

Slow to Learn: Bargaining, Uncertainty, and the Calculus of Conquest*

William Spaniel[†] Peter Bils[‡]

June 20, 2016

Abstract

If peace fails due to incomplete information and incentives to misrepresent power or resolve, war is supposed to serve as a learning process and allow parties to reach a mutually preferable bargain. We explore crisis bargaining under a third type of uncertainty: the extent to which one side wishes to conquer the other. With incomplete information and take-it-or-leave-it negotiations, this type of uncertainty is isomorphic to incomplete information about the probability of victory. However, with incomplete information and bargaining while fighting, standard convergence results fail: types fail to fully separate because there is no differential cost for delay. Wars correspondingly last longer while benefiting no one. These results help explain empirical differences between territorial versus non-territorial conflicts and interstate versus intrastate wars.

*Forthcoming at *The Journal of Conflict Resolution*. We thank Karyn Albert, James Fearon, and Brenton Kenkel for their comments on earlier drafts.

[†]Stanton Nuclear Security Postdoctoral Fellow, Center for International Security and Cooperation, Stanford University. (williamspaniel@gmail.com, <http://williamspaniel.com>).

[‡]Department of Political Science, University of Rochester, Harkness Hall 330, Rochester, NY 14627 (phbils@gmail.com).

1 Introduction

Consider the Taliban's dilemma in September 2001. While the United States made its desire to capture Osama Bin Laden evident, whether Washington and its allies wanted to pursue regime change in Afghanistan was unclear. Indeed, in the weeks between September 11 and the invasion of Afghanistan, declassified documents reveal that the United States attempted to reach a settlement that would keep the Taliban in power.¹ Yet these negotiations failed. Communication alone could not resolve the Taliban's uncertainty; if wanting to impose maximalist post-war divisions increases an opponent's willingness to make concessions, then more minimalist types might want to bluff accordingly.²

Similar issues have been at the root of conflict involving Russia, Georgia, and the South Ossetian and Abkhazian autonomous regions. The Georgian government has sought to reign in autonomy with a pair of civil wars, one at the end of the Cold War and another in 2008. Throughout the process, however, it has remained unclear whether Georgia would ultimately find complete control over these regions to be worth the cost, or if its ideal feasible outcome is to merely increase integration. This is problematic for Russians, South Ossetians, and Abkhazians, who might wish to maximize autonomy. Indeed, standing firm could risk future military conflict with Georgia, while conceding some autonomy would increase the likelihood of peace at the price of some self-determination.

Scholars of international relations appreciate these strategic choices states make after victory (Ikenberry, 2009). However, the literature remains quiet on how post-victory expectations affect bargaining behaviors before and during a war. This paper takes a step in that direction. While the logic of crisis bargaining is straightforward when states know of their opponents' post-war intentions, asymmetric uncertainty over these plans leads to complex and unexpected strategic behaviors.³

Initially, a reader familiar with the literature on bargaining and war may assume that uncertainty over conquest may quickly reduce to a simple risk-return tradeoff; that is, states

¹<http://nsarchive.gwu.edu/NSAEBB/NSAEBB358a/doc09.pdf>

²To wit, during his address to Congress following the attacks, President George W. Bush threatened that the Taliban would share bin Laden's fate if did not comply with American demands.

³International relations scholars have increasingly accepted the combination of asymmetric information and incentives to misrepresent as a major driver of costly conflict. Extant research, however, focuses on only two sources of uncertainty: the relative military power between the states and the costs they expect to pay in war. The cases above do not fit either category. Rather, the uncertainty was over the portion of the bargaining pie that a party found valuable; one potential type may only seek limited concessions while another might have a broad appetite for conquest.

weigh the relative likelihood of facing a conquest-willing type versus a conquest-hesitant type against the costs of war, and select a strategy that maximizes the tradeoff. The results we present below indicate otherwise. Uncertainty over the extent one side wishes to conquer the other leads to identical results as uncertainty over the probability of victory *but only in ultimatum games*. If, however, we allow bargaining to continue while the states fight, then equilibrium expectations diverge: whereas learning can occur relatively rapidly with uncertainty over power, the incentive to bluff is stronger in the case we develop. Thus, wars last longer when uncertainty over conquest is at the root of conflict.

Some intuition will clarify our findings. When uncertainty is over power, opposing types have differential rewards for fighting each battle. This is because more powerful types are more likely to prevail in each confrontation, putting themselves in a better bargaining position that a less powerful type cannot mimic. In turn, if the opponent wishes to, it can offer an amount up front that a weaker type would be willing to accept but that a stronger type would not accept. Therefore, the negotiating process leading up to battle reveals substantial information, allowing the opponent to settle with the stronger type afterward.⁴

With uncertainty over the preferred post-war implementation, such a differential cost for fighting does not necessarily exist. This is because each type is equally likely to win a battle, and they pay the same cost to do so. Consequently, the opponent cannot adequately offer an amount that completely screens out more moderate types. We show that incentive compatibility constraints lead to inefficient fighting if the opponent wishes to gear its offer toward the moderate. Indeed, over the whole course of fighting, it consistently offers an amount that only the moderate would accept. The moderate sometimes accepts at each of these stages but not often enough to lead to a meaningful change in the opponent's negotiation strategy.

Empirically, this result indicates that fighting to reveal the extent of one's demands is not inherently effective. For example, if the Taliban were unsure whether the United States had broad aims or only wished to capture Bin Laden, American initiation of costly conflict is insufficient to convey that information. This has unfortunate second-order effects for types that wish to capture larger shares after military victory. Because only a marginal amount

⁴A similar learning process can occur with uncertainty over an opponent's resolve. Although all types have the same probability of winning each battle, the proposer can still structure a series of offers to induce separation. Less resolved types accept at earlier stages because the proposer selects an early offer such that bluffing strength by fighting a battle provides no more than the proposed division. Meanwhile, more resolved types reject these offers because their lower costs of war imply that they prefer separating by fighting a battle.

of learning occurs, the proposer continues offering an amount unacceptable to these more expansionist types. In turn, these types never settle; rather, they fight the war until its military conclusion.

This paper speaks to two related literatures on war intensity and duration. First, the literature on convergence sees war as a continuation of bargaining (Slantchev 2003; Powell 2004; Filson and Werner 2002; Smith and Stam 2004). As actors fight, they learn about their relative capabilities and resolve. Inefficient conflict thus ends as the cause of the original bargaining breakdown (uncertainty) dissolves over time. These models, however, only address uncertainty about power or resolve. Each state completely understands the division of the stakes its opponent would want to impose if it were at the opponent's mercy. While it may initially appear that the logic of uncertainty about the extent of ideal post-war settlements would follow similarly, our results below indicate that fighting provides far less information about those preferences.

Wolford, Reiter and Carrubba (2011) follows from this literature and has similarities and differences with our paper worth highlighting. Both our work and theirs note that conflicts may last longer than the initial convergence literature suggested even though fighting provides information in both our models. Our papers differ in the source of uncertainty. The uninformed party in Wolford et al uses screening offers to determine whether it should pursue a settlement or fight a preventive war; the state pursues expansive war aims (the destruction of its opponent) in the latter case as a means to an end. In our paper, the uninformed party does not know whether the opponent values part of the policy space or territory in contention. Thus, the source of uncertainty is more accurately the “peace aims” of a state, rather than the war aims.⁵

Second, our paper addresses a puzzle in explaining long wars. In a review of the bargaining literature, Walter (2009, 246) finds that “[i]nformation asymmetries . . . do not explain the subset of conflicts where combatants do not sign or implement settlements even after long wars have been fought and much information revealed.” Many scholars also cast doubt on the information narrative of long wars and instead argue for commitment-based explanations (Powell 2006, 170; Fearon 2004, 290; Reiter 2009, 34). Our model reveals that uncertainty over such peace aims creates particularly intractable conflicts. Indeed, conditional on war occurring, the proposer never offers an amount that the more expansionist type would accept. Rather, it continually tries to screen the moderate type. Wars against the expansionist

⁵We are therefore careful not to use the phrase “war aims” throughout this paper.

type must therefore end with the military defeat of one party.

This helps explain why two types of wars last longer than others. Whereas most interstate wars end in negotiated settlement, only around one in five civil wars ends short of complete military destruction of one party. Correspondingly, civil wars tend to last a long time (Fearon 2004). Walter (1997) argues that the inability to credibly commit to the terms of a settlement once implemented helps explain why negotiations fail. Meanwhile, Fearon (2007) points to a similar commitment issue over a shorter time-horizon—once a party accepts an offer, the opponent cannot help but update its belief about its strength and then demand a larger amount. In turn, civil war combatants engage in non-serious bargaining for an extended period of time. Our paper demonstrates that information problems can explain these long civil wars, and thus all long civil wars are not necessarily the product of commitment problems.

Second, a large literature on territorial conflict (Hensel 1996; Vasquez and Henehan 2001; Walter 2003; Huth 2009) claims that wars fought over territorial issues tend to be intense and long-lasting. A common theoretical argument is that territorial wars involve inherently more valuable stakes, hence the reluctance to capitulate. Our model provides an alternative explanation for the puzzle: the fighting process cannot efficiently communicate the combatants' post-war plans. Therefore, in such wars, we would expect casualties to accrue and for the parties to reach fewer negotiated settlements.

2 The Calculus of Conquest

Wars do not simply end. Rather, even warring parties that achieve complete military victory over their opponents must choose the type of post-war settlement to implement. Ikenberry (2009, 4) provides a useful typology. On one end of the spectrum, a victor may assert complete dominance over the befallen, pillage the countryside, loot all plunderable resources, and enslave the population. Other winners may elect for more moderate post-war policies, such as abandoning the region entirely or transforming domestic political institutions to be more compliant with the conquerer.

Post-conflict settlements in the World Wars illustrate the variance in outcomes. The Treaty of Versailles, for example, implemented a highly extractive policy against Germany, resulting in loss of territory and billions in reparation payments. The Western settlement with Germany and Japan following World War II was comparatively kind, with the United

States funneling large sums of money to both nations to rebuild. The U.S. also permitted the Emperor of Japan to remain a figurehead despite insisting on unconditional surrender. Yet we also observe variance within World War II. Nazi treatment of its fallen targets varied greatly (Hollander 2008). And while the United States sought to jump start the German economy, the Soviet Union desperately raced to strip East German factories of all their useful parts (Naimark 1995, 141-204; Stone 2002, 27-28) and completely devalued the Reichsmark (Turner 1987, 24).

What accounts for such variation? Following existing research on post-war settlement implementation, five factors are at play: logistics, the target population, domestic negotiations, adherence to territorial norms, and the preferences of the conquering state. First, logistics form a necessary condition. A conquering force requires supply lines and adequate military support to coerce local populations into complying with the demands. Yet resolving military logistics problems is not simple (Van Creveld 2004). Even the United States—as powerful as it was—had difficulties administering Afghanistan. Indeed, at one point Washington halted construction on a European missile defense shield to just convince Russia to allow shipments through its border. As the scope of the post-war demands increase, so too do the logistical problems. For example, capturing critical targets like Osama Bin Laden has fewer logistical hurdles than administering an entire country. Weaker states must find the optimal tradeoff between what they might wish to achieve and what they can realistically accomplish.

Second, attributes of the conquered society partially determine the profitability of extraction. Societies better at footdragging or hiding valuable production (Scott 2008) increase the cost of administration and monitoring, decreasing the desirability of extreme post-war demands. Similarly, societies better equipped to coordinate resistance make less attractive targets (Acemoglu, Verdier and Robinson 2004). Existing internal divisions (Kenkel 2013) and the presence of extractable natural resources also affect a conquerer's calculus. And going back to logistics, a conquerer's ability to overcome each of these problems leads to greater war demands.

Third, domestic negotiations between a leader and his or her coalition determine the extent to which a state can extract a greater share of the good post-war. One determinant here is the availability of labor and capital to conduct extraction projects (Kedziora 2012).⁶ The spoils of war do not always go to those who pay the costs of conflict. In turn, ideal

⁶See also Bueno de Mesquita et al. (2003) for a more general framework of internal negotiations for power and resources.

post-war policies—and therefore reservation values in bargaining—depend on the leader’s ability to convince the population to fund the project and the outcomes a leader expects to suffer in the absence of success (Goemans 2000). Greater leeway permits a leader to push for greater rates of extraction, even if the country as a whole must pay the startup costs. Such tensions may create uncertainty for opposing countries, which may not be as readily able to anticipate the outcome of domestic negotiations (Iida 1993). This played an important role in the Mexican-American War, which we discuss later as an illustration of a couple of our key propositions.

Fourth, scholars have observed a growing norm for territorial boundaries (Zacher 2001). Under this premise, land is no longer an object that states may acquire through the use of force. Rather, they ought to respect existing divisions unless peaceful agreements dictate otherwise. Compliance to this norm undoubtedly varies. Nevertheless, states wishing to maintain the norm might concede territory acquired in the process of fighting.⁷

Finally, humanitarian concerns might limit a conquerer’s extraction policies. While conquest can pay (Lieberman 1998; Herbst 2014), the process invites small-scale militarized resistance. Effective countermeasures are often brutal, so much so that democratic audiences might wish to forgo the process entirely (Arreguin-Toft 2001). Thus, for states with a clear preference against extraction, the other factors (logistics, footdragging, and domestic negotiations) are irrelevant if the winner would still wish to limit its imposition even under rosy conditions.

Several other factors could also provide micro-foundations for the assumption that states may find part of the bargaining pie undesirable. The conquering state may worry that consuming a larger amount of the good could create a bargaining problem with a third party in the future (Chiba and Reed 2014; Siverson and Starr 1990; Blainey 1988). Third-party responses through international institutions may also create an incentive for a state to only consume a portion of the disputed good (Huth, Croco and Appel 2011). Additionally, leaders may make limited claims if they are worried about suffering domestic audience costs should they not achieve their stated goals (Fearon 1994; Tarar and Leventoglu 2009). Lastly, reluctance to capture a large share of a good may come from a more complex post-war bargaining problem with the conquerer and groups of conquered citizens (Spolaore and Alesina 2005;

⁷Such concessions are common throughout history, even before scholars believe the norm became widespread following World War II. For example, the United States had captured Mexico City and portions of Baja California but relinquished them at the end of the war.

[Kenkel 2013](#)).

Regardless of the microfoundation, existing models are silent on the variation in the extent of conquering. Instead, modelers standardize the bargaining pie to value 1, and the winner simply takes the entirety of that good.⁸ The above discussion, however, indicates that the scope of conflict varies from case to case. It is thus unclear how willingness to conquer alters the bargaining environment. The model below shows that the implication is not immediately obvious. Despite a substantial discussion above on the cost tolerance for conquering, more expansionary desires manifest in the model as isomorphic to greater military power.

Further, by not explicitly modeling optimal conquest, existing treatments implicitly assume that states have complete information on the subject. That is, everyone wishes to conquer the entire good, and this is common knowledge.⁹ However, one cannot standardize the good if an opposing party is not actually sure what ultimately is at stake. And it is reasonable to believe that would-be conquering states have private information about the extent of their post-war settlement wishes. For example, the result of domestic bargaining should be readily understood internally but is difficult to decipher externally.¹⁰ It may also be unclear how costly a state finds breaking the norm to respect existing divisions. In addition, foreign states would have a hard time knowing the exact details their opponents' logistical concerns, willingness to pursue extractive policies, and recognition of territorial norms. The standard model remains silent here. Yet, as we show in the model below, this type of uncertainty behaves differently from existing sources in environments with bargaining and learning while fighting.

3 Ultimatum Bargaining over Policy

We begin with a one-shot bargaining model similar to [Fearon \(1995\)](#). In addition to showing the setup's close relationship with the standard bargaining model of war, the game played here will be a subgame of our complete model.

⁸This standardization accounts for how much a state values the good in dispute. More resolved states—those that have a higher value for the prize—receive lower costs for war via the standardization. Resolve thus affects only one party's payoff. Here, in contrast, uncertainty about the post-war policy ultimately imposed affects both parties' payoffs, as minimalist demands reduce one side's payoff while increasing the other's.

⁹Although this setup is standard, see [Schultz and Goemans \(2014\)](#) for a recent exception.

¹⁰To wit, researchers see uncertainty over just a single leader as a source of conflict ([Wolford 2007](#); [Rider 2013](#)). Uncertainty with multiple domestic actors is exponentially more complicated.

Suppose two states, A and B are bargaining over a good, represented by the unit interval, in the shadow of war. A begins by making a demand $x \in [0, 1]$. After observing x , B can accept or reject A's offer. If B accepts, then the game ends, with A receiving x and B receiving $1 - x$. If B rejects, then the states fight a war; A prevails with probability p_A , B wins with complementary probability, and the sides pay respective costs $c_A, c_B > 0$. Following the war, the interaction the ends.

Our model breaks from the traditional setup in the following way. Ordinarily, the model assumes that the winner takes the entire good and the loser receives nothing. Here, we generalize the standard model by allowing for limited claims. Specifically, if state B is victorious, it lets A keep some of the good. Formally, $m \in [0, 1]$ denotes the amount donated back to A, leaving B to keep $1 - m$ of the good for itself. Thus, the parameter m represents B's level of moderation. We view this setup as a reduced-form game, with m resulting from the various factors discussed in the previous section. That is, states that have limited claims, find conquest too expensive, or worry about third-party intervention have correspondingly larger m values.

The game is straightforward to solve with complete information. State A's reservation value for war equals $(p_A)(1) + (1 - p_A)(m) - c_A$. As such, A is willing to accept any offer x such that

$$x \geq p_A(1 - m) + m - c_A. \quad (1)$$

Meanwhile, B's payoff equals $(p_A)(0) + (1 - p_A)(1 - m) - c_B$. Thus, B is willing to accept any division such that

$$1 - x \geq (1 - p_A)(1 - m) - c_B. \quad (2)$$

Combining 1 and 2 shows that a set of mutually acceptable outcomes exist if:

$$c_A + c_B \geq 0.$$

This inequality holds by assumption. Thus, if state A made an ultimatum offer, it would demand $x = p_A + m(1 - p_A) + c_B$ in the unique subgame perfect equilibrium and B would accept if and only if $x \geq p_A + m(1 - p_A) + c_B$.

Note that this model is isomorphic to the standard bargaining model of war with complete information. The standard format centers the bargaining range around p . In this case, the

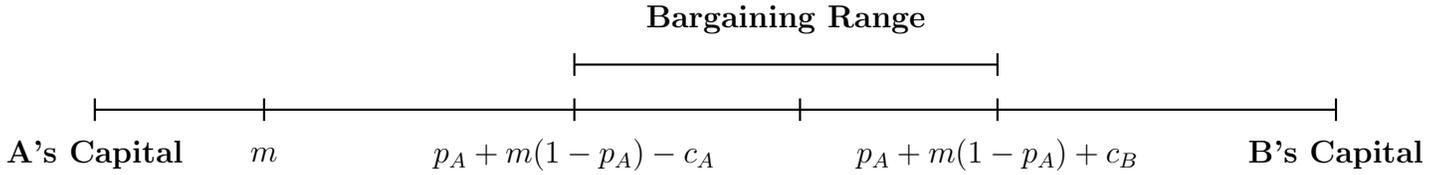


Figure 1: A geometric interpretation of the bargaining problem with complete information. Both parties find any settlement on the interval $[p_A(1 - m) + m - c_A, p_A(1 - m) + m + c_B]$ to be mutually acceptable.

range forms around $p_A + m(1 - p_A)$. However, $p_A + m(1 - p_A) \in [0, 1]$. Since the standard model requires $p_A \in [0, 1]$ as well, the models are isomorphic by setting the value of p_A in the original game equal to $p_A + m(1 - p_A)$ from our model.

What explains this connection? The standard interpretation of p_A in most crisis bargaining models is the probability that A prevails in conflict and takes the entire bargaining good. That said, an alternative interpretation is the portion of the good A expects to receive through war.¹¹ Whether a state pursues moderate policy goals post-war in part determines the expected distribution of the good through war. As a result, the difference (in a one-shot game) is merely one of notation— p_A for probability, m for moderation.

4 Ultimatum Bargaining with Incomplete Information

Regardless, we are more interested in the incomplete information case. That is, we acknowledge that actors might be unaware of their opponent’s willingness to conquer, their desire to uphold territorial norms, and the result of domestic negotiations of claims. This type of uncertainty is fundamentally different from uncertainty over power or costs of war and is consequently worth modeling to investigate whether the standard results apply.

Consider the previous setup with the following modification. The game now begins with Nature drawing the division B would prefer to choose to implement if it won a war from a commonly known prior distribution. Specifically, state B’s preferred division is $m \in [0, 1]$ with probability r and is, without loss of generality, 0 with probability $1 - r$. We say that the m type is a *moderate* and the 0 type is an *extremist*. State B observes its type, while state A remains uncertain whether B is a moderate or extremist. After B observes its type,

¹¹Some authors (Fey and Ramsay 2011; Arena 2015b) use the notation w_i in place of p for a similar purpose.

the parties negotiate in the manner previously described.

We now solve for the game's perfect Bayesian equilibrium (PBE). A PBE is a set of strategies and beliefs such that the strategies are sequentially rational and beliefs are updated via Bayes' rule wherever possible. Proposition 1 below gives the game's solution for the non-corner solution.¹² Because the uninformed actor moves first, the equilibrium does not require any Bayesian updating:

Proposition 1. *Let $r^* = \frac{c_A + c_B}{m(1-p_A) + c_A + c_B}$. The following defines State A's unique equilibrium demand x^* :*

$$x^* = \begin{cases} p_A + m(1 - p_A) + c_B & \text{if } r > r^* \\ p_A + c_B & \text{if } r < r^* \end{cases}$$

The moderate accepts an offer x if and only if $x \leq p_A + m(1 - p_A) + c_B$ and the extremist accepts if and only if $x \leq p_A + c_B$.

The intuition is the standard risk-return tradeoff. Moderate types are easier to buy off because they have less at stake, while extremists require deeper concessions because they want to take more of the good in the event of war. As such, the proposer must decide whether to be conservative with its demands and induce all possible types to accept or go aggressive and force the extremist into war. The former case avoids the costs of conflict but comes at the price of missing out on potential concessions from the moderate type.

To briefly illustrate this dynamic, consider the uncertainty surrounding the United States' position on westward expansion on the eve of the Mexican-American War. Although President James Polk was consistent in his desire to acquire California, he faced opposition both with the public and in the Senate.¹³ Indeed, Northerners and Whigs feared that westward expansion along the southern border would upset delicate balance between free and slave states (Tutorow 1978, 132). They instead wished to focus Westward expansion in the Oregon/Washington area and were already upset at Polk for signing the Oregon Treaty. This

¹²That is, $p_A + m(1 - p_A) + c_B < 1$.

¹³Although we focus on the uncertainty Mexico faced in understanding the result of domestic bargaining in Washington, it is worth noting that the familiar incentives to misrepresent exist in this type of interaction. That is, because more expansionist types earn more concessions in negotiations, less expansionist types have incentive to mimic their behaviors in cheap talk communications. Consequently, even if the United States had presented a consistent position on westward expansion, Mexico should not have necessarily taken the U.S. at its word.

group made substantial electoral gains during Polk’s midterm election (Wheelan 2007, 296-297). Others thought that expansion in general would “doom democracy” (Henderson 2007, 134).

Overall, according to future President Ulysses S. Grant (then an army lieutenant), support for the war was so flimsy that the goal of the first military expedition was to provoke a Mexican attack. This, in theory, was the key to giving Polk leverage over the Senate and his Democratic resistance (Grant 2006, 21). Thus, even if Mexico knew the policy position of all the key players involved, the country faced asymmetric uncertainty in understanding which domestic group would ultimately prevail in the event of military defeat and also whether Polk’s gambit would win over his domestic audience. Mexican politicians correspondingly called for bold action (Henderson 2007, 150). This contributed to the incompatible demands that eventually led to war.

Before moving on, it is again worth noting the isomorphism to standard results:

Remark 1. Consider the standard crisis bargaining game in which A is uncertain whether its probability of victory equals p_A or $p'_A > p_A$ from a common prior distribution $(r, 1 - r)$. The game is isomorphic to the game with uncertainty about outcome implementation by using the substitution $p'_A = p_A + m(1 - p_A)$.

Put differently, when we black box war as a one-shot costly lottery, uncertainty about outcome implementation is identical to uncertainty about relative power. Intuitively, this is because the probability of victory is isomorphic to the average distribution of the good that war produces. Consequently, when state B is a moderate, A’s share of the war outcome increases because B will not take m portion of the good regardless of who wins or loses. This effectively gives A a level of power equal to $p'_A = p_A + m(1 - p_A)$. In contrast, when B is an extremist, A’s share of the war outcome diminishes. This effectively gives A the smaller power level of p_A .

More explicitly, by “isomorphic to the game with uncertainty about outcome implementation,” we mean that one could take the states’ best response correspondence for either game, use the substitution $p'_A = p_A + m(1 - p_A)$, and finish with the best response correspondence for the other game. Despite having two disparate sources of uncertainty, the only formal difference is one of notation. The underlying strategic considerations turn out to be identical. Our appendix gives a formal proof of this. The process requires solving for the standard model with uncertainty about power and deriving the comparison between the two.

This isomorphism might be surprising. After all, one might suspect that a state’s differing appetite for conquest might manifest itself in the cost parameter, as costs in bargaining model of war implicitly represent a state’s value for the bargaining good. Nevertheless, there is a clear intuition as to why uncertainty over conquest is instead isomorphic to uncertainty over power. With uncertainty over power, the uninformed party does not know what to expect to receive from war due to the different possible probabilities of victory. With uncertainty over moderation, the uninformed party does not to what to expect to receive from war due to the different possible post-war divisions. Although the origins differ, the uninformed state lacks the knowledge of the *interdependent* value of war in both cases. In contrast, with uncertainty over costs, the uninformed state simply does not know its opponent’s value for war. As a result, uncertainty over moderation matches uncertainty over power, not uncertainty over costs.

Under normal circumstances, we would ignore the knife-edge case of $r = r^*$ for the standard reasons. In practice, state A is indifferent between either of the two offers here. Thus, the PBE is not unique under these circumstances. This becomes critical in the bargaining-while-fighting game. As we will see, no fully separating equilibria exist for certain parameter spaces. In these cases, the moderate type must mix. The indifference conditions necessary to sustain the equilibrium require creating that exact posterior, which then permits state A to mix between the high offer and low offer in the second bargaining stage.

5 Bargaining while Fighting

The previous section contained a model that black-boxed war. In practice, actors continue the bargaining process as fighting continues, allowing the parties to reach a settlement short of complete military defeat of one side. Indeed, most interstate wars end with a negotiated settlement, while a sizeable portion of intrastate wars do as well.

Researchers have already offered a number of models of bargaining while fighting, though none investigates limited claims as the source of incomplete information. Because the model in [Filson and Werner \(2002\)](#) allows for the easiest interpretation, our setup most closely resembles the structure of their model. We study the simplest possible interaction of bargaining while fighting, solving a model that allows for *at most* two periods of negotiations. This simplicity permits us to derive and analyze an explicit solution. From there, we discuss how the intuition elucidates how these results would extend to a longer process, in which bargaining and fighting occurs over a large number of periods.

5.1 Setup

Nature begins the game by drawing state B as a moderate with probability q and an extremist with complementary probability.¹⁴ As before, these types only differ in their preferred level of consumption. The states negotiate and potentially battle over the issue.

State A begins the first period of bargaining by demanding $x_1 \in [0, 1]$. State B sees the demand and accepts it or rejects it. Accepting ends the game. Rejecting leads to a battle. State A prevails in the battle with probability p_A and state B wins with complementary probability. Each pays a respective cost of $c_A, c_B > 0$.

If state B loses the battle then it is militarily defeated and state A can implement its desired outcome. If state A loses the battle then bargaining continues. State A begins the second period of bargaining by demanding $x_2 \in [0, 1]$. State B sees the demand and accepts or rejects it. Accepting still ends the game while rejecting leads to a battle. Once more, state A prevails in the battle with probability p_A and state B wins with complementary probability. Each pays a respective cost of $c_A, c_B > 0$.

If State B is defeated, state A again implements its desired outcome. If state A loses, and so has lost two battles, it is defeated and state B can choose its preferred outcome. Regardless of the outcome, the game ends after the second period.¹⁵

Note that the subgame beginning in the second period is identical to the one-shot bargaining game discussed above. The only change is that we set r as state A's posterior belief that B is a moderate, where r is derived using Bayes' rule whenever possible. Since we have already solved this subgame, our inquiry only has two remaining questions, both about the first period of bargaining. First, given any particular proposal from A, how do the types of B wish to influence A's beliefs for the next period? And second, given B's equilibrium response to all initial proposals, which proposal maximizes A's overall welfare? We answer these questions below.

¹⁴We now use r to denote state A's posterior belief in the second stage, as that subgame is the one-shot game we just solved for when the total number of possible battles is 2.

¹⁵For concreteness, one might imagine the states having military units, with state A having two military units and state B having one military unit. Losing a battle depletes a military unit. If a state depletes all of its units then it is defeated by the opposing state.

5.2 Equilibrium

We begin searching for equilibria by looking for non-corner solutions. Specifically, we look at the condition where m is less than the minimum of $p_A + c_B$ and $2p_A - p_A^2 + 2c_B - p_A c_B$. This condition ensures that state B will fully consume the remainder of state A's demand in both stages of bargaining. In standard crisis bargaining models, such solutions are normally substantively identical to corner solutions. Later, we will show that this is not the case with uncertainty over outcome implementation—convergence works (i.e., a fully separating equilibrium exists) for the corner solution but it does not work (i.e., a semi-separating equilibrium exists, but no fully separating equilibria exist) for the non-corner solution.

We are now present for our first result, characterizing the semi-separating equilibrium.

Proposition 2. *If q is sufficiently high, A demands $x_1 = 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_A c_B$. The extremist rejects this offer with probability 1, while the moderate rejects with probability $\sigma_R^* = \frac{(1-q)(c_A + c_B)}{q[m(1-p_A)]}$. After observing a rejection, in the second period, A's posterior equals r^* . A demands $p_A + m(1 - p_A) + c_B$; the extremist rejects with probability 1 while the moderate accepts with probability 1.*

See the appendix for a complete proof and formal derivation of the cutpoint for q , which is notationally cumbersome. To recap, in the first period, A tries to screen out the moderate type by offering concessions insufficient to induce the extremist type to accept. To some degree, this is successful—a portion of the moderate types accepts, and A updates its belief to r^* accordingly. But even as A learns something about B, the amount of separation too small to meaningfully change A's behavior. Indeed, if the parties reach the second period of negotiations, A once again offers an amount that only the moderate type accepts. In contrast, a successful screen would induce the moderate to accept with certainty in the first stage, allowing A to tailor its offer in the second stage to ensure the extremist type's compliance.

Why does convergence fail here? The key is to first understand why convergence normally succeeds. When the proposer faces uncertainty about the probability of victory, the differing types find war differentially risky. The proposer begins his thought process by calculating the amount of concessions he will have to offer to induce the strong type to accept in the second stage. He then calculates the weak type's payoff for bluffing strength by fighting in the first stage and accepting an offer geared toward the stronger type in the second stage.

The proposer then offers that amount. The weak type accepts. The strong type, however, rejects. This is because the weak type was indifferent between accepting that offer and trying to survive the battle and reach the next stage. But the strong type is inherently more likely to survive the battle and therefore finds accepting that offer strictly worse than continuing. This leads to a clean screening process—the weak type accepts in the first stage, and the strong type accepts in the second stage.¹⁶ The proposer has no need to continue the war to its final conclusion since it does not need to protect against potential bluffers.

The story is similar when the uncertainty is over the opposing party’s resolve. The proposer again begins by calculating the amount of concessions he will have to offer to induce the resolved type to accept in the second stage. He then calculates the unresolved type’s expected utility for fighting a battle in the first period and offers that amount. The unresolved type is willing to accept by construction. On the other hand, the resolved type has a strict preference to reject. This is because the unresolved type is indifferent and the resolved type pays a smaller per period cost of war, so fighting yields a greater expected utility. Once more, this leads to clean separation; the first offer filters out the unresolved type while the second offer filters out the resolved type.¹⁷

Neither of these screening mechanisms succeeds when uncertainty is over outcome implementation. If the moderate rejects an offer, it pays the same costs to fight a war and wins with the same probability as the extremist. Differential costs—necessary for separating behavior (Arena 2015a)—do not exist here. To understand why, imagine the proposer calculates the amount of concessions he will have to offer to induce the extremist type to accept in the second stage. Next he use that information to calculate the moderate type’s expected utility for fighting a battle in the first period. However, because the types have identical costs and probabilities of victory, the amount necessary to buy off the moderate is also the amount necessary to buy off the extremist. Further, the proposer learns nothing from the battlefield result because each type has the same probability of winning. Thus, attempting to screen in this manner ends up yielding universal acceptance.

As a result, the only way the proposer can induce the moderate to accept in the first stage is if it will again attempt to appease only the moderate in the second stage. Under such

¹⁶With more than two types, the uninformed party can also update its belief following a battle via Bayes’ rule.

¹⁷This is true for models in which each round of fighting entails some probability of military breakdown. Research into circumstances in which this is not true remains comparatively unexplored. We therefore stress that our mechanism is but one explanation for why full separation cannot work; others may exist as well.

conditions, however, the moderate cannot accept immediately—doing so would guarantee that the proposer would appease the extremist in the second stage, which in turn gives the moderate incentive to bluff by rejecting in the first stage. Instead, the moderate sometimes accepts and sometimes rejects. Rejection occurs frequently enough that the proposer would again wish to buy off the moderate type in the second stage. Since the proposer aims all of these offers at the moderate, the extremist rejects throughout.

The earlier discussion of the Mexican-American War helps illustrate this point as well. As history has shown, Polk ultimately held enough political capital to capture Alta California. Thus, in the language of our model, the United States was the more expansionary type. But fighting alone could communicate this to Mexico, as moderate types would be willing to reject initial offers if they knew that doing so would result in settlements targeting the expansionary types later. As the model predicts, conditional on war occurring, an expansionary type must fight the conflict to its military completion. The U.S. did just that. This helps explain the puzzle as to why Mexico fought as long as it did despite being at a steep military disadvantage (Henderson 2007, 148-149).

Our appendix contains a full proof. The bulk of the work is in showing that no fully separating proposal strategy works in equilibrium and that the above semi-separating proposal strategy is the best option for state A.

The remaining case occurs when state B is likely to have maximalist post-victory preferences. Proposition 3 presents the equilibrium analysis for this case.

Proposition 3. *If q is sufficiently low, A demands $x_1 = 2p_A - p_A^2 + 2c_B - p_{AcB}$ in the first period. Both types of B accept. Off the path, A can have any belief and plays according to Proposition 1.*

The intuition here is far more straightforward. The semi-separating proposal strategy from Proposition 2 is costly and risky to implement. This is because it leads to a high probability of war against the moderate in the first stage and guaranteed war throughout versus the extremist. Consequently, for state A to tailor its proposal strategy for the moderate, it must think that the moderate type is likely. Under Proposition 3's conditions, however, this is not the case. Hence, state A lowers its demands and induces immediate acceptance from all types. The appendix contains the proof, which is a natural complement to the proof for Proposition 2.

Unfortunately, these results have negative welfare implications. The fighting process is

supposed to reveal information and open up negotiated resolutions that were not possible at the start of war. No such meaningful learning occurs here. Rather than select a demand in the first stage that screens out the moderate type, it selects an amount that the moderate sometimes rejects. Its demand strategy in the second stage then yields acceptance from the moderate but guarantees an absolute war with the extremist type. Because the moderate type mixes in the first stage, A's belief about its level of power grows more pessimistic in the second stage. Yet that updating does not yield substantive change—it still stubbornly buys off only the moderate.

Why does meaningful learning not occur under these circumstances? As outlined above, the problem is the moderate type's incentive to bluff. To fully convince the moderate not to bluff, A must make its demand satisfactory to extremist type. Doing so requires state A to make substantial concessions. The alternative requires demanding an amount in both stages that only the moderate type finds acceptable. Despite such a strategy, A still suffers some amount of fighting against the moderate in the first stage. Because q is sufficiently high in Proposition 2—that is, the probability of facing the moderate is great—A prefers going through the long and inefficient process because the alternative requires giving even greater unnecessary concessions to the moderate. The process is fruitless—all parties would be better off if the moderate always accepted the initial offer up front and then A tailored its demand to the extremist in the second stage. Yet incentive compatibility constraints doom the more efficient solution.

Further, note that these results do not depend on having only two possible battles. The underlying issue is that the types do not pay differential costs or face distinct risks when they fight a battle. Consequently, the same screening problem would persist even if the states could fight many more battles. Rather, the solution is to find circumstances under which the types would find equilibrium *offers* differentially valuable. We address these circumstance below.

5.3 How the Distribution of Power Matters

As stressed earlier, the previous propositions investigated bargaining outside of the corner case. Under these conditions, if the proposer attempted to appease the extremist, the moderate would happily accept the same amount. In the corner case, the moderate would only want to consume a portion of the good and return the rest to the proposer. Ignoring these

cases ordinarily does not affect results. Here, however, the moderate's unwillingness to accept a greater share means that the moderate and extremist have different continuation values for rejecting demands in the first stage. In turn, as Proposition 4 summarizes, convergence can succeed:

Proposition 4. *If $m > p_A + c_B$, a separating equilibrium can exist. The moderate accepts state A's demand in the first stage. The extremist rejects the first demand but then accepts A's demand in the second stage.*

As just described, the intuition is in understanding the difference in continuation values for the two types. If $m > p_A + c_B$, then the most the moderate type can hope to earn in equilibrium by rejecting in the first stage equals:

$$(1 - p_A)(1 - m) - c_B.$$

This is because if the moderate wins the first battle it then receives a demand intended for the extremist, which is less than m . The moderate, however, only wants to consume a total of $1 - m$ of the good. As such state A consumes remainder.

In contrast, the extremist would want to consume the whole good. Because state A would never give more than what is necessary to the extremist, the extremist's continuation value for rejecting equals:

$$(1 - p_A)(1 - p_A - c_B) - c_B$$

Since $m > p_A + c_B$, this continuation value is greater than the best case continuation value for the moderate type. Consequently, whereas the $m < p_A + c_B$ case could not produce effective screening, the $m > p_A + c_B$ can yield a separating equilibrium in which state A demands $1 - (1 - p_A)(1 - m) - c_B$ in the first stage (inducing the moderate to accept) and demands $p_A + c_B$ in the second stage (inducing the extremist to accept if it survived the first battle).

Although the distinction between a corner case and an interior solution may seem to only be of mathematical concern, the substantial variance in outcomes yields three empirical implications. First, increasing the value of m means that the parameters are more likely to fulfill the $m > p_A + c_B$ condition. This permits full separation, which implies shorter durations of conflict. Yet larger values of m indicate more precise information—as m goes

to 0, the moderate type becomes increasingly similar to the extremist type. Thus, exacerbating the asymmetric information problem may *reduce* war. This contrasts with a common argument that uncertainty has a monotonic relationship with inefficient conflict (Reed 2003; Kydd 2010, 104).¹⁸

Second, decreasing the value of p_A also makes fulfilling the $m > p_A + c_B$ condition more likely. Noting that $1 - p_A$ represents B's probability of winning a battle, one interpretation is that full separation becomes possible when B is close to conquering A. In other words, negotiations are more likely to result in agreements if winning the initial battle means that B is very likely to win the war. Substantively, we would then expect an actor to end its wars more frequently through negotiated settlement when losing battles makes complete military collapse a real possibility. This appears to hold empirically: warring parties closer to military parity fight longer than when one state holds a preponderance of power (Slantchev 2004). Although the standard explanation for this is that states at military parity have more uncertainty to sort through, our model indicates that learning occurs more slowly under those circumstances. Parity, it appears, is an informational double whammy.

Finally, decreasing c_B also makes full separation more likely for the same reasons as above. This contrasts with standard results, which suggest that peace is most likely when the costs of war are high, given that fighting would lead to disastrous outcomes (Mueller 1989) or sacrifice large trade surpluses (Keohane and Nye 1977).¹⁹ Yet those high costs also mean that full separation cannot occur, which in turn leads to more fighting in the first stage and guaranteed conflict against the extremist in the second stage.²⁰

6 Discussion

The formal analysis revealed a straightforward claim: parties are less likely to settle crises with uncertainty over conquest than with uncertainty over power or resolve. Translating this

¹⁸See also Fey (2014) for a different argument increasing information can increase war. Note that the full relationship between m and the probability of war is non-monotonic. When m goes to 0, the extremist and moderate converge to the same preferences, which mimics the complete information case. This makes the parameters more likely to fulfill the conditions for Proposition 3, which in turn yields a safe demand and no war at any stage.

¹⁹See also Fey and Ramsay 2011.

²⁰Like the case for the parameter m , the full relationship between the cost of war and probability of fighting is non-monotonic—decreasing c_B increases the chances of hitting the parameters for Proposition 3, which yields peace with certainty.

proposition to empirical implications then requires theorizing the conditions under which we are more likely to observe the former type of uncertainty. This section offers three possibilities.²¹

First, as discussed earlier, one precondition for conquest is a willingness to engage in the enterprise in any form. If a certain subset of states generally wish to avoid conquest, uncertainty over related issues (manageable logistics, domestic negotiations, respect for norms against territorial capture, cleavages between the leader and the population) becomes moot. Put differently, while uncertainty may exist, opposing states still understand that the outcome of war will not result in conquest. Thus, if democratic states are generally known to prefer limited claims (Ikenberry 2009), this model helps explain why we would expect democracies to terminate their conflicts more quickly.²²

Ikenberry similarly argues that institutions can lock-in moderate post-war outcomes. Further, a large literature argues that institutions provide information by reducing transaction costs (Koremenos, Lipson and Snidal 2001). These two factors would reduce the relevance of conquest by reducing the possibility of it and strengthening the uncertain actor's prior regarding the situation. In turn, we would expect information transmission through war to occur more rapidly here because states can credibly communicate the remaining uncertainty (on power or resolve) through inefficient fighting.

Finally, the model provides a clear explanation for why civil wars more frequently end in military defeat and last longer than interstate wars (Walter 1997; Fearon 2004; Fearon 2007). For uncertainty over conquest to matter, warring parties must first fight over stakes that could conceivably result in occupation and regime change. However, because occupation requires massive and sustained power projection, few states can credibly threaten such actions against others for most interstate crises. Further, according to The Correlates of War project, regime change is the primary issue for initiators of militarized interstate disputes less than 5% of the time.

In contrast, by the very nature of the conflicts, all civil wars involve some element of conquest. Occupation and power projection are not concerns. After all, a rebel governments intend to replace the existing institutions, while existing regimes seek to reconquer lost regions. Whether a resistance group or government intends to forgive past transgressions

²¹Of course, the previous section offered a fourth: fighting can be informative when there is a preponderance of military power.

²²On the other hand, in cases where the democracy would actually prefer expansion (as with the Mexican-American War), Proposition 2 predicts long conflicts, as q is high due to prior beliefs about democracies.

(as in South Africa) or execute the leadership (as in Libya) has a great impact on a party's willingness to settle. So does the extent to which each is willing to accept autonomy or independence in sectional civil wars. Unfortunately, the fighting process cannot effectively transmit that information. As Proposition 2 states, the informational issue would lead to longer conflicts with higher risk of military collapse in civil wars than in interstate wars.

That being said, our theory still has important implications for some interstate disputes. Many conflict scholars note that territorial wars tend to escalate and endure longer than the average dispute (Hensel 1996; Vasquez and Henehan 2001; Walter 2003; Huth 2009). Frequently, the cited mechanism is that territorial issues tend to be highly salient or have greater military consequences. Thus, "once a territorial dispute becomes a salient domestic political issue within the challenger, there are few political incentives for leaders to make concessions to settle a dispute" (Huth 2009, 145).

We do not wish to dispute this mechanism. However, our model produces a complimentary explanation. When states fight over territory, they may be uncertain of what their opponents will ultimately take in the event of military collapse. Fighting does not effectively communicate that information. In turn, they must continue engaging in costly confrontations longer than they would if the source of uncertainty regarded resolve or the likelihood of victory. Due to issues with power projection, the extent of the problem might be as great as in civil wars. Nevertheless, this helps explain variation in interstate conflict.

7 Conclusion

This paper has examined how information revealed through fighting may affect bargaining between combatants. Specifically, we study the role of incomplete information and limited consumption in this interaction. Contrary to previous research that looks at incomplete information over power or resolve, we show that fighting may not quickly resolve bargaining problems between states when uncertainty is about the extent to which one state wants to conquer another. Our results indicate that the existence of differential costs of delay are crucial for fighting to effectively reveal information.

Three policy implications naturally stem from the model. First, if invasion and territorial capture is a goal, policymakers should anticipate such conflicts to last longer. Second, the United States in particular ought to take this to heart. Since 2001, potential conquest has lingered in the wars the United States has been involved in—Afghanistan, Iraq, Libya, and

their associated civil wars and insurgencies. American policymakers need to be cognizant that short periods of fighting in these types of wars might not result in larger offers from the opposing party. Lastly, in the context of civil war, international peacekeeping missions might need to wait a long time before deploying. After all, if the peacekeepers are only there to ensure the long-term credibility of a treaty, it may be a while before the parties find a mutually agreeable solution.

Our work leaves open many avenues for future research. The model we present is a first-cut at interstate (or intrastate) negotiations with uncertainty over conquest. We imagine it as a reduced-form interaction, with the assumption that states' taste for conquest varies and that the precise preference may be unknown to its opponents. This was sufficient to derive the isomorphism and convergence failure results. However, open questions remain about how states reach their conquest decisions. Further research in that vein may prove fruitful.

8 Appendix

8.1 Proof of Proposition 1

There are only two possible equilibrium demands: $x \in \{p_A + c_B, p_A(1 - m) + m + c_B\}$. This is for the standard reasons. A demand strictly greater than both results in war against all types, but appeasing just the strongest type is a profitable deviation. A demand strictly less than both generates peace versus both types, but state A could profitably deviate by demanding the midpoint between that offer and $p_A + c_B$. Finally, a demand strictly between the two results in peace versus the moderate and war against the extremist, but demanding the midpoint between that demand and $p_A(1 - m) + m + c_B$ gives a greater payoff against the moderate and still induces war versus the extremist and is thus a profitable deviation.

As such, we investigate whether demanding $p_A + c_B$ and inducing both types to accept is better than demanding $p_A(1 - m) + m + c_B$ and fighting a war against just the extremist.²³ That calculation is as follows:

$$p_A + c_B > r[p_A(1 - m) + m + c_B] + (1 - r)(p_A - c_A)$$

²³For the standard reasons, for all one-shot bargaining proofs, we assume that the receiver accepts with probability 1 when indifferent. While this is without generality here, it fails in the game with bargaining while fighting.

$$r < \frac{c_A + c_B}{m(1 - p_A) + c_A + c_B}$$

By analogous argument, A demands $p_A(1 - m) + m + c_B$ if r is greater than that critical amount and is indifferent between the two iff r equals that critical value. \square

8.2 Proof of Remark 1

Similar to before, there are only two possible equilibrium demands: $x = p'_A + c_B$ and $x = p_A + c_B$.²⁴ If A demands $p'_A + c_B$, the weak type of B accepts and the strong type rejects. If A demands $p_A + c_B$, both types accept. As such, A prefers demanding $p_A + c_B$ if:

$$p_A + c_B > r(p'_A + c_B) + (1 - r)(p_A - c_A)$$

$$r < \frac{c_A + c_B}{p'_A - p_A + c_A + c_B}$$

Recall that the remark said that these two models are isomorphic using the substitution $p'_A = p_A + m(1 - p_A)$. There are two elements to check. First, note that if B is weak, it accepts if $x \leq p'_A + c_B$. Making the substitution, it accepts if $x \leq p_A + m(1 - p_A) + c_B$. This is identical to the case in the original model when B is a moderate. Second, the critical cutpoint here is $\frac{c_A + c_B}{p'_A - p_A + c_A + c_B}$. Making the substitution yields $\frac{c_A + c_B}{m(1 - p_A) + c_A + c_B}$, which is the original cutpoint. \square

8.3 Proof of Propositions 2 and 3

We group the proofs for these two propositions together because they follow the same general strategy. We proceed by exhausting possible equilibrium values of x_1 .

To begin, note that both types must accept $x_1 < 2p_A - p_A^2 + 2c_B - p_A c_B$ with probability 1. First, consider the extremist's decision. In the second period, the extremist can receive *at most* $1 - p_A - c_B$. This follows directly from Proposition 1—either A demands $p_A + c_B$ and the extremist accepts or A demands $p_A + m(1 - p_A) + c_B$. Regardless, the extremist earns $1 - p_A - c_B$. We can use this information to calculate the extremist's best continuation value for rejecting. With probability p_A , it loses in the first period and receives 0. With

²⁴An analogous argument from the proof for Proposition 1 rules out all other possible cases, so we omit proofs for them.

probability $1 - p_A$, it advances to the second stage and earns at most $1 - p_A - c_B$. Either way, it pays c_B . Consequently, its overall best possible payoff for rejecting equals:

$$p_A(0) + (1 - p_A)(1 - p_A - c_B) - c_B$$

$$1 - 2p_A + p_A^2 - 2c_B + p_Ac_B$$

In turn, any demand $x_1 < 2p_A - p_A^2 + 2c_B - p_Ac_B$ leaves strictly more leftover for the extremist than if it rejects and fights a battle. So the extremist must accept any such demand.

The logic follows analogously for the moderate type. The moderate can expect *at most* $p_A + c_B$ if it reaches the second period. Unlike the extremist, the moderate could receive strictly less than that if $r > r^*$ and A demands $p_A + m(1 - p_A) + c_B$. However, that point is moot—the calculation for the moderate’s most optimistic continuation value is identical to the extremist’s, so it too must accept any demand $x_1 < 2p_A - p_A^2 + 2c_B - p_Ac_B$.

On the other end of the spectrum, both types must reject $x_1 > 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_Ac_B$. This is because the moderate receives *at least* its war payoff of $1 - p_A - m(1 - p_A) - c_B$ in the second period. Working through the cost of and the probability of being eliminated through the first battle, the moderate’s continuation value for rejection is *at least*:

$$p_A(0) + (1 - p_A)[1 - p_A - m(1 - p_A) - c_B] - c_B$$

$$1 - 2p_A + p_A^2 - m(1 - p_A)^2 - 2c_B + p_Ac_B$$

In contrast, any demand $x_1 > 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_Ac_B$ leaves an amount strictly smaller than the moderate’s minimum continuation value. Therefore, the moderate must reject. But because the moderate type’s war payoff is strictly less than the extremist’s (since both win with the same probability and pay the same costs but the extremist earns an additional m if it wins), the extremist must reject as well.

This leaves values on the interval $[2p_A - p_A^2 + 2c_B - p_Ac_B, 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_Ac_B]$ as the only remaining possibilities. From here, we consider two divisions of the parameter space: $q < r^*$ and $q > r^*$. First, suppose $q < r^*$. Regardless of the offer, if the moderate type pools with the extremist, the second period subgame has $r < r^*$ and therefore A makes the “safe” demand of $p_A + c_B$ that both types accept. Note that this gives the moderate a payoff equivalent to the extremist’s war payoff. This means that both types accept iff

$x_1 = 2p_A - p_A^2 + 2c_B - p_{ACB}$. Thus, in the unique PBE, A offers that amount and both types accept; any other amount leads to unnecessary deadweight loss that ultimately comes out of A's payoff.

Second, suppose $q > r^*$. Consider any demand $x_1 \in (2p_A - p_A^2 + 2c_B - p_{ACB}, 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB})$. The extremist must reject with probability 1. This follows from the above result that the extremist earns *at least* $1 - 2p_A + p_A^2 - 2c_B + p_{ACB}$ but any such demand does not leave enough leftover to appease the extremist. If the moderate type separates from the extremist, the second period subgame has $r = 0 < r^*$. Per Proposition 1, A demands $p_R + c_B$. Note that if the moderate type were to deviate to rejecting, this would give it a payoff equivalent to the extremist's war payoff. However, the extremist rejects in the first period because its war payoff is greater than its payoff for accepting. But this in turn means that the moderate's payoff for accepting a value in the interior is less as well. So the moderate could profitably deviate, and thus it cannot separate in equilibrium.

Next, suppose the moderate pools with the extremist by rejecting. Then $r > r^*$, so state A demands $p_A + m(1 - p_A) + c_B$ in the second stage. This gives the moderate type its war payoff. However, the moderate strictly prefers accepting any $x_1 < 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}$ in the first stage because this is strictly greater than its war payoff. So pooling on rejecting cannot be a best response.

Finally, consider semi-separating strategies. The moderate's indifference condition requires its expected utility for rejecting to be equal to its expected utility for accepting. In the second period, only two demands are possible in equilibrium: $p_A + m(1 - p_A) + c_B$ and $p_A + c_B$. Note that the remainder the moderate receives for accepting $p_A + m(1 - p_A) + c_B$ is strictly less than its payoff for rejecting x_1 while the remainder the moderate receives for accepting $p_A + c_B$ is strictly greater than its payoff for rejecting x_1 . Thus, for the moderate to be indifferent between accepting and rejecting, state 1 must offer a convex combination of the two. For state 1 to mix between those offers, its posterior must equal r^* .

From here, it might seem that A would need to calculate a complicated expected utility function for all such demands $x_1 \in (2p_A - p_A^2 + 2c_B - p_{ACB}, 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB})$ and then optimize that function. However, *no* such demand is optimal. To understand why, recall that A is indifferent between demanding $p_A + m(1 - p_A) + c_B$ and $p_A + c_B$ when $r = r^*$. Therefore, its expected utility for the second period is a flat $p_A + c_B$. In turn, the probability of reaching the second period and state 1's payoff for that period is unchanging in the original demand x_1 on the interval.

Nevertheless, A could deviate to demanding the midpoint between that demand and $2p_A - p_A^2 + 2c_B - p_{ACB}$. Since that new value still falls in the interval, the parties continue playing under the same equilibrium strategies and A receives the same payoffs afterward. However, it keeps slightly more in the case where the moderate accepts. This is a profitable deviation. As such, no equilibrium involves a demand $x_1 \in (2p_A - p_A^2 + 2c_B - p_{ACB}, 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB})$.

This leaves two possibilities: $x_1 \in \{2p_A - p_A^2 + 2c_B - p_{ACB}, 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}\}$. In the latter case, for the standard reasons, no equilibrium can exist in which the extremist rejects with positive probability. Thus, we consider the case in which the extremist accepts with certainty. This gives state A a payoff of $2p_A - p_A^2 + 2c_B - p_{ACB}$ if it demands that much.

If state A demands $2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}$ instead, its expected utility equation is substantially more involved. For the reasons covered in the $x_1 \in (2p_A - p_A^2 + 2c_B - p_{ACB}, 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB})$ case, the moderate must mix while the extremist rejects. As such, state A earns a convex combination of $2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}$ (its demand in the first period that the moderate sometimes accepts) and its payoff if state B rejects.

The first step is to calculate the probability that state B accepts state A's initial offer. Recall that r must equal r^* . Let σ_R be the probability the moderate rejects. With the extremist rejecting with probability 1, we can calculate the mixed strategy that generates r^* as follows:

$$\frac{q\sigma_R}{q\sigma_R + (1 - q)(1)} = \frac{c_A + c_B}{m(1 - p_A) + c_A + c_B}$$

$$\sigma_R^* = \frac{(1 - q)(c_A + c_B)}{q[m(1 - p_A)]}$$

As such, demanding $2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}$ yields that value with probability $q(1 - \sigma_R^*)$.

The remaining portion of the time, state B rejects. Regardless of the outcome of the first battle, state A pays c_A . With probability p_A , state A wins the war decisively in the first battle and receives 1. With probability $1 - p_A$, B prevails and the game moves to the second stage of bargaining. Note that the posterior belief guarantees that state A is indifferent between demanding $p_A + c_B$ and $p_A + m(1 - p_A) + c_B$. Accordingly, we can calculate A's

expected payoff as its expected payoff for demanding $p_A + c_B$, which is simply $p_A + c_B$.

Overall, state A prefers demanding $2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}$ in the first period to demanding $2p_A - p_A^2 + 2c_B - p_{ACB}$ if:

$$q(1 - \sigma_R^*)(2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}) + [1 - q(1 - \sigma_R^*)][(p_A)(1) + (1 - p_A)(p_A + c_B) - c_A]$$

$$> 2p_A - p_A^2 + 2c_B - p_{ACB}$$

$$q > q^* \equiv \frac{(c_A + c_B)(c_A + c_B + m(2 - p_A)(1 - p_A))}{(c_A + c_B + m(1 - p_A)^2)(c_A + c_B + m(1 - p_A))}.$$

So if q is greater than q^* and r^* , state A demands $x_1 = 2p_A - p_A^2 + m(1 - p_A)^2 + 2c_B - p_{ACB}$. The extremist rejects with certainty while the moderate rejects with probability σ_R . In the second period, state A demands $x_2 = p_A + m(1 - p_A) + c_B$. The extremist rejects and the moderate accepts. If q is less than q^* , state A demands $2p_A - p_A^2 + 2c_B - p_{ACB}$ and both types accept. \square

References

- Acemoglu, Daron, Thierry Verdier and James A Robinson. 2004. "Kleptocracy and Divide-and-Rule: A Model of Personal Rule." *Journal of the European Economic Association* 2(2-3):162–192.
- Arena, Philip. 2015a. "Costly signaling, resolve, and martial effectiveness." *Working paper*.
- Arena, Philip. 2015b. "Crisis bargaining, domestic opposition, and tragic wars." *Journal of Theoretical Politics* 27(1):108–131.
- Arreguin-Toft, Ivan. 2001. "How the weak win wars: A theory of asymmetric conflict." *International Security* 26(1):93–128.
- Blainey, Geoffrey. 1988. *Causes of War*. New York: Simon and Schuster.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph M Siverson and James D Morrow. 2003. *The logic of political survival*. Cambridge: MIT Press.

- Chiba, Daina and William Reed. 2014. "Peaceful Territorial Transfers & Third Party Conflict." *Working paper* .
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(03):577–592.
- 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.
- Fearon, James D. 2007. Fighting rather than bargaining. In *Annual Meetings of the American Political Science Association*.
- Fey, Mark. 2014. "Learning to fight: Information, uncertainty, and the risk of war." *Working paper* .
- Fey, Mark and Kristopher W Ramsay. 2011. "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict." *American Journal of Political Science* 55(1):149–169.
- Filson, Darren and Suzanne Werner. 2002. "A Bargaining Model of War and Peace: Anticipating the Onset, Duration, and Outcome of War." *American Journal of Political Science* pp. 819–837.
- Goemans, Hein Erich. 2000. *War and punishment: The causes of war termination and the First World War*. Princeton Princeton University Press.
- Grant, Ulysses S. 2006. *Personal Memoirs of Ulysses S. Grant*. Cosimo, Inc.
- Henderson, Timothy J. 2007. *A glorious defeat: Mexico and its war with the United States*. New York: Hill and Wang.
- Hensel, Paul R. 1996. "Charting a course to conflict: Territorial issues and interstate conflict, 1816-1992." *Conflict Management and Peace Science* 15(1):43–73.
- Herbst, Jeffrey. 2014. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton: Princeton University Press.

- Hollander, Ethan J. 2008. "The Final Solution in Bulgaria and Romania: A Comparative Perspective." *East European Politics & Societies* 22(2):203–248.
- Huth, Paul. 2009. *Standing your ground: Territorial disputes and international conflict*. University of Michigan Press.
- Huth, Paul K, Sarah E Croco and Benjamin J Appel. 2011. "Does international law promote the peaceful settlement of international disputes? Evidence from the study of territorial conflicts since 1945." *American Political Science Review* 105(02):415–436.
- Iida, Keisuke. 1993. "When and how do domestic constraints matter? Two-level games with uncertainty." *Journal of Conflict Resolution* 37(3):403–426.
- Ikenberry, G John. 2009. *After victory: Institutions, strategic restraint, and the rebuilding of order after major wars*. Princeton: Princeton University Press.
- Kedziora, Jeremy. 2012. "Endogenous War Aims and State Resolve." *Working paper* .
- Kenkel, Brenton. 2013. "Dividing the conquered." *Working paper* .
- Keohane, Robert Owen and Joseph S Nye. 1977. *Power and interdependence: World politics in transition*. Little, Brown Boston.
- Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The rational design of international institutions." *International Organization* 55(04):761–799.
- Kydd, Andrew H. 2010. "Rationalist approaches to conflict prevention and resolution." *Annual Review of Political Science* 13:101–121.
- Liberman, Peter. 1998. *Does conquest pay?: the exploitation of occupied industrial societies*. Princeton University Press.
- Mueller, John. 1989. *Retreat from doomsday: The obsolescence of modern war*. New York: Basic Books.
- Naimark, Norman M. 1995. *The Russians in Germany: a history of the Soviet Zone of occupation, 1945-1949*. Harvard University Press.

- Powell, Robert. 2004. "Bargaining and learning while fighting." *American Journal of Political Science* 48(2):344–361.
- Powell, Robert. 2006. "War as a commitment problem." *International organization* 60(01):169–203.
- Reed, William. 2003. "Information, power, and war." *American Political Science Review* 97(04):633–641.
- Reiter, Dan. 2009. *How wars end*. Princeton University Press.
- Rider, Toby J. 2013. "Uncertainty, salient stakes, and the causes of conventional arms races." *International Studies Quarterly* 57(3):580–591.
- Schultz, Kenneth and Hein Goemans. 2014. "Aims, Claims, and the Bargaining Model of War." *Working paper* .
- Scott, James C. 2008. *Weapons of the weak: Everyday forms of peasant resistance*. New Haven: Yale University Press.
- Siverson, Randolph M and Harvey Starr. 1990. "Opportunity, willingness, and the diffusion of war." *American Political Science Review* 84(01):47–67.
- Slantchev, Branislav L. 2003. "The Principle of Convergence in Wartime Negotiations." *American Political Science Review* 97(04):621–632.
- Slantchev, Branislav L. 2004. "How initiators end their wars: The duration of warfare and the terms of peace." *American Journal of Political Science* 48(4):813–829.
- Smith, Alastair and Allan C Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48(6):783–813.
- Spolaore, Enrico and Alberto Alesina. 2005. *The size of nations*. Cambridge: MIT Press.
- Stone, Randall W. 2002. *Satellites and Commissars: Strategy and Conflict in the Politics of Soviet-Bloc Trade*. Princeton: Princeton University Press.
- Tarar, Ahmer and Bahar Leventoglu. 2009. "Public commitment in crisis bargaining." *International Studies Quarterly* 53(3):817–839.

- Turner, Henry Ashby. 1987. *The two Germanies since 1945*. New Haven: Yale University Press.
- Tutorow, Norman E. 1978. *Texas annexation and the Mexican War: a political study of the Old Northwest*. Palo Alto: Chadwick House.
- Van Creveld, Martin. 2004. *Supplying war: logistics from Wallenstein to Patton*. Cambridge: Cambridge University Press.
- Vasquez, John and Marie T Henehan. 2001. "Territorial disputes and the probability of war, 1816-1992." *Journal of Peace Research* 38(2):123–138.
- Walter, Barbara F. 1997. "The critical barrier to civil war settlement." *International organization* 51(03):335–364.
- Walter, Barbara F. 2003. "Explaining the Intractability of Territorial Conflict1." *International Studies Review* 5(4):137–153.
- Walter, Barbara F. 2009. "Bargaining failures and civil war." *Annual Review of Political Science* 12:243–261.
- Wheelan, Joseph. 2007. *Invading Mexico: America's continental dream and the Mexican War, 1846-1848*. New York: Carroll and Graf Publishers.
- Wolford, Scott. 2007. "The turnover trap: new leaders, reputation, and international conflict." *American Journal of Political Science* 51(4):772–788.
- Wolford, Scott, Dan Reiter and Clifford J Carrubba. 2011. "Information, commitment, and war." *Journal of Conflict Resolution* 55(4):556–579.
- Zacher, Mark W. 2001. "The territorial integrity norm: International boundaries and the use of force." *International Organization* 55(02):215–250.