and high, inherently weak challengers are prone to rebel in rare periods when they pose a high threat. However, if only inherently strong challengers pose a high maximum threat, then they are harder to buy off. These theoretical insights uncover key parameters on which empirical research must focus, which we apply to existing debates about democratization.

*Many thanks to Otto Kienitz for excellent research assistance.
†Associate Professor, Department of Political Science, University of California Berkeley.
‡Associate Professor, Department of Political Science, University of Rochester.

1 Introduction

Why do countries vary in their incidence of civil or international conflict? Why do some countries democratize? Under what conditions do dictators share power? To explain these various substantive phenomena, scholars often appeal to dynamic commitment problems.[1] The core intuition for this mechanism is that rulers face impediments to buying off challengers that pose transient threats. A temporarily strong challenger can leverage their present threat to garner a favorable distribution of spoils. However, the transitory nature of their threat inhibits the ruler from committing to spoils distribution in the future. The ruler's commitment problem can trigger either conflict or institutional reform.

This style of argument is pervasive because the core intuition is straightforward, compelling, and broadly applicable. However, a key question remains ambiguous: how does the coercive strength of a challenger affect prospects for bargaining breakdown and/or institutional reform? We demonstrate that a common and seemingly innocuous assumption to simplify the distribution of threats leads to conclusions that do not generalize. By extending this class of models to consider a more general distribution of threats, we provide new theoretical implications for bargaining breakdown and institutional reform, clarify key parameters on which empirical tests of these models must focus, and explore more general foundations for a commonly studied model.

Our departure point from existing work is to scrutinize a common simplifying assumption: the distribution of threats for the challenger is binary, either strong or weak. In this setup, a natural way to capture the "strength of the challenger" is the probability that the challenger is strong in any period. Intuitively, one might expect that "stronger" challengers are harder to buy off and more likely to initiate conflict. However, a higher frequency of strong periods produces the opposite

[1] For civil conflict, see Fearon (2004); Chassang and Padro-i Miquel (2009); Walter (2009); Powell (2012); Gibilisco (2021). For international war, see Fearon (1995); Powell (1999); Debs and Monteiro (2014); Krainin (2017). For democratization, see Acemoglu and Robinson (2006); Ansell and Samuels (2014); Dower et al. (2018). For authoritarian power sharing and democratic separation of powers, see Helmke (2017); Christensen and Gibilisco (2020); Meng (2019); Paine (2021).

result: conflict occurs along the equilibrium path only if the challenger is sufficiently weak, and thus only rarely face opportunities to coerce the regime. Their rare moments in the sun are too tempting to pass up and forgo revolution, given their poor prospects for gaining concessions in the future if the status quo regime remains intact. For the same reason, the ruler faces greater incentives to extend the franchise or share power with weak challengers.

We extend this class of stochastic infinite-horizon bargaining models by considering a more general distribution of threats. To our knowledge, ours is the first to consider any discrete or continuous distribution of threat levels. Our main theoretical finding is that prospects for conflict in equilibrium hinge on the relationship between the *average* and *maximum* threat posed by the challenger. Different notions of "increasing the strength" of the challenger have different implications for this relationship, and can either improve or hinder prospects for peaceful bargaining or institutional reform.

Despite the virtue of analytical tractability, the common assumption that the challenger fluctuates between a strong and a weak state is a special case in which the maximum probability with which the challenger wins a fight is fixed (often at 1), and the comparative statics on challenger strength alter only the frequency of such "strong" periods. Under this distribution, increasing the strength of the challenger reduces prospects for bargaining breakdown by bolstering their average threat—which raises their total concessions across the infinite horizon—without affecting their maximum threat.

However, an increase in the strength of the challenger could also raise both their maximum and average threat levels. This makes the opposite relationship possible: stronger challengers are more likely to fight or gain institutional concessions. This occurs whenever a shift in the strength parameter raises the maximum threat by at least as much as the average, and can occur even if the average increases at a somewhat higher rate.[2] For example, a uniform rightward shift of the

[2]This finding complements that from Powell (2004). He presents general conditions under which a large and rapid shift in the distribution of power triggers fighting. We tie a similar mechanism to the "strength" of the challenger and show that a stronger challenger can make the sufficient condition for fighting in Powell's model either more or less

distribution of challenger strength makes them harder to buy off peacefully. Such a shift also makes institutional reform more likely.

Beyond "challenger strength" specifically, our theoretical results provides a new lens to study the effects of many possible stimuli. For example, exercising repression may either increase or decrease prospects for conflict, depending on how it changes the distribution of the challenger's probability of winning. If repression creates a uniform downward shift in these probabilities, the probability of conflict and the need to offer institutional reform will decrease. By contrast, if repression usually prevents people from mobilizing but creates rare instances where they are able to forge cross-class coalitions, such regimes might be subject to revolutionary outbursts because the maximum threat is high whereas the average threat is low—hence leaving challengers "no other way out" than revolution (Goodwin, 2001).

We conclude by discussing implications for debates about democratization and authoritarian power sharing. In models such as Acemoglu and Robinson (2006) and Dower et al. (2018), weak challengers trigger institutional reform. A low average threat makes the shadow of the future unfavorable. This, combined a high maximum threat, bolsters their bargaining leverage in a rare strong period. However, other seemingly similar models yield different implications about challenger strength (Ansell and Samuels, 2014; Meng, 2019; Paine, 2021). Our model explains the conditions under which we recover each implication. These findings also carry implications for empirical research designs that test these models. Although existing studies propose innovate ways to measure key parameters, they do not consider the countervailing effects of higher maximum and average threats. Future work must push on this frontier to extend our understanding of how the strength of societal challenges affects prospects for conflict and institutional reform.

likely to bind.

2 Model

2.1 Setup

A ruler and challenger bargain over spoils in periods $t = 1, 2, ...$ with a common discount factor $\delta \in (0, 1)$. We normalize total spoils in each period to 1. In each period, the ruler makes a take-it-or-leave it offer $x_t \leq 1$. That is, we impose the common assumption in this literature that the ruler cannot transfer more than the entire contemporaneous budget in any period, and hence cannot borrow across periods. To ease exposition, we first impose an unconventional assumption that the ruler can offer $x_t < 0$, which means that the ruler can demand a transfer from the challenger. Later we show that the core insights are qualitatively similar when we impose a lower bound on x_t.[3]

If the challenger accepts in some period t, then the ruler and challenger respectively consume $(1-x_t, x_t)$ and engage in a strategically identical interaction in period $t+1$. If instead the challenger rejects in period t, then conflict occurs. Fighting is a game-ending move that permanently destroys $\phi \in (0, 1)$ in each period, and the winner consumes all remaining spoils.

The challenger's probability of winning a conflict varies by period. The parameter is p_t, which depends on an independent and identically distributed choice by Nature revealed to both players at the outset of each period.[4] Thus, at the bargaining stage, both actors are perfectly informed about p_t. We call p_t the *threat* posed by the challenger in period t. The distribution function of p_t is $F(p; s)$, where s is a parameter that captures the general strength of the challenger. The distribution has mean $\bar{p}(s) \equiv \mathbb{E}[p; s]$ and support on $\left[p^{\min}(s), p^{\max}(s)\right]$, for $0 \leq p^{\min} < p^{\max} \leq 1$. To capture the general notion that stronger challengers tend to pose a higher threat, assume that $\bar{p}(s)$, $p^{\min}(s)$, and $p^{\max}(s)$ each weakly increase in s. To streamline the exposition, we suppress s where it does not cause confusion.

[3]The value of the lower bound simply shifts the threshold at which a peaceful equilibrium exists. The case without a lower bound is analytically simpler because the ruler can hold down the challenger to his reservation value in every period. This implies the offer is linear in p_t, allowing for a straightforward characterisation of how the distribution of p_t affects equilibrium offers.

[4]In Appendix A.2, we allow for path-dependent states. If the state remains the same in the next period with some probability, the analysis is qualitatively unchanged.

2.2 The Distribution of Threats and Conflict

We examine the conditions under which a Markov Perfect Equilibrium (MPE) exists in which conflict occurs with probability 0 along the equilibrium path. We refer to this as a peaceful equilibrium.

Along a peaceful equilibrium path, in every period t, the ruler makes an offer $x_t \leq 1$ that the challenger accepts. In any equilibrium, the challenger accepts only offers for which its lifetime expected stream of consumption along a peaceful path weakly exceeds the value of its fighting outside option. Thus, if we write the challenger's future continuation value along a peaceful path as V^C, a necessary condition for peaceful bargaining in any period t is:

$$\underbrace{x_t + \delta \cdot V^C}_{\text{Accept}} \geq \underbrace{p_t \cdot \frac{1-\phi}{1-\delta}}_{\text{Fight}}. \tag{1}$$

Given our present assumption that x_t is not bounded from below, the ruler never makes offers that the challenger strictly prefers to accept. Otherwise, the ruler could profitably deviate by making a slightly lower offer that the challenger would accept. Consequently, Equation (1) must hold with equality for every period t. The optimal transfer in every period must satisfy:

$$x^*(p_t) = p_t \cdot \frac{1-\phi}{1-\delta} - \delta \cdot V^C. \tag{2}$$

In a peaceful MPE in which the ruler uses this offer function in every period, we can write the continuation value as equal to the average offer made divided by $1-\delta$. A convenient aspect of the optimal offer is that it is linear in the current-period threat p_t, and hence the average value of p_t is the only aspect of the distribution which affects the continuation value. As demonstrated in Appendix A.3, this property holds in any equilibrium with conflict as well.

Formally, we can write the continuation value as:

$$V^C = \frac{1}{1-\delta} \cdot \underbrace{\int_{p^{\text{min}}}^{p^{\text{max}}} \left[p_t \cdot \frac{1-\phi}{1-\delta} - \delta \cdot V^C \right] \cdot dF(p)}_{\text{Average per-period transfer}}$$

$$\implies V^C = \frac{1}{1-\delta} \left[\bar{p} \cdot \frac{1-\phi}{1-\delta} - \delta \cdot V^C \right]$$

$$\implies V^C = \frac{1-\phi}{1-\delta} \cdot \bar{p} \tag{3}$$

Combining Equations 2 and 3 enables us to solve for the equilibrium per-period offer:

$$x^*(p_t) = \frac{1-\phi}{1-\delta} \cdot \big(p_t - \delta \cdot \bar{p}\big). \tag{4}$$

A peaceful equilibrium requires that the challenger can be bought off in every period. Equation 4 makes clear that this condition is most difficult to satisfy in a period that the challenger poses the highest threat, which we formalize in Proposition 1.[5]

Proposition 1 (Existence of a peaceful equilibrium)**.** *The following inequality is a necessary and sufficient condition for a peaceful equilibrium to exist:*

$$\frac{1-\phi}{1-\delta} \cdot \big(p^{max} - \delta \cdot \bar{p}\big) \leq 1.$$

The condition in Proposition 1 enables us to take comparative statics on the challenger's strength, s. Rearranging the no-fighting constraint in Equation 1 and explicitly writing the distribution parameters as a function of s gives:

$$\frac{1-\delta}{1-\phi} \geq p^{\text{max}}(s) - \delta \cdot \bar{p}(s) \equiv \tau(s). \tag{5}$$

[5]This finding requires the assumption of iid shocks across periods, as we discuss in Appendix A.2.

The overall effect of increasing the strength of the opposition s can be summarized by how it affects the $\tau(s)$ term. If increasing s raises τ, then Equation 1 is harder to meet and hence a peaceful equilibrium is harder to sustain. We can write this derivative

$$\frac{\partial \tau(s)}{\partial s} \equiv \frac{\partial p^{\max}(s)}{\partial s} - \delta \cdot \frac{\partial \bar{p}(s)}{\partial s}. \tag{6}$$

This equation clearly highlights our main point about the need to compare the maximum and average probabilities of winning. These parameters exert countervailing effects on prospects for conflict. On the one hand, higher s makes conflict more likely by raising $p^{\max}$. When the challenger poses its maximum threat, the opportunity cost of not fighting is largest. The challenger's expected probability of winning in the next period is lower than in the current period because they currently have the maximum draw. This creates the temptation to fight now to "lock in" their temporary advantage.

On the other hand, higher s makes conflict less likely by raising $\bar{p}$. From the perspective of a period with the maximum draw, the magnitude of the adverse shift in the future distribution of power depends on the challenger's average probability of winning. High $\bar{p}$ lowers the opportunity cost of not fighting in a period with $p_t = p^{\max}$. The challenger expects to continue to get favorable draws of p_t in the future along a peaceful path, which diminishes their incentives to fight now.

General binary distribution To connect this result to past work, suppose the challenger threat takes on one of two values, which we write as $p_t \in \{p^{\min}, p^{\max}\}$, with $q = Pr(p_t = p^{\max})$. In this case, average threat is $(1-q) \cdot p^{\min} + q \cdot p^{\max}$. Substituting this into Equation 6 and taking comparative statics yields

$$\frac{\partial \tau}{\partial s} = \underbrace{(1-\delta \cdot q) \cdot \frac{\partial p^{\max}}{\partial s}}_{\text{Maximum strength (+)}} - \underbrace{\delta \cdot (1-q) \cdot \frac{\partial p^{\min}}{\partial s}}_{\text{Minimum strength (-)}} - \underbrace{\delta \cdot \frac{\partial q}{\partial s} \cdot (p^{\max} - p^{\min})}_{\text{Frequency of strong periods (-)}} \tag{7}$$

This equation clarifies the intuition for the result from Acemoglu and Robinson (2006) and other models that an increase in the challenger's strength makes it *easier to buy them off*; or, conversely in Fearon (2004), that an decrease in the government's strength makes civil war less likely to occur. In a distribution in which the minimum and maximum threats are fixed, the first two terms in Equation 7 are 0. Hence, challenger strength affects only the frequency of strong periods. In this case, higher s improves the shadow of the future along a peaceful path. However, higher s does not increase the opportunity cost of fighting in the maximum-threat state because $p^{\max}$ is not a function of s.

Our analysis also suggests a sense in which we can generalize this finding. For any distribution shift such that the bounds (and in particular the upper bound) are fixed but the average increases, it will be easier to buy off the challenger peacefully. With a binary distribution, this implies fixing $p^{\max}$ and raising either $p^{\min}$ or q.

However, this result does not hold for all increases in strength, even if the distribution of threats is binary. For example, a natural case to consider is one in which both the minimum and maximum threat increase at the same rate, $\frac{\partial p^{\max}}{\partial s} = \frac{\partial p^{\min}}{\partial s} = d_s > 0$; but the relative probability of the two values does not, $\frac{\partial q}{\partial s} = 0$. This would arise if, for example, increasing s leads to a uniform rightward shift of F. In this case, *peace is harder to sustain* in the face of a strong challenger because $\frac{\partial \tau}{\partial s} = (1-\delta)d_s > 0$ (the result follows directly from substituting this case into Equation 7).

This example highlights a useful fact for future theorizing: a binary distribution in of itself does not greatly limit the generality of insights from models with dynamic commitment problems. Even with a simple distribution, increasing the challenger's strength can either increase or decrease prospects for conflict. Instead, the important takeaway is that *how the researcher structures the parameters* in the distribution determine the direction of the comparative statics prediction. In a binary distribution, there are three key parameters, and different changes carry divergent implications for the prospect of peace.

2.3 Prospects for Institutional Reform

We have shown that the challenger's endowed strength parameter exhibits ambiguous consequences for conflict. The intuition is qualitatively similar when we allow the possibility of institutional reform, which in reality can mean either sharing power within an authoritarian regime or full-blown democratization. Specifically, in Appendix A.1, we assume that at the outset of the game, the ruler chooses a basement level of spoils that the challenger receives in every period, $\underline{x} \in [-\infty, 1]$. In each period, the choice set for the transfer is $x_t \in [\underline{x}, 1]$. We can interpret higher levels of $\underline{x}$ as capturing a power-sharing agreement, democratization, or any other institutional reform which checks the ruler's ability to dictate the division of spoils.

The challenger's strength has similar effects on institutional reform as its effects on prospects for conflict. If greater strength shifts the challenger's maximum threat by a larger magnitude than their average threat, then the ruler will offer (weakly) greater basement spoils to a stronger challenger. In this scenario, a stronger challenger is harder to buy off, which compels institutional reforms. On the other hand, if greater strength has the opposite effect on the maximum and average threats, then the ruler will offer (weakly) more reforms to a less strong challenger. In this case, greater strength substitutes for the need to raise the basement level of spoils. Each mechanism raises the challenger's average per-period consumption, which lowers their desire to fight in a maximally strong period. This, in turn, undermines the ruler's willingness to offer institutional reforms.

3 Application

To illustrate the substantive importance of our findings, we engage with debates about causes of democratization and authoritarian power sharing.

Adjudicating divergent theoretical implications In Acemoglu and Robinson's baseline model of authoritarian politics,[6] economic elites (the equivalent to our generic reference to a "ruler") control the regime. They interact with the masses (equivalently, "challenger"), who alternate between periods in which they are coercively strong or weak. When "strong," the masses can threaten to stage a revolution, which succeeds with probability 1 and removes elites from power forever. In every strong period, elites would like to buy off the masses by setting a high tax rate and redistributing wealth. However, elites cannot credibly commit to make concessions in any future period in which the masses are "weak" in the sense that a revolutionary attempt succeeds with probability 0. If strong periods arise rarely, then in such periods the masses stage a revolution to establish a new regime—given their unfavorable shadow of the future. Consequently, costly fighting occurs in equilibrium because of the confluence of two factors: the distribution of power fluctuates over time, and elites cannot commit to compensate the challenger in weak periods.

Acemoglu and Robinson then extend the framework to explain institutional reform. If the commitment problem is binding, elites can extend the franchise. This enables the masses to set the tax rate in all future periods, and prevents the catastrophic destruction unleashed by a revolution.[7]

In this model, a stronger challenger is one that can mobilize more frequently. Thus, strength affects the average but not the maximum threat. As we highlighted in our analysis of the general binary distribution, this implies that weaker challengers have a more credible threat to revolt.[8] This, in turn, compels the ruler to offer institutional concessions.

Ansell and Samuels (2014, 70-71) challenge a core assumption underlying these results. They contend that the material resources of a group should influence their probability of winning. Industrialization should create a stronger capitalist class that is better-positioned to challenge landed

[6]See Acemoglu and Robinson (2006), Chapter 5.

[7]See Acemoglu and Robinson (2006), Chapter 6. They also introduce the strategic option for the elites to repress the masses, which lies outside the scope of our discussion here.

[8]Fearon's (2004) model of civil wars yields a fundamentally similar implication because the frequency of strong periods is uncorrelated with prospects for winning a war, as in Acemoglu and Robinson. He phrases strength in terms of the government, and thus a "strong challenger" in his model means the government has a high chance of being weak in a given period.

elites who monopolize power. Rather than fix the maximum threat at 1, they parameterize the challenger's probability of winning in a similar fashion to our term $p^{\max}(s)$. However, their model is a one-shot game and hence threats does not fluctuate over time. They conclude that stronger challengers have better bargaining leverage and induce institutional reform, but this is a special case in which strength affects the maximum threat and its effect on the average threat is perfectly autocorrelated.

A parallel, although previously unrecognized, debate exists about motives for authoritarian power sharing. Dower et al. (2018) extend the Acemoglu and Robinson framework to incorporate the case of partial institutional reform within an authoritarian regime, as opposed to the all-or-nothing choice of full democratization. Once again, strength affects the average but not the maximum threat.

By contrast, in Meng (2019), any challenger will grow weaker over time as the dictator consolidates power between periods 1 and 2. Consequently, challengers that begin strong (or, in her phrasing, dictators who begin their tenure in a weak position) anticipate a larger adverse shift in the future distribution of power. This makes them more prone to stage a coup if the ruler does not share power with them at the outset, which induces him to do so. Here, strength affects the maximum threat more than the average threat.

In Paine (2021), the relationship between challenger strength and prospects for fighting (and power-sharing deals) are inverted U-shaped. Very weak challengers have a very low chance of prevailing (low maximum threat), and very strong challengers mobilize very frequently (high average threat). Only when challengers are endowed with intermediate strength will a ruler offer to share power because the maximum threat is large relative to the average threat.

In sum, we can recover implications from both sides of these debates as special cases of our more general model. These models produce divergent comparative statics on challenger strength because they yield varying relationships between the maximum and average threat. Understanding that these are the key theoretical quantities in these models should help to advance future theoretical work.

Implications for empirical research designs Our analysis also helps clarify impediments to empirical research that tests the relationship between challenger strength and either conflict or institutional reform. Despite advances from proposing innovative measures and research designs to assess this relationship, we highlight that theoretical implications about direction of the effect challenger strength are ambiguous.[9] They depend on the value of conditional factors that we encourage researchers to address in future studies.

Exemplifying this concern, leading empirical evaluations of these models assess opposing hypotheses.[10] Dower et al. (2018) study endogenous representation for peasants in Imperial Russia. Reforms created district-level assemblies, but varied in the extent of representation for peasants. They use the frequency of historic protests in a district to proxy for the ability to protest in the future, i.e., the q parameter. Consistent with the case in which higher strength primarily means that high threats occur more frequently, they find that high levels of past unrest engendered less representation for peasants. However, deriving this hypothesis from the model requires the additional assumption that historical threat levels had little if any effect on the magnitude of the threat posed when institutional reforms were offered.

By contrast, Aidt and Franck (2015) focus on the present threat posed by the masses. Specifically, they leverage incidence in the so-called Swing Riots to measure threat perception of British MPs in their districts, and how this affected their votes on the bill that became known as the Great Reform Act of 1832. Drawing explicitly from Acemoglu and Robinson's theory, they interpret widespread protests and rioting as a credible signal to autocratic elites that the generic hurdles to mobilizing and coordinating popular support have been temporarily overcome, i.e., $p^{\max}$ is high.

[9] Further, the theoretical maximum threat may be a parameter which is fundamentally impossible to pin down exactly.

[10] Many other studies empirically assess predictions from Acemoglu and Robinson (2006) about the relationship between economic inequality and democratization. Because these theoretical implications follow directly from underlying assumptions about the effects of challenger strength, the considerations raised here apply to these empirical tests as well.

Hence, they anticipate that MPs are *more* likely to vote for reform when more riots and protests occur in their district. However, comparing this hypothesis to that in Dower et al. (2018) highlights the problem with linking the theory to empirics. Aidt and Franck assume that strong challengers pose purely transitory threats and hence their average threat, or q, is low. However, if instead riots and protests proxy for permanently strong threats, then these models anticipate that MPs in high-protest districts should be able to pacify the recalcitrant masses with temporary transfers rather than permanent reforms. In this case, they should vote against the reform act.

Similarly, Ansell and Samuels (2014) anticipate that higher levels of industrialization and a stronger capitalist class improve prospects for democratization. The problem, though, is that if the capitalist class is permanently strong, the dynamic model implies that institutional reform is unnecessary. A high average threat enables capitalists to constantly pressure landowning elites for temporary concessions. Of course, in the real world, bargaining through such non-institutional channels may be prohibitively difficult to sustain over time because of transaction costs or costs of mobilizing. However, these are exactly the elements of these models that need to be developed in future research, and measured empirically.

While our analysis highlights some fundamental impediments to empirically measuring key parameters from models of dynamic commitment problems, it also suggests theoretical and empirical paths forward. On the theoretical end, we show that a general distribution of challenger threats can be quite tractable, while also highlighting when restricting to a binary distribution entails minimal loss. Future work can build on this to answer questions like how repression, technology for mobilization, and economic factors affect the prospects of conflict and institutional reform. Future empirical work should seek to tease apart average versus maximum threats, or perhaps more realistically how the volatility of threats relates to conflict and reform.

References

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

Aidt, Toke S and Raphaël Franck. 2015. "Democratization under the Threat of Revolution: Evidence from the Great Reform Act of 1832." *Econometrica* 83(2):505–547.

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

Chassang, Sylvain and Gerard Padro-i Miquel. 2009. "Economic Shocks and Civil War." *Quarterly Journal of Political Science* 4(3):211–228.

Christensen, Darin and Michael Gibilisco. 2020. "How Budgets Shape Power Sharing in Autocracies." Working paper.

Debs, Alexandre and Nuno P. Monteiro. 2014. "Known Unknowns: Power Shifts, Uncertainty, and War." *International Organization* 68(1):1–31.

Dower, Paul Castañeda, Evgeny Finkel, Scott Gehlbach and Steven Nafziger. 2018. "Collective Action and Representation in Autocracies: Evidence from Russia's Great Reforms." *American Political Science Review* 112(1):125–147.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

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

Gibilisco, Michael. 2021. "Decentralization, Repression, and Gambling for Unity." *Journal of Politics* 83(4):1353–1368.

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

Helmke, Gretchen. 2017. *Institutions on the Edge: The Origins and Consequences of Inter-branch Crises in Latin America*. Cambridge University Press.

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

Meng, Anne. 2019. "Accessing the State: Executive Constraints and Credible Commitment in Dictatorships." *Journal of Theoretical Politics* 33(4):568–599.

Paine, Jack. 2021. "Strategic Power Sharing: Commitment, Capability, and Authoritarian Survival." *Journal of Politics, forthcoming* .

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton University Press.

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

Powell, Robert. 2012. "Persistent Fighting and Shifting Power." *American Journal of Political Science* 56(3):620–637.

Walter, Barbara F. 2009. "Bargaining Failures and Civil War." *Annual Review of Political Science* 12:243–261.

Online Appendix

A Extensions

A.1 Endogenous Institutions

We now extend the baseline model. At the outset of the game, we allow the ruler to choose a lower bound on the transfer made to the challenger in each period. This enables us to study endogenous institutional change while also relaxing the assumption that there is no lower bound on the offer that the ruler can make in each period. Allowing the ruler to choose a lower bound could capture a wide range of institutional choices, like a power-sharing agreement, expanding the franchise, or civil rights protections. We refer to the lower bound $\underline{x}$ as the "level of reform," with higher levels of reform guaranteeing more transfers for the challenger in future periods.

Formally, at the outset of the game, the ruler chooses $\underline{x} \in [-\infty, 1]$. In each period, the choice set for the transfer is $x_t \in [\underline{x}, 1]$.

Working backwards, we analyze how the game plays out for a fixed choice of $\underline{x}$, and then move back to consider the optimal choice from the ruler's perspective. To simplify, we consider the case of a binary challenger strength, though importantly we allow strength to affect both the threat posed when strong and weak in addition to the probability of being strong.[11]

Binary strength. To simplify, we analyze the general binary case that challenger strength takes on one of two values, $p_t \in \{p^{\text{min}}, p^{\text{max}}\}$, with $q = Pr(p_t = p^{\text{max}})$. Let $x(p_t, \underline{x})$ be the offer made when the current period challenger strength is p_t and the lower bound on the offer is $\underline{x}$.

In the unbounded case, write these $x(p_t, -\infty)$. By the analysis above, in any peaceful MPE these offers are:

$$x^*(p^{\text{min}}, -\infty) = \frac{1-\phi}{1-\delta}(p^{\text{min}} - \delta((1-q)p^{\text{min}} + qp^{\text{max}})) = \frac{1-\phi}{1-\delta}((1-\delta(1-q))p^{\text{min}} - \delta q p^{\text{max}})$$
$$x^*(p^{\text{max}}, -\infty) = \frac{1-\phi}{1-\delta}(p^{\text{max}} - \delta((1-q)p^{\text{min}} + qp^{\text{max}})) = \frac{1-\phi}{1-\delta}(p^{\text{max}}(1-\delta q) - \delta(1-q)p^{\text{min}})$$

If $\underline{x} \leq x^*(p^{\text{min}}, -\infty)$, then the lower bound is irrelevant, and the analysis is equivalent to the unbounded case. That is, if equation 1 is met, then there is a peaceful MPE with no reform. If not, then there will be conflict the first time $p_t = p^{\text{max}}$.

At the other extreme, if $\underline{x}$ is sufficiently high, then the challenger would accept this minimal offer (expecting to receive it again in every future period) even when posing their maximum threat. Formally, this is true when:

$$\frac{\underline{x}}{1-\delta} \geq \frac{p^{\text{max}}(1-\phi)}{1-\delta}$$

or $\underline{x} \geq p^{\text{max}}(1-\phi) > x^*(p^{\text{min}}, -\infty)$.

[11] Preliminary analysis indicates the more general case leads to similar results.

If $\underline{x}$ is in-between these extremes, then in a potential peaceful MPE the ruler will offer $\underline{x}$ when the challenger is weak and make a higher offer when the challenger is strong.

In such an MPE, the offer made when the opposition is strong must make the challenger indifferent between accepting and not, or:

$$x(p^{\max}, \underline{x}) + \frac{\delta}{1-\delta}(qx(p^{\max}, \underline{x}) + (1-q)\underline{x}) = p^{\max}\frac{1-\phi}{1-\delta}.$$

This is solved by:

$$x^*(p^{\max}, \underline{x}) = \frac{p^{\max}(1-\phi) - \delta\underline{x}(1-q)}{1-\delta(1-q)}$$

Since we retain the upper bound on an offer as 1, the requirement for a peaceful MPE with lower bound $\underline{x}$ is that $x^*(p^{\max}, \underline{x}) \leq 1$.

This offer is decreasing in $\underline{x}$, as it takes less to buy off the challenger when they expect favorable deals in the future. Rearranging, $x^*(p^{\max}, \underline{x}) \leq 1$ if and only if:

$$\underline{x} \geq 1 - \frac{1-(1-\phi)p^{\max}}{\delta(1-q)} \equiv \underline{x}^{\text{peace}} \tag{8}$$

which is less than 1, meaning it is always possible to set $\underline{x}$ such that there is a peaceful MPE.

Note that as $q \to 1$ or $\delta \to 0$ this always holds, but it might not if q is low. Further, $\underline{x}^{\text{peace}} < p^{\max}(1-\phi)$, which verifies that if the ruler does need to offer more than $\underline{x}$ when the challenger is strong.

Optimal institutional choice Now consider the optimal institutional choice.

If equation 1 is met, there is no reason to pick a $\underline{x} \geq -\infty$. Intuitively, this will either have no impact on the outcome the game or redistribution some of the surplus to the Challenger.

The interesting case is when 1 is not met, and hence there will be conflict without reform.

There is no reason to set $\underline{x}$ strictly below $\underline{x}^{\text{peace}}$, as this will just entail giving weakly more to the challenger in any period where $p_t = p^{\min}$ and still fighting in the first period where $p_t = p^{\max}$. And there is no reason to set $\underline{x}$ strictly higher than $\underline{x}^{\text{peace}}$, as this will just lead to a peaceful MPE where the challenger gets more in periods where $p_t = p^{\min}$.

So, the relevant question to ask when there is no peaceful MPE without a lower bound is whether the challenger prefers to set $\underline{x}^{\text{peace}}$ and play the peaceful MPE or to pick a lower bound which never binds, expect that conflict will happen when $p_t = p^{\max}$.

If setting the lower bound to $\underline{x}^{\text{peace}}$, the ruler will get 0 in periods where the challenger is strong (since they need to offer 1) is and $1 - \underline{x}^{\text{peace}}$ in periods where the challenger is weak. This gives expected utility

$$U_R(\underline{x}^{\text{peace}}) = \frac{(1-q)(1-\underline{x}^{\text{peace}})}{1-\delta} = \frac{1-p^{\max}(1-\phi)}{\delta(1-\delta)} \tag{9}$$

To compute the expected utility when setting a lower bound which does not bite, note that conflict will occur in period 1 with probability q, giving $\frac{(1-p^{\max})(1-\phi)}{1-\delta}$ for the remainder of the game and with probability $(1-q)$ the ruler gets to keep $1-x^*(p^{\text{min}}, -\infty)$ and the expected utility moving forward is the discount rate times the expected utility starting in period 1. That is,

$$U_R(-\infty) = q\frac{(1-p^{\text{max}})(1-\phi)}{1-\delta} + (1-q)(1-x^*(p^{\text{min}}, -\infty) + \delta U_R(-\infty))$$

Which has a unique solution.

While this expression is complex, if we want to know how the parameters of the distribution affect the propensity for power-sharing, we need to check how they affect the difference in the expected utility for picking the reform level which guarantees peace vs allowing conflict to occur:

$$D_{\text{peace}} \equiv U_R(\underline{x}^{\text{peace}}) - U_R(-\infty)$$

If increasing one parameter increases this expression, it increases the set of other parameters where reform will happen. These derivatives have the same interpretation as how these parameters affect the prospects for conflict.

Proposition 2. *The propensity for reform D_{peace} is increasing in p^{min} and q, and decreasing in p^{max}.*

Proof

$$\begin{aligned}
\frac{\partial D_{\text{peace}}}{\partial p^{\text{min}}} &= \frac{(1-\phi)(1-q)}{1-\delta} > 0 \\
\frac{\partial D_{\text{peace}}}{\partial p^{\text{max}}} &= -\frac{(1-\phi)(1-\delta q)}{\delta(1-\delta)} < 0 \\
\frac{\partial D_{\text{peace}}}{\partial q} &= \frac{(1-\phi)(p^{\text{max}}-p^{\text{min}})}{1-\delta} + \frac{\phi}{(1-\delta(1-q))^2} > 0
\end{aligned}$$

As in the case of how the distribution affects conflict, the answer depends on the precise nature of what it means for the challenger to get stronger. The result about a shift of the distribution holds as well. If increasing s increases p^{min} and p^{max} at equal rates (and does not affect q), then the relative value of reform, the overall effect of increasing strength is proportional to $\frac{\partial D_{\text{peace}}}{\partial p^{\text{min}}} + \frac{\partial D_{\text{peace}}}{\partial p^{\text{max}}} = -\frac{1-\phi}{\delta} < 0$. So, making the challenger stronger in this sense decreases the prospect for reform.

Finally, we can ask how the amount of institutional reform change as the challenger gets stronger. If the ruler needs no institutional reform or prefers to just let the challenger fight, then (local) changes in the challenger strength do not affect the equilibrium choice.

Within the range where the ruler selects $\underline{x}^{\text{peace}}$, we can see how the challenger become stronger affects this choice by differentiating $\underline{x}^{\text{peace}}$, which is clearly increasing in p^{max} and decreasing in q.

Proposition 3. *With binary threat levels, if $\underline{x}_{peace} \geq x^*(p^{min}, -\infty)$ and $D_{peace} \geq 0$, then the ruler picks a non-trivial lower bound on the offer made in each period, and this ensures a peaceful MPE. The level of reform is increasing in the maximum strength of the challenger and decreasing in how frequently they are strong.*

As with the case with no lower bound on offers, the question of whether stronger challengers induce more institutional reform depends critically on how one conceives of strength. If getting stronger corresponds to a "rightward shift" in the the opposition probability of winning conflict, then the ruler is more likely to institute reform and picks a higher level of reform when they do. If getting stronger corresponds to posing a relatively high threat more often, then the ruler is less apt to choose reform, and chooses a lower level of reform is so.

A.2 Path-Dependent States

While the distribution we use for the baseline model is quite general, one strong assumption is that threat levels are independent across periods. A simple way to capture "path dependent" states is to assume that with probability $q \in (0,1)$, the challenger threat in period t is equal to p_{t-1}, and is drawn from $F(p; s)$ otherwise. So, higher values of q correspond to "more sticky" threat levels. If so, the continuation value depends on the current value of p_t. Let $V^C(p_t)$ be the continuation value for entering the next period with threat p_t. The indifference condition can be written:

$$x_t(p_t) = p_t \cdot \frac{1-\phi}{1-\delta} - \delta \cdot (qV^C(p_t) + (1-q)V_n^C) \tag{10}$$

Let $V_n^C = \mathbb{E}[V^C(p_t)]$ is the continuation value if the threat changes. We can write the continuation value with threat p_t as:

$$V^C(p_t) = x_t(p_t) + \delta \cdot (qV^C(p_t) + (1-q)V_n^C)$$

And so:

$$V^C(p_t) = \frac{x_t(p_t) + \delta(1-q)V_n^C}{1-\delta q}$$

Plugging this back into equation (10) gives:

$$x_t(p_t) = p_t \cdot \frac{1-\phi}{1-\delta} - \delta \cdot (q\frac{x_t(p_t) + \delta(1-q)V_n^C}{1-\delta q} + (1-q)V_n^C) \tag{11}$$

$$x_t(p_t) = \frac{(1-\phi)(1-\delta q)}{1-\delta}p_t - \delta(1-q)V_n^C \tag{12}$$

which is linear in p_t. As a result we can solve for V_n^C by:

$$V_n^C = \mathbb{E}[x_t(p_t)] + \delta V_n^C$$

$$V_n^C = \frac{(1-\phi)(1-\delta q)}{1-\delta}\bar{p} - \delta(1-q)V_n^C + \delta V_n^C$$

$$V_n^C = \frac{(1-\phi)}{(1-\delta)}\bar{p}$$

I.e., the same as in the $q = 0$ (no path-dependence) case.

Plugging this back into equation (12) gives an explicit formula for the offer in each period:

$$x_t(p_t) = \frac{(1-\phi)(1-\delta q)}{1-\delta} p_t - \delta(1-q)\frac{(1-\phi)}{(1-\delta)}\bar{p} \tag{13}$$

$$= \frac{1-\phi}{1-\delta}((1-\delta q)p_t - \delta(1-q)\bar{p}) \tag{14}$$

As $q \to 0$ this nests the no path dependence case. As $q \to 1$ the offer approaches $(1-\phi)p_t$, i.e., what the offer would be in the static version of the model. Note this is always less than 1, and so as q increases it is easier to sustain peace. More generally, peace is possible when:

$$\frac{1-\phi}{1-\delta}((1-\delta q)p^{\max}(s) - \delta(1-q)\bar{p}(s)) \equiv \tau(s,q) =\leq 1 \tag{15}$$

This is harder to sustain when the opposition is stronger if $\tau(s,q)$ is increasing in s, or:

$$(1-\delta q)\frac{\partial p^{\max}}{\partial s} - \delta(1-q)\frac{\partial \bar{p}}{\partial s} > 0$$

If q is sufficiently large, the second term approaches zero (while the first does not as long as $\delta < 1$), so this is always true as long as $\frac{\partial p^{\max}}{\partial s} > 0$.

In other words, as the threat posed by the challenger is sufficiently stable over time, then stronger challengers are always hard to buy off (provided this has some impact on the maximum threat).

A.3 When Does Conflict Happen?

A natural MPE with conflict is one where there exists a p^* such that when $p_t \leq p^*$ an offer is made which is accepted, and when $p_t > p^*$ conflict occurs in period t. In such an MPE, the continuation value for accepting an offer must equal the conflict payoff $p_t\frac{1-\phi}{1-\delta}$. Further, in an period where conflict occurs, this must be the continuation value as well. As a result, the optimal offer in periods with peace must be the same as in the peaceful MPE. As a result, conflict will happen in the first period such that:

$$p_t\frac{1-\phi}{1-\delta} \geq 1 + \delta\frac{1-\phi}{1-\delta}\bar{p}$$

or

$$(p_t - \delta\bar{p})\frac{1-\phi}{1-\delta} \geq 1$$

A straightforward implication is that conflict happens once the challenger achieves a period threat sufficiently higher than their average threat.