Original Article

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

Comparative Political Studies
2023, Vol. 0(0) 1–20

DOI: 10.1177/00104140231169031
journals.sagepub.com/home/cps
Sage

Andrew T. Little[1] and **Jack Paine**[2]

Abstract

Prominent theories propose that commitment problems drive phenomena such as war and democratization. However, existing work disagrees about a basic question: how does a challenger's coercive strength affect prospects for conflict and/or institutional reform? We establish that the relationship depends on how challenger strength affects the average and maximum probability of winning a conflict in a given period ("threat"). We analyze a formal model with a general distribution of threats, and conceptualize challenger strength as affecting this distribution. If the maximum threat is fixed and stronger challengers pose a higher average threat, then weak challengers will rebel (absent reform) during the rare periods they pose a high threat. However, if stronger challengers pose a greater maximum threat, then they are harder to buy off. Applying these insights advances theoretical and empirical debates about democratization.

Keywords

conflict processes, democratization and regime change, formal theory, political economy, political regimes

[1]University of California Berkeley, Berkeley, CA, USA
[2]Emory University, Atlanta, GA, USA

Corresponding Author:
Andrew T. Little, University of California Berkeley, 210 Social Sciences Building, Berkeley, CA 94720, USA.
Email: andrew.little@berkeley.edu

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? Much existing research points to *dynamic commitment problems* to explain these varied phenomena.[1] The core premise is that an actor who controls a flow of rents, such as a government, cannot commit to promises about how it will distribute spoils in the future. Limited commitment ability can prompt conflict and/or institutional reform for the following reason. A challenger—for example, a domestic opposition group or foreign adversary—can compel the government for concessions at any time in which it poses a coercive threat. However, in most foreseeable real-world scenarios, the threat that a challenger poses to overthrow the government fluctuates over time. Sometimes the domestic masses have favorable opportunities to mobilize anti-government demonstrations, and sometimes they do not. Sometimes foreign states enjoy economic booms that bolster their military strength, and sometimes they do not. Consequently, a challenger that poses a stark threat today may not tomorrow. A temporarily strong challenger will seize its window of opportunity to fight because the government has limited ability to commit to future concessions. Alternatively, this pressure might prompt the government to reform institutions (e.g., democratization, sharing power) and thereby bolster its commitment to share future spoils with the challenger.

This style of argument is pervasive because the core intuition is straightforward, compelling, and broadly applicable. However, existing theories that incorporate this mechanism propose divergent answers to a basic, substantively important question: do coercively stronger challengers make conflict and/or institutional reform more or less likely? How does a bigger non-elite class, a better-organized civil society, or a more advanced neighboring state affect these outcomes? Our answer is simple: the effect of challenger strength depends on how it affects the challenger's average and maximum threats, with "threat" measured by the probability that the challenger would win a conflict in a given period. A challenger is easy to buy off when its average threat is high but difficult to buy off when its maximum threat is high. Thus, we explain how different ways of conceptualizing the challenger's coercive strength yield divergent answers. Understanding this mechanism is crucial for resolving seeming contradictions in existing theories of democratization and power sharing, while also carrying important implications for empirical testing.

Our theoretical innovation is to examine a more general distribution of threats than in existing models. To our knowledge, we are the first to consider any discrete or continuous distribution of threat levels or, more generally, any variable which affects payoffs and fluctuates over time. Our findings establish that seemingly technical model assumptions about this distribution in fact matter greatly for substantively important questions.

Much existing work makes a simplifying assumption about the distribution of threats: the challenger fluctuates between two threat levels, minimal and maximal. Minimal threats usually represent periods in which the challenger would lose for sure or in which conflict is prohibitively destructive. By contrast, maximum threats arise when the challenger would win a (moderately costly) conflict with certainty, which creates a dire threat. In this setup with binary threat levels and fixed values for each, a natural way to capture coercive strength is the probability that the challenger will pose its maximum threat—with a stronger challenger posing the maximal threat more frequently. Counterintuitively, in this setup, *a weaker challenger is more prone to fight*. Its rare windows of opportunity are too tempting to pass up and to forgo revolting, given poor prospects to gain concessions in the future if the status quo regime remains intact. Thus, weaker challengers lead to either more conflict or, if the ruler is able, more institutional reform.

We indeed recover this scenario as a special case in our model. The underlying commitment problem is less pressing for a challenger who poses a higher *average* threat. This source of strength enables the challenger to compel more concessions from the ruler over time, which lessens its motives to fight.

However, as we demonstrate, this is not a general result about the consequences of stronger challengers. The seemingly innocuous assumption to construct the distribution of threats such that only the average threat is affected by strength creates an unrecognized problem: it holds fixed a crucial margin, the *maximum* threat is affected by strength. If stronger challengers correspond with a higher maximum threat, then *a stronger challenger is more prone to fight* upon realizing the maximum threat. A higher contemporaneous threat raises the opportunity cost to forgoing conflict. In this scenario, a stronger challenger is also better positioned to force institutional reform.

Generally, coercively strong challengers should pose greater threats both on average and when maximally strong. This means that the overall effect of greater coercive strength is theoretically ambiguous. A stronger challenger is more likely to fight (or gain institutional concessions) whenever a shift in the strength parameter raises the maximum threat by at least as much as the average. In fact, the direction of the result is unchanged even if the average threat increases at a somewhat higher rate because the challenger does not enjoy the benefits of a higher average threat until future periods. For example, a uniform upward shift in the distribution of threats makes the challenger harder to buy off peacefully and makes institutional reform more likely. Even in a simple binary threats model, we can recover our core result as long as we allow both the maximum and average threats to vary. Overall, we cannot understand the consequences of challenger strength without taking into account how coercive strength affects both maximum and average threats.

After analyzing the model, we highlight new insights for debates about democratization and authoritarian power sharing. In models such as Acemoglu and Robinson (2006) and Castañeda Dower et al. (2018), coercively weak challengers trigger institutional reform. A low average threat makes the shadow of the future unfavorable. This feature, combined with an assumption that the maximum threat is high, bolsters the challenger's bargaining leverage in a rare maximum-threat period. However, other seemingly similar models yield the opposite implication about challenger strength (Ansell & Samuels, 2014; Meng, 2019; Paine, 2022). By disaggregating maximum and average threats, our model explains the conditions under which we recover each implication.

Our findings also offer guidance for empirical research designs that test these models. Recent studies propose innovative ways to measure key parameters but do not consider the countervailing effects of higher maximum and average threats. Fortunately, our analysis highlights paths forward. For tests that compare challenger strength across cases, it is important to scrutinize the perceived permanence of the group who organizes against the government, as more durable organizations yield higher average threats. This may help to account for the divergent empirical relationships between mobilization and democratic reform found in Aidt and Franck (2015) and Castañeda Dower et al. (2018). Tests that compare challenger strength over time within cases are less subject to our main theoretical insight because any model in this class predicts that conflict (or institutional reform) is more likely in periods in which the challenger's *current* threat is high relative to the *average* threat it poses. This applies to research designs that study, for example, rainfall-induced economic shocks (Aidt & Leon, 2016; Brückner & Ciccone, 2011). Yet more discussion is needed, for any particular substantive setting, to establish why the maximum threat would be high relative to the average threat—which is necessary for high-threat periods to yield institutional reform. We conclude by discussing the importance of formalization and directions for future empirical and theoretical research.

Formal Model

Setup

A ruler and a challenger bargain over spoils in periods $t = 1, 2, \ldots$ 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 $x_t \leq 1$. This incorporates the standard assumption that the ruler cannot transfer more than the entire contemporaneous budget in any period, and hence cannot borrow across periods. If the challenger accepts an offer 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 a fraction $\mu \in (0, 1)$ of total spoils, with the winner consuming all the remaining spoils.

The challenger's probability of winning a conflict varies by period. The parameter is p_t, which depends on an independently and identically distributed choice by Nature revealed to both players at the outset of each period. 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. Formally, assume the distribution function of p_t is $F(p; s)$, where s is a parameter that captures the challenger's latent coercive capabilities, or strength. The distribution has mean $\overline{p}(s) \equiv \mathbb{E}[p; s]$ and support on $[p^{\min}(s), p^{\max}(s)]$, for $0 \leq p^{\min} < p^{\max} \leq 1$. To capture the general notion that stronger challengers tend to pose a higher threat, we assume that $\overline{p}(s)$, $p^{\min}(s)$, and $p^{\max}(s)$ each weakly increase in s.[2] To streamline the exposition, we suppress s when doing so does not cause confusion.

Following the model analysis, we discuss simplifying assumptions and summarize extensions presented in full in the appendix: imposing a lower bound on the bargaining offer, allowing for a path-dependent distribution of threats, modeling fluctuations in the cost of conflict rather than the probability of winning to more closely parallel the setup in Acemoglu and Robinson (2006), and allowing the ruler to engage in institutional reform.

Analysis: How the Distribution of Threats Affects Conflict

We start by asking when a Markov Perfect Equilibrium (MPE) exists in which conflict occurs with probability 0 in every period 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

$$\underbrace{x_t + \delta V^C}_{\text{Accept}} \geq \underbrace{\frac{p_t(1-\mu)}{1-\delta}}_{\text{Fight}} \tag{1}$$

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, and thus the optimal transfer in every period satisfies:

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

The next step is to solve for the continuation value V^C. In a peaceful MPE in which the ruler uses the offer function from Equation (2) in every period, we can write the continuation value as equal to the average transfer divided by $1-\delta$. An analytically convenient aspect of the optimal offer is that it is linear in the current-period threat p_t, and hence the average value $\overline{p}$ is the only aspect of the distribution that affects the continuation value. As demonstrated in Appendix A.1, this property holds in any equilibrium with conflict as well.

Formally, we can write the continuation value as

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

$$\Rightarrow V^C = \frac{\overline{p}(1-\mu)}{1-\delta}. \tag{3}$$

This term is simply the challenger's average probability of winning multiplied by the lifetime expected consumption from winning a conflict. Combining Equations (2) and (3) enables us to explicitly solve for the equilibrium per-period offer:

$$x^*(p_t) = \frac{(p_t - \delta\overline{p})(1-\mu)}{1-\delta}. \tag{4}$$

Two conditions are needed for a peaceful MPE to exist. First, the offer in Equation (4) never exceeds the budget constraint of 1. Second, the ruler prefers to make this offer in each period rather than to endure conflict. The optimal offer $x^*(p_t)$ strictly increases in p_t, as we can see in Equation (4), and therefore the maximum value is $x^*(p^{\max})$. Consequently, the first condition requires $x^*(p^{\max}) \leq 1$. The second condition always holds because, in any period in which a peaceful bargain is struck, the ruler extracts all surplus saved from not fighting. The existence of bargaining surplus also implies that, when the first condition holds, the peaceful MPE is unique; in any strategy profile in which conflict occurs, the ruler can profitably deviate to make the challenger an acceptable offer. We summarize this result in Proposition 1. If instead $x^*(p^{\max}) > 1$, then the equilibrium is qualitatively similar to the strategy profile just described despite featuring conflict along the equilibrium path. Specifically, there is a unique MPE in which the ruler buys off the challenger whenever possible but conflict occurs in the first period in which the contemporaneous threat p_t exceeds a critical threshold.

Proposition 1. (Equilibrium). *If* $(p^{max} - \delta\overline{p})(1 - \mu)/(1 - \delta) \leq 1$, *then the following strategy profile is the unique MPE: in every period, the ruler proposes* $x_t = x^*(p_t)$ *according to Equation* (4) *and the challenger accepts if and only if* $x_t \geq x^*(p_t)$. *Along the equilibrium path, the challenger accepts the offer in every period. If* $(p^{max} - \delta\overline{p})(1 - \mu)/(1 - \delta) > 1$, *then any MPE has conflict along the equilibrium path of play (see Proposition* A.1 *for details and proof).*

The inequality that forms the crucial scope condition for Proposition 1, $(p^{max} - \delta\overline{p})(1 - \mu)/(1 - \delta) \leq 1$, enables us to analyze how changing challenger strength affects the possibility of a peaceful equilibrium by taking comparative statics on s. If we move the threat parameters to one side of the inequality and write them explicitly as a function of s, this condition becomes:

$$\frac{1 - \delta}{1 - \mu} \geq p^{\max}(s) - \delta\overline{p}(s) \equiv \tau(s) \tag{5}$$

We say that prospects for conflict are increasing in s if $\tau(s)$ increases in s because this expands the range of parameters in which conflict occurs along the equilibrium path. Similarly, we say prospects for conflict are decreasing in s if $\tau(s)$ decreases in s. The overall effect of increasing the challenger's strength on the prospects for peace and conflict can be summarized as follows:

Proposition 2. *Increasing challenger strength raises prospects for conflict, that is,* $\tau(s)$ *strictly increases in s, if and only if:*

$$\underbrace{\frac{\partial p^{\max}(s)}{\partial s}}_{\uparrow\ maxthreat} > \delta \underbrace{\frac{\partial \overline{p}(s)}{\partial s}}_{\uparrow\ averagethreat} \tag{6}$$

The proof follows directly from the following derivative:

$$\frac{\partial \tau(s)}{\partial s} = \frac{\partial p^{\max}(s)}{\partial s} - \delta \frac{\partial \overline{p}(s)}{\partial s}$$

This result expounds our main point about the need to compare the maximum and average threats, which exert countervailing effects on prospects for conflict. On the one hand, higher s increases prospects for conflict by raising the challenger's maximum probability of winning, $p^{\max}$. This effect raises the challenger's opportunity cost to forgoing the conflict option in a maximum-threat period. When we raise the probability of winning in a maximum-threat period while the average threat (and hence future expected transfers) is fixed,[3] we increase the discrepancy between the challenger's threat in the current period and its threat in future periods. This creates the temptation to fight now to "lock in" this temporary advantage. Consequently, the inequality from Proposition 1 needed for a peaceful path holds for a smaller range of parameter values.

On the other hand, higher s diminishes prospects for conflict by raising $\overline{p}$. When the challenger contemplates fighting in a maximum-threat period, it considers the magnitude of the adverse shift in the future distribution of power. A higher average threat lowers the opportunity cost to forgoing fighting. The challenger expects more favorable draws of p_t in the future along a peaceful path, which diminishes its incentives to fight now. This makes the inequality from Proposition 1 easier to meet. Notice that $\frac{\partial \overline{p}(s)}{\partial s}$ is multiplied by δ whereas $\frac{\partial p^{\max}(s)}{\partial s}$ is not. The reason is that the challenger reaps the benefits of higher $p^{\max}$ in the present period whereas the consumption gains from higher $\overline{p}$ begin to accrue starting only in the next period.

General Binary Distribution

To connect this result more directly to past work, suppose the per-period threat takes either of two values. We write these as $p_t \in \{p^{\min}, p^{\max}\}$ with $0 \leq p^{\min} < p^{\max} \leq 1$ and $q = Pr(p_t = p^{\max})$. In this case, the average threat is $\overline{p} = (1-q)p^{\min} + qp^{\max}$. Substituting this term into Equation (6) and taking comparative statics yields:

$$\frac{\partial \tau(s)}{\partial s} = (1-\delta q) \underbrace{\frac{\partial p^{\max}}{\partial s}}_{\uparrow \text{ max threat}} - \delta(1-q) \underbrace{\frac{\partial p^{\min}}{\partial s}}_{\uparrow \text{ min threat}} - \delta(p^{\max} - p^{\min}) \underbrace{\frac{\partial q}{\partial s}}_{\uparrow \text{ max-threat periods}} \tag{7}$$

In Figure 1, we graphically summarize some key comparative statics predictions. It is a region plot with $p^{\max}$ on the x-axis and q on the y-axis; all other parameters are fixed at values stated in the accompanying note. The white region corresponds with parameter values in which the equilibrium path of play is peaceful (that is, the inequality in Proposition 1 holds), whereas conflict occurs in the dark region.

Equation (7) and Figure 1 clarify the intuition for the result from Acemoglu and Robinson (2006) and other models in which a stronger challenger is *easier to buy off*. Holding fixed the values of the minimum and maximum threats means the first two terms in Equation (7) are 0. Hence, higher s improves the shadow of the future for the challenger along a peaceful path by raising q. This effect bolsters the challenger's average threat without altering the opportunity cost of fighting in the maximum-threat state, which is dictated by $p^{\max}$. This

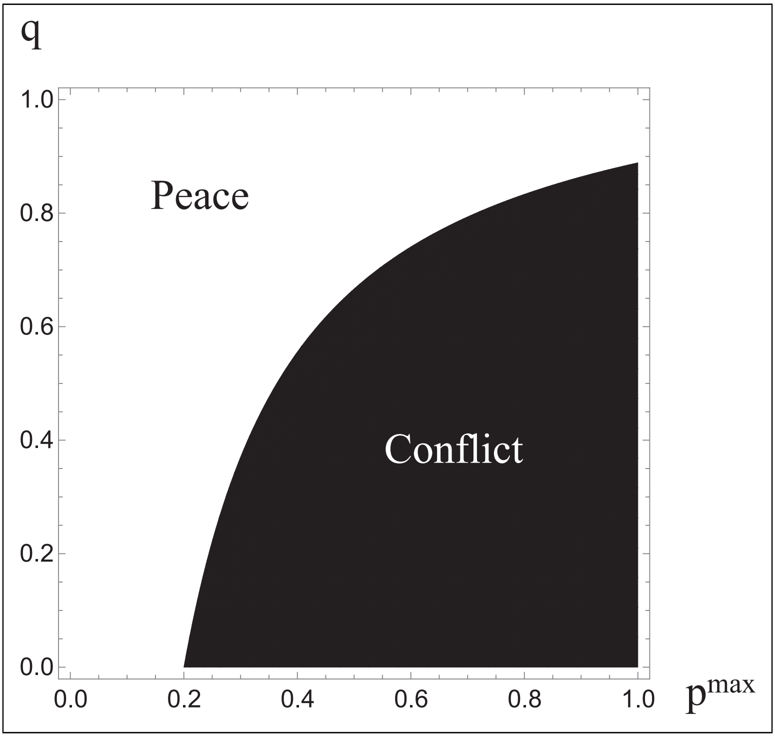

Figure 1. Peace and Conflict in the Binary Threats Model. *Parameter values*: $\delta = .9$, $\mu = .5$, $p^{\min} = 0$.

corresponds with an upward shift in Figure 1, which can move parameter values from conflict to peace.

Our analysis also suggests a sense in which we can generalize this finding. For any distribution shift such that the upper bound is 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, even with a binary distribution of threats, raising the challenger's strength can instead produce the opposite effect. The simplest case is one in which greater coercive strength raises $p^{\max}$ while q remains fixed. This corresponds with a rightward shift in Figure 1, which can move parameter values from peace to conflict.

Another natural case is when increasing s shifts the distribution of threats F uniformly to the right. This corresponds to raising the challenger's probability of winning by a fixed amount in each period. Hence, the minimum and maximum threats each increase by some constant $d > 0$ but the per-period probability of each threat realization, q, is unchanged. In this case, facing a stronger challenger makes *peace harder to sustain* for the reason discussed after Proposition 2: the challenger immediately reaps the benefits of a higher maximum threat, but does not gain the benefits of a higher minimum (and therefore average) threat until future periods. Formally, we can see this by substituting this case into Equation (7), which yields $\partial\tau(s)/\partial s = (1 - \delta)d > 0$.

These examples highlight a useful fact for future theorizing: a binary distribution in of itself does not discernibly 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 conceptualizes challenger strength and structures the parameters in the distribution of threats determines the direction of the comparative statics prediction. A binary distribution of threats contains three key parameters, and different changes carry divergent implications for the prospect of peace.

Discussion of Assumptions and Extensions

In the baseline model, we do not impose a lower bound on the bargaining offers. This makes it possible for the ruler to offer $x_t < 0$ and hence to demand a net transfer from the challenger. Substantively, we can interpret the "spoils" under consideration as only part of the resources available in the society, which enables the ruler in principle to extract more whenever the challenger poses a weak threat. Mathematically, the case without a lower bound is analytically simpler because the ruler can hold the challenger down to its reservation value in every period. This makes the optimal offer linear in the challenger's strength, and hence the average threat is the only part of the distribution that matters for the continuation value. With a lower bound, the continuation value depends on other aspects of the distribution. However, the core insights are sometimes identical and otherwise qualitatively similar when we assume that offers must be above some lower bound $\underline{x}$, which we demonstrate in Appendix A.2. To preview the intuition, suppose $\underline{x} = 0$, and therefore net transfers must favor the challenger. From Equation (4), it is immediately apparent that all interior-optimal offers strictly exceed zero if $p^{\min} \geq \delta\overline{p}$. Thus, the lower bound never binds if there is a small range of feasible values of p and the actors are not too patient. If instead $p^{\min} < \delta\overline{p}$, then the lower bound of zero is binding. This case adds additional terms but does not qualitatively alter the main insight that we need to compare the maximum and average threats.

Another simplifying assumption is that threats are drawn independently and identically across periods. In Appendix A.3, we relax this assumption and demonstrate that our key findings hold when we allow for a specific type of path dependence. With positive probability, Nature does not change the state of the world in the next period; but with complementary probability, Nature draws from the same underlying distribution of threats as in the baseline model. Although our model does not nest all forms of path dependence or deterministic shifts (Gibilisco, 2021; Krainin, 2017), this extension demonstrates that our core findings do not require iid shocks.

In our model, the probability of winning p_t is the random variable. Some related models instead fix the probability of winning and allow the cost of conflict to vary across periods, for example, Acemoglu and Robinson (2006). In their setup, the challenger is "stronger" when this cost, which we (and they) denote as μ, is more frequently low. In our view, allowing the probability of winning to vary is a more natural way to capture the notion of challenger strength. However, all that matters for our results is how the challenger's expected value to fighting changes over time—and the calculus here is identical regardless of whether the probability of winning or the costliness of fighting fluctuates over time. In Appendix A.4, we present a version of the model where the cost of conflict μ_t fluctuates over time, which yields qualitatively identical results.

Prospects for Institutional Reform

Our final extension is more substantively oriented and addresses endogenous institutional reform. We have shown that the challenger's strength parameter, s, has ambiguous consequences for conflict. The intuition is identical when we allow the ruler to strategically reform institutions. In Appendix A.5, we assume that the ruler in each period can choose to permanently increase the basement level of spoils the challenger consumes in all periods (that is, to choose the value of $\underline{x}$, introduced in Appendix A.2). We interpret a higher basement level of spoils as capturing a power-sharing agreement, democratization, or any other institutional reform that constrains the ruler's ability to dictate the division of spoils.

The parameter region in which institutional reform occurs in this extension is identical to that in which conflict occurs in the baseline game. Along the equilibrium path, in the first maximum-threat period, the ruler offers a sufficient level of institutional reforms to enable buying off the challenger then and in all future periods. The continuous choice of institutional reform enables the ruler to hold the challenger down to indifference, and the ruler would immediately incur the costs of conflict if she did not reform institutions. Consequently, the ruler never lets conflict occur along the equilibrium path.

The equivalence of the institutional reform region with the conflict region implies that all comparative statics from the baseline model carry over to explain institutional reform: a greater average threat diminishes incentives for institutional reform, and a greater maximum threat increases incentives for institutional reform. Higher $p^{\max}$ also increases the extent of institutional reform (that is, raises the optimal choice of $\underline{x}$), conditional on any reform occurring. A challenger with high $p^{\max}$ requires greater assurances to compensate for the higher opportunity cost of not fighting in a maximum-threat period.

Application: Adjudicating Divergent Theoretical Implications

To illustrate the substantive importance of our findings, we engage with debates about causes of democratization and authoritarian power sharing. In this section we adjudicate divergent theoretical implications, and in the next we discuss implications for empirical research designs.

In Acemoglu and Robinson's baseline model of authoritarian politics,[4] economic elites (the equivalent to our generic reference to a "ruler") control the political regime. Elites interact with the masses (equivalently, "challenger"), whose threats alternate over time according to a binary distribution with $p^{\min} = 0$ and $p^{\max} = 1$.[5] Thus, in maximum-threat periods, the masses can credibly threaten to stage a revolution, which succeeds with probability 1 and removes elites from power forever. In every maximum-threat period, elites would ideally like to buy off the masses with a temporary concession: setting a high tax rate and redistributing wealth in that period only. However, elites cannot credibly commit to make similarly generous concessions in any future periods in which the masses pose the minimum threat, in the sense that a revolutionary attempt succeeds with probability 0. If maximum-threat periods arise rarely, then in any such period, temporary concessions are insufficient to pacify the masses because their shadow of the future is unfavorable. Frequent minimum-threat periods imply low future consumption, which prompts the masses to revolt when temporarily strong.

Acemoglu and Robinson then extend their framework to explain endogenous institutional reform.[6] If revolution would otherwise occur along the equilibrium path, then elites will extend the franchise. The drawback for elites is that democratization enables the masses to set the tax rate in all future periods. However, elites benefit by preventing the catastrophic destruction that a revolution would unleash. In our model, increasing the lower bound offer $\underline{x}$ corresponds with franchise expansion.

In the Acemoglu and Robinson model, a stronger challenger is synonymous with more frequent maximum-threat periods. Thus, strength affects the average but not the maximum threat, which is fixed at $p^{\max} = 1$. As we highlighted in our analysis of the general binary distribution, this implies that weaker challengers have a more credible threat to revolt. This, in turn, compels the ruler to offer institutional concessions to weak but not strong challengers.

Ansell and Samuels (2014) confront a core assumption underlying these results (see especially pp. 70–71). They contend that the material resources of a group should influence its probability of winning. In 19th century European countries, industrialization created a stronger capitalist class that was better positioned to challenge landed elites who monopolized power. Rather than fix $p^{\max} = 1$, they parameterize the challenger's probability of winning in a similar fashion to our term $p^{\max}(s)$. A stronger challenger therefore wins with higher probability, which enables them to compel institutional reform—which yields the opposite result as in Acemoglu and Robinson. However, Ansell and Samuels' model is a one-shot game, which means that threats do not fluctuate over time. As we demonstrate with our more general model, this is a special case in which challenger 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. Castañeda Dower et al. (2018) extend the Acemoglu and Robinson framework to incorporate the possibility of partial institutional reform within an authoritarian regime, as opposed to the all-or-nothing choice of full democratization. Once again, challenger strength affects the average but not the maximum threat, and thus weaker challengers compel power sharing.

By contrast, in Meng's (2019) two-period game, the challenger grows weaker over time as the dictator consolidates power between periods 1 and 2. Consequently, a challenger that initially poses a high maximum threat anticipates a larger adverse shift in the future distribution of power. This makes a stronger challenger more prone to stage a coup if the ruler does not share power at the outset. This mechanism induces power sharing with strong but not weak challengers. Here, greater challenger strength affects the maximum threat more than the average threat.

In Paine (2022), the relationship between challenger strength and prospects for both fighting and power-sharing deals are inverted U-shaped. Using a binary threats model, he assumes that the maximum threat is positively correlated with the frequency with which the maximum threat is realized. Consequently, very weak challengers have a low chance of ever prevailing (low maximum threat) and very strong challengers frequently enjoy maximum-threat periods (high average threat). Only intermediate-strong challengers have a credible threat to fight, which induces the ruler to share power. In this range, the maximum threat is large relative to the average threat.

In sum, we can reconcile seemingly incompatible implications about democratization and authoritarian power sharing as special cases of our more general model. Existing models yield divergent comparative statics for challenger strength because of varying, and usually undiscussed, assumptions that affect the relationship between the maximum and average threat. Understanding that these are the key theoretical quantities in these models should help to advance future theoretizing. Seemingly technical model assumptions carry important substantive implications.

Application: Implications for Empirical Tests

Empirical examinations of democratization models with commitment problems typically address either of two questions. (a) Across cases, where should we expect democratic reform? (b) Within cases over time, when should we expect democratic reform? These models offer ambiguous implications for the first question for the reasons we have discussed. Challenger strength differs across cases, but without further specification, we do not know whether comparatively stronger challengers are more or less likely to fight or to gain institutional concessions (see Proposition 2). By contrast, these models offer an unambiguous prediction for the second question: a challenger is more likely to fight (or gain institutional reforms) in time periods when it poses a higher-than-average threat. This follows from the straightforward point, discussed in the analysis, that the opportunity cost of forgoing fighting strictly increases in p_t. Consequently, the optimal offer $x^*(p_t)$ strictly increases in p_t (see Equation (4)), which makes the challenger strictly harder to buy off. We discuss existing empirical tests from this perspective.

Cross-Case Comparisons

Drawing conclusions from comparisons of challenger strength across cases is inherently difficult because the theoretical implications are ambiguous. Exemplifying this point, two leading empirical evaluations of the Acemoglu and Robinson (2006) model assess opposing hypotheses about challenger strength. In both cases, further discussion of key scope conditions is needed to more closely tie the empirical test to theoretical implications.[7]

Castañeda Dower et al. (2018) study endogenous representation for peasants in Imperial Russia. Reforms in 1864 created district-level assemblies, *zemstva*, which varied in the fraction of seats reserved for peasants. The authors use the frequency of protests in each district over the preceding decade to proxy for the average threat that peasants in the district would pose in the future, which we (and they) formalize as the q parameter. They find that high levels of past unrest engendered less representation for peasants, hence demonstrating that higher average threats led to less institutional reform. Alone among existing empirical tests, they explicitly engage with the idea from the Acemoglu and Robinson (2006) model that the masses and institutional reformers consider not only the present threat, but also the expectation about future threats.

However, our analysis of a more general distribution of threats identifies a key scope condition for the theoretical implication: coercively stronger challengers (as proxied by the frequency of unrest in the 1850s) must have posed a higher average than maximum threat. Otherwise, if such districts posed (comparatively) very large threats in the mid-1860s when the reforms were implemented, we would instead expect them to gain greater reforms. The authors' careful discussion of the historical context does not address this specific point.

By contrast, Aidt and Franck (2015) focus on the present threat posed by the masses. They leverage local variation in the intensity of Swing Riots to measure how British MPs perceived the threat level in their districts, and they analyze how these perceptions affected roll-call 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 had been temporarily overcome, that is, the masses posed their maximum threat and this threat was ominous (i.e., high $p^{\max}$). By focusing on the current threat, they anticipate that MPs were *more* likely to vote for reform when greater riots and protests occurred in their district.

Comparing this hypothesis to the opposing one tested in Castañeda Dower et al. (2018) highlights the additional steps needed to link the theory to empirics. Aidt and Franck (2015) implicitly assume that strong challengers pose purely transitory threats and hence their average threat is low. However, suppose instead that riots and protests proxy for districts in which the masses posed *persistently* strong threats, even if not activated at all points in time. By refocusing on average threats, the model would anticipate that MPs in high-protest districts would be able to pacify the recalcitrant masses with temporary concessions rather than permanent reforms. Under these scope conditions, we would expect them to vote against the Reform Act.

The historical setting of each study differs in an important way that motivates the plausibility of the specific hypothesis assessed in each. For Aidt and Franck (2015), the Swing Riots were unprecedented attacks by peasants on agricultural infrastructure such as threshing machines and barns. Although the underlying economic grievances were long standing, the movement itself lacked coherent organization. Therefore, it is plausible that the contemporaneous (maximum) threat was high but the future (average) threat was low. By contrast, the types of anti-serfdom riots analyzed in Castañeda Dower et al. (2018) could, plausibly, have been sustained over longer periods, which would raise their average threat. More generally, it is important in empirical tests to identify which groups and organizations are inherently more transitory and which have greater staying power. This determines a crucial parameter, their average threat.

Within-Case Comparisons

Our analysis suggests a more direct specification to test theories of dynamic commitment problems with time series data: comparing within a single case an estimate of the threat *at a particular time* to an estimate of the expected threat in the future. Assuming one can come up with reliable estimates of both quantities (and setting aside other causal inference challenges), the current threat should be positively associated with conflict (or institutional reform) whereas the expected future threat should be negatively associated with conflict. Put another way, although models of dynamic commitment problems make ambiguous *cross-case* predictions about challenger strength, they make straightforward *within-case* predictions about what kinds of periods are most likely to involve conflict or institutional reform.

One example is empirical tests that examine how shocks in rainfall and the onset of droughts in Africa have influenced democratic reforms (Aidt & Leon, 2016; Brückner & Ciccone, 2011). These studies link their research designs to the Acemoglu and Robinson (2006) model and emphasize that transitory economic shocks create the conditions under which the masses pose a temporary, but not permanent, threat of revolution. Their empirical tests focus on Sub-Saharan Africa because these economies are predominantly agricultural by contemporary global standards. Therefore, fluctuations in rainfall should more greatly affect economic output.

These empirical tests analyze changes over time within cases, a design for which the ambiguous theoretical effects of challenger strength are less problematic. However, even in these cases, more extensive discussion of scope conditions would further tighten the connection between the empirical tests and the theoretical model. Transitory economic shocks lead to democratic reforms in equilibrium only in societies for which the average threat is low and the maximum threat is high. This scope condition appears plausible in this setting. During the Cold War, most African countries were closed dictatorships with no meaningful mass political participation, which made mobilization difficult (low average threat). Yet, many of these states had tenuous control over their entire territory and weak command over the national military, which made the maximum threat high in the rare periods in which the masses could solve their collective action problem. In light of our theoretical discussion, we encourage authors to routinely address (and elaborate upon) points such as these.

Conclusion

How does a challenger's coercive strength affect prospects for conflict and/or institutional reform? We established that the relationship depends on how challenger strength affects its average and maximum threats. Higher average threats lead to less conflict and fewer institutional concessions, whereas higher maximum threats yield the opposite implications. In general, a stronger challenger—such as a bigger non-elite class, a better-organized civil society, or a more advanced neighboring state—poses a greater average and maximum threat, which yields ambiguous theoretical implications. We summarized existing theoretical debates and empirical tests from this perspective.

Our analysis underscores the importance of formalizing theoretical intuitions. We highlight that existing work conceptualizes challenger strength in different ways that lead to divergent theoretical implications. However, the ambiguous consequences of challenger strength are not a product of nor a flaw with game-theoretic modeling per se. Instead, formalization enables us to clarify the conceptual difference between average and maximum threats, explain precisely why the implications are ambiguous, and characterize the conditions under which the implications cut in one direction or the other. Our modeling approach, in turn, is possible because of the advances in applied formal theory pioneered by the authors referenced throughout this article as well as many others. Clarifying theoretical implications is also crucial for empirical testing.

We highlight some fundamental impediments to empirically measuring key parameters from models of dynamic commitment problems, but we also provide some suggestions for which empirical tests are most convincing (at least with regard to the considerations raised here). Within-case comparisons elide the main source of theoretical ambiguity, although cross-case comparisons are viable if authors specify the extent to which the challenger they study has long-run staying power. Another possible approach in future work is to use structural models.[8] In general, this estimation technique is helpful when models make countervailing predictions about comparative statics and the theoretical parameters are difficult to measure directly (see, for example, Crisman-Cox & Gibilisco, 2018). The goal would be to find the version of the model, in particular the distribution of threats $F(p)$, that best fits the data and then to compare the frequency of conflict and democratic reforms under counterfactual distributions. Of course, given limitations to available data, it may be difficult to distinguish distributions that vary only on the maximum threat level.

Our model can also be used as a foundation for future theoretical work. We show that modeling a general distribution of challenger threats can be quite tractable while also highlighting when restricting to a binary distribution entails minimal loss in generality. Beyond “challenger strength” specifically, our theoretical results provide 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 repression changes the distribution of the challenger’s probability of winning. If repression creates a uniform downward shift in these probabilities, then 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). Overall, future work can build on our insights to examine how factors such as repression, technology for mobilization, and economic fundamentals affect prospects for conflict and institutional reform.

Acknowledgments

Many thanks to Johnathan Guy and Otto Kienitz for excellent research assistance, and to Scott Gehlbach, Anne Meng, and Zhaotian Luo for comments and suggestions.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Andrew T. Little https://orcid.org/0000-0001-9911-4291

Supplemental Material

Supplemental material for this article is available online.

Notes

1. For democratization, see Acemoglu and Robinson (2006); Ansell and Samuels (2014); Leventoğlu (2014); Castañeda Dower et al. (2018). For authoritarian power sharing and democratic separation of powers, see Helmke (2017); Christensen and Gibilisco (2023); Meng (2019); Powell (2021); Paine (2022). 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 (2006); Debs and Monteiro (2014); Krainin (2017). For the general mechanism, see Powell (2004).
2. A sufficient but not necessary condition for this to hold is if $F(p, s_1)$ has first-order stochastic dominance over $F(p, s_2)$ for any $s_1 > s_2$.
3. Of course, changing $p^{\max}$ also exerts an indirect effect by increasing the average future threat. For example, if p_t follows a discrete distribution in which the probability of the maximum threat is q, then increasing $p^{\max}$ by one unit also increases $\overline{p}$ by q. To keep $\overline{p}$ fixed, the increase in $p^{\max}$ could be offset by decreasing other threat levels. Alternatively, even when allowing $\overline{p}$ to increase commensurately with an increase in $p^{\max}$, the net effect is to increase $\tau(s)$ by $1 - \delta q > 0$, hence unambiguously making conflict more likely.
4. See Acemoglu and Robinson (2006), Chapter 5.

5. Again, this is a slightly different interpretation of their stochastic cost-of-fighting parameter, but is conceptually equivalent (see Appendix A.4).
6. See Acemoglu and Robinson (2006), Chapter 6. They also introduce a strategic option for elites to repress the masses, which lies outside the scope of our discussion here and hence we ignore it.
7. 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.
8. We thank an anonymous reviewer for highlighting this point.

References

Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.

Aidt, T. S., & Franck, R. (2015). Democratization under the threat of revolution: Evidence from the Great reform Act of 1832. *Econometrica*, *83*(2), 505–547. https://doi.org/10.3982/ecta11484

Aidt, T. S., & Leon, G. (2016). The democratic window of opportunity: Evidence from riots in sub-saharan Africa. *Journal of Conflict Resolution*, *60*(4), 694–717. https://doi.org/10.1177/0022002714564014

Ansell, B. W., & Samuels, D. J. (2014). *Inequality and democratization: An elite competition approach*. Cambridge University Press.

Brückner, M., & Ciccone, A. (2011). Rain and the democratic window of opportunity. *Econometrica*, *79*(3), 923–947.

Chassang, S., & Padro-i Miquel, G. (2009). Economic shocks and civil war. *Quarterly Journal of Political Science*, *4*(3), 211–228. https://doi.org/10.1561/100.00008072

Christensen, D., & Gibilisco, M. (2023). How budgets shape power sharing in autocracies. *Quarterly Journal of Political Science*. forthcoming.

Crisman-Cox, C., & Gibilisco, M. (2018). Audience costs and the dynamics of war and peace. *American Journal of Political Science*, *62*(3), 566–580. https://doi.org/10.1111/ajps.12347

Debs, A., & Monteiro, N. P. (2014). Known unknowns: Power shifts, uncertainty, and war. *International Organization*, *68*(1), 1–31. https://doi.org/10.1017/s0020818313000192

Castañeda Dower, P., Finkel, E., Gehlbach, S., & Nafziger, S. (2018). Collective action and representation in autocracies: Evidence from Russia's Great reforms. *American Political Science Review*, *112*(1), 125–147. https://doi.org/10.1017/s0003055417000454

Fearon, J. D. (1995). Rationalist explanations for war. *International Organization*, *49*(3), 379–414. https://doi.org/10.1017/s0020818300033324

Fearon, J. D. (2004). Why do some civil wars last so much longer than others? *Journal of Peace Research*, *41*(3), 275–301. https://doi.org/10.1177/0022343304043770

Gibilisco, M. (2021). Decentralization, repression, and gambling for unity. *The Journal of Politics*, *83*(4), 1353–1368. https://doi.org/10.1086/711626

Goodwin, J. (2001). *No other way out: States and revolutionary movements, 1945–1991*. Cambridge University Press.

Helmke, G. (2017). *Institutions on the edge: The origins and consequences of inter-branch crises in Latin America*. Cambridge University Press.

Krainin, C. (2017). Preventive war as a result of long term shifts in power. *Political Science Research and Methods*, *5*(1), 103–121. https://doi.org/10.1017/psrm.2015.35

Leventoğlu, B. (2014). Social mobility, middle class, and political transitions. *Journal of Conflict Resolution*, *58*(5), 825–864. https://doi.org/10.1177/0022002713478563

Meng, A. (2019). Accessing the state: Executive constraints and credible commitment in dictatorship. *Journal of Theoretical Politics*, *31*(4), 568–599. https://doi.org/10.1177/0951629819875515

Paine, J. (2022). Strategic power sharing: Commitment, capability, and authoritarian survival. *The Journal of Politics*, *84*(2), 1226–1232. https://doi.org/10.1086/715593

Powell, R. (2004). The inefficient use of power: Costly conflict with complete information. *American Political Science Review*, *98*(2), 231–241. https://doi.org/10.1017/s000305540400111x

Powell, R. (2006). War as a commitment problem. *International Organization*, *60*(01), 169–203. https://doi.org/10.1017/s0020818306060061

Powell, R. (2012). Persistent fighting and shifting power. *American Journal of Political Science*, *56*(3), 620–637. https://doi.org/10.1111/j.1540-5907.2011.00575.x

Powell, R. (2021). Power sharing with weak institutions. Working paper.

Walter, B. F. (2009). Bargaining failures and civil war. *Annual Review of Political Science*, *12*(1), 243–261. https://doi.org/10.1146/annurev.polisci.10.101405.135301

Author Biographies

Jack Paine is an associate professor of political science at Emory University. His research interests include authoritarian politics, democratization, colonialism, and formal theory. Recent articles appear in American Journal of Political Science, American Political Science Review, Journal of Politcs, International Organization, and Quarterly Journal of Political Science.

Andrew Little is an associate professor of political science at the University of California, Berkeley. Most of his current research uses formal models to study the origins and political consequences of biased beliefs.