Fanatical Peace: How Fundamental Disagreements Can Discourage Conflict

Peter Bils
Vanderbilt University

Richard Jordan
Baylor University

Kristopher Ramsay
Princeton University

Abstract

It is generally held that a leader's optimism about the likely outcome of a war can be a cause of war. When two rivals disagree about what they can expect to win by force, both might prefer to fight rather than settle a dispute. This explanation is usually referred to as war resulting from "mutual optimism." Conventionally, the literature suggests that the more extreme this mutual optimism grows—the wider the gap between leaders' assessments—the more likely is conflict. We show that there is no necessary, logical reason to believe that divergent beliefs tend to make conflict more likely. In a dynamic bargaining model, we find that increasing leaders' mutual optimism can *prevent* war. In fact, conflict can be more likely between actors with similar underlying beliefs than between actors whose beliefs are radically different. We call this fanatical peace.

1 Motivation and Argument

Some argue countries go to war because they believe the eventual outcome of a war will leave them better off than the peace agreement that is available at the time (Blainey, 1973; Fearon, 1995). The "principle of convergence" establishes that, during war, battlefield outcomes inform the players' beliefs and over time, once beliefs converge enough, players can conclude the war (Slantchev, 2003). In other words, battlefield outcomes are a costly method for changing people's minds. Leaders can use war as a tool to teach adversaries they are too optimistic about eventual success.

This observation implies the cost of war is endogenous to players' beliefs about victory. Their expectations about how beliefs will change over time—and, subsequently, about the prospective costs of war—will likewise feed back into peacetime and wartime bargaining. As a result, there are two competing forces influencing the likelihood of war or peace. When both sides are optimistic about the outcome of a campaign, they may prefer war to a settlement their opponent will accept; when this happens, we say that mutual optimism has caused conflict. At the same time, a very optimistic — or "fanatical" — opponent will need to see several adverse battlefield outcomes to agree to a peace, and such a war will be very costly. So while mutual optimism may increase the naive value of war relative to peace, a bit more thought reveals that a war against a fanatical opponent will also be more costly to pursue. This logic favors a settlement. Increasing the gap between players' beliefs thus has an ambiguous net effect.

In this paper, we show that war, as a teaching tool, is most useful when leaders' beliefs are relatively similar. The more their beliefs diverge, the less useful this tool becomes. Sometimes, it is too costly to teach an opponent they are wrong; in this case, differences in leaders' beliefs can cause peace. By introducing initial worldviews into a rationalist model, we show how a situation of entrenched, nonviolent disagreement can occur. We call this phenomenon *fanatical peace*.

We believe this prediction sometimes describes reality. 'Clashes of civilization' trace wars along fundamental cultural and religious divides (Huntington, 1993). Yet, intrastate conflicts have consistently outnumbered interstate conflicts for the past half-century, often by a factor of four or more (Davenport and Gates, 2014). For whatever reason, the deepest fissures seem to be *within* civilizations. Edward Said makes the point most effectively: "Huntington," he observes, has no "time to spare for the internal dynamics and plurality of every civilization, or for the fact that the major contest in most modern cultures concerns the definition or interpretation of each culture."[1] We believe that "fanatical peace" can account for some

[1] Said (2001). See also Fox (2002, 433): "it is minorities which are more culturally similar to the majority

of intrastate war's relative prevalence. Likewise, the worst religious wars are often fought between schisms of the same confession—against heretics, not infidels. The Third Crusade offers perhaps the most famous example, when the Byzantines aligned with Saladin against the Holy Roman Emperor.[2]

We are not trying here to offer a single model for all wars. Rather, we hope to contribute to an expanding stable of explanations. As with the Third Crusade, we believe there are important wars and non-wars that current models cannot explain, but that can be explained by the logic of "fanatical peace." After presenting the model, we will offer some examples—of wars that were fought because an enemy was "teachable" and that were *not* fought because an enemy's beliefs were too extreme to try to alter.

We aim to make three contributions to the literature. First, we show why the relationship between beliefs and the likelihood of war is not as simple as is sometimes assumed: there is a clear theoretical reason that increasingly divergent beliefs can decrease the incentive to fight. Second, we offer an explanation of an underappreciated phenomenon: conflict is sometimes more likely between groups who (mostly) agree than between groups with little in common. Third, we use a rationalist framework while incorporating a non-Bayesian foundation for leaders' initial beliefs. This is a step toward a rationalist/ideational synthesis. Finally, while our analysis is rooted in bargaining models of war, the broader logic may also apply to domestic politics, especially to partisanship and polarization.

1.1 Sources of Divergent Beliefs

To study how divergent beliefs cause war or peace, we first need to discuss how beliefs come to differ. Recent scholarship distinguishes three mechanisms. The first is *overconfidence*.

A simple causal chain links overconfidence to war: cognitive biases cause leaders to have positive illusions; these illusions cause leaders to overestimate their odds of victory; as these illusions grow more severe, the range of possible bargains shrinks, making a peaceful compromise more difficult. Because it depends on leaders interpreting information inaccurately (not according to Bayes's Rule), overconfidence is a non-rationalist explanation for war.[3]

Dominic Johnson offers the most consistent recent work on overconfidence (Johnson et al., 2006; Johnson and Tierney, 2011). Johnson links war to positive illusions through the over-

groups in their state which opt more often for the violent alternative or rebellion as opposed to the more peaceful option of political protest."

[2] "The early surrender of Jerusalem was brought about by the knowledge that, out of hatred for the Latins, the Greek Orthodox residents...were ready to betray the city" (Brand, 1962, 170). For recent work on other "frenemies," see Haas (2022).

[3] The literature following Jervis (1976) is too vast to canvass here. For recent work linking first-image beliefs to conflict, see Mercer (2013), Horowitz and Stam (2014), Holmes (2015), and Duelfer (2011).

estimation of one's own capability, the underestimation of a rival's, and the persistent neglect of intelligence. The combination of these factors implies that the "sum of opponents' winning estimates > 1" (Johnson, 2004, 36). Johnson and Fowler (2011) go on to demonstrate that overconfidence is evolutionarily desirable, such that overconfident individuals will tend to multiply while unbiased individuals will tend to be eliminated from a political ecosystem. These results seem confirmed by experimental studies (Sheffer et al., 2018) which find that policymakers, if anything, are *more* irrational than their citizens. They suggest that we should expect a population of overconfident, overly optimistic politicians, and so we should also not be surprised at the frequency of their conflicts.[4]

A second mechanism causing beliefs to diverge is *private information*, which can lead to mutual optimism between the parties. Unlike overconfidence, private information offers a rationalist explanation for war. Its logic is straightforward: actors' private information leads them to different expectations about the likely outcome of a war. If both players "expect to win," then neither side will back down.[5] Like the literature on overconfidence, work on private information explains how divergent beliefs are related to the likelihood of conflict (Wittman, 1979; Filson and Werner, 2002; Slantchev, 2003), and the claim that belief polarization leads to more war is repeated in standard IR textbooks.[6] Some scholarship in this vein has uncovered a nonmonotonic relationship between uncertainty and war, so that increasing uncertainty can in fact prevent conflict (Arena and Wolford, 2012; Debs and Weiss, 2016; Joseph, 2021). Nonetheless, this scholarship sees a purely monotonic relationship between the likelihood of war and the gap separating actors' beliefs. To our knowledge, there is no crisis bargaining model that demonstrates how increasing the distance between players' beliefs can reduce the likelihood of conflict.

We suggest these two mechanisms, overconfidence and private information, are missing

[4]See also Kahneman and Renshon (2007); Little and Zeitzoff (2017). Johnson (2020) further argues that these traits are widely beneficial in foreign policy. Johnson and Fowler (2011) assume that increasing overconfidence leads to conflict. Importantly, in their evolutionary model, there is no possibility for the equivalent of limited war—that is, a war which might communicate information without resulting in the death of an actor. It is this possibility which, in our model, will link increasing confidence to peace.

[5]Of course, both players do not need to assess their probabilities of victory as more than 50%. It will suffice for the sum of their estimates to exceed 1. Note that third parties can exacerbate this dynamic (Bas and Schub, 2016). Importantly, the theoretical consistency of this explanation is contested. Fey and Ramsay (2007) argue that another player's willingness to fight should itself communicate information about their private signals, thus leading the sum of players' beliefs to converge to 1, so that no player would ever want to fight an opponent who was willing to fight him. Slantchev and Tarar (2011) find this objection unpersuasive. Other scholars, such as Debs (2022), find it compelling, and so argue that uncertainty must be multidimensional for mutual optimism to explain war. We discuss how our results relate to this debate in the concluding section.

[6]Two examples: "[w]ar is more likely when states fall prey to false optimism about its outcome" (Van Evera, 1999, 14); "the underestimation of adversary capabilities...increases the probability of war." (Levy and Thompson, 2010, 136).

an important kind of belief. Explanations drawing on non-rational beliefs tend to emphasize the non-rational mechanism whereby they cause war, i.e. biased or inaccurate inferences. On the other hand, rationalist explanations tend to emphasize only the rational origins of beliefs, i.e. private information. Both traditions thus neglect a third, important way that beliefs lead to conflict, for there are non-rational beliefs that *rationally* cause war. While scientific or technical knowledge might be easily incorporated into a Bayesian belief model, humans have many other sources of beliefs, such as religious teachings or cultural upbringing.[7] As Blainey (1973, p. 53-54) argues in his discussion of the dreams and delusions of coming wars:

> Optimism does not arise from a mathematical assessment. It does not simply represent one nation's careful calculation that its military and economic capacities exceed those of the potential enemy. . . . The process by which nations evade reality is complicated. Patriotism, national languages and a sense of a nation's history are all dark glasses.

If rational people start from fundamentally different ideas about how the world works, then it makes sense that, even after observing the same new information, they may still hold different beliefs. Like overconfidence or private information, this third kind of disagreement can also cause conflict. It hinges on worldviews.

1.2 Worldviews

There is no universally-endorsed, rigorous definition of "worldview." The idea dates to Immanuel Kant, was popularized in the 19th century, and has now spread throughout the social sciences and humanities as an object of study. The word's usage usually falls between two poles.

In its simplest sense, a worldview refers to ideas that are in some sense large or encompassing, organize secondary notions about causes and effects, and perhaps give life meaning. On this understanding, a worldview is a kind of necessary cognitive architecture for processing new information and acting upon it.

At the other extreme, a worldview might be a kind of coping mechanism or even willful stupidity. These accounts of worldview tend to portray it as something that distorts human

[7] Recently, Allan (2018) argues that a scientific cosmology is itself a laden worldview, which shapes (and potentially distorts) the ways new information is interpreted. In this, he is building on Emanuel Adler (1992)'s work on epistemic communities and their role in shaping how new ideas, including seemingly neutral scientific ones, are incorporated into policymakers' belief structures. If Allan is right, then we have the more reason to study how fundamental worldviews shape even otherwise-rational Bayesian agents. Moreover, it is worth noting that, while we will assume that beliefs can be updated obviously in response to new information according to Bayes's Rule, in reality this is rarely the case—in which case, a worldview's very plasticity might make it *more* effective when assimilating new information (Black, 1983).

judgment—as something that makes people behave irrationally. For instance, it might serve as a crutch "when the unexpected undermines or overturns our most respected theories" (Katzenstein, 2022, 3). A long tradition in political psychology studies worldviews and ideologies not just as necessary cognitive heuristics, but as a kind of motivated reasoning or an effort to "justify the system" (Jost, Banaji and Nosek, 2004). Constructivists argue that wars and enmities arise from underlying conceptions of reality—what Wendt calls cultures of anarchy—rather than from the material distribution of capabilities (Wendt, 1999). Feminist IR scholars go further still. Many have argued that studying worldviews may be incompatible with the standard "Enlightenment" tools of political science (Tickner, 2009, 2014). These scholars suggest that IR theory has rarely taken worldviews like religion seriously because so much of IR theory is built on rationalist foundations. They argue that worldviews should be studied using "dialogic" methods rather than "rationalist" ones.

This is a serious challenge. Can something as seemingly nonrational as a worldview be incorporated into a rationalist account of politics? We believe it can. In this paper, we will model how worldviews can impact rational beliefs without undermining the rationality of those beliefs. We will also show that the *lack* of a shared worldview can mollify conflict rather than inflame it.

To do so, we return to a more Kantian notion of worldview as something "supersensible" (Kant, 1983, 111). On this understanding, a worldview is perhaps *pre*rational but not therefore *ir*rational. Facts do not interpret themselves. Individuals, when presented with the same information, will often interpret that information differently. It is tempting to attribute such differences to bias, but this conclusion is too hasty: "rational intuition" is not a contradiction (Holmes and Traven, 2015), and rational Bayesians can hold different posterior beliefs after observing the same new information. As Smith and Stam (2004, 785) point out, Marxist, liberal, and realist academics "interpret the same political events very differently." Presumably the same holds for Marxist and liberal policymakers.

Philip II offers an instructive example. Philip pursued a policy of "messianic imperialism" throughout his reign. His Catholic faith informed this policy and, indeed, his interpretation of almost all political events of the period.[8] In light of the setbacks that bedeviled his rule, it might seem like such a man must have been suffering from premature cognitive closure or an unwillingness to consider new evidence, and yet the opposite is the case—if anything, he operated on "information overload." He enjoyed reports of such "superior quality" that he often knew his enemies' secrets before they did; and, far from seeking irrational closure,

[8]For instance, Philip interpreted the victory at Lepanto and the near-simultaneous birth of an heir as more than mere coincidence. "In Philip II's world everything was felt to have a direct cause. Everyone therefore searched for the hand of Providence in human affairs and tried to discern and implement God's purpose" (Parker, 1998, 104).

he "constantly sought to update and extend his personal knowledge" (Parker, 1998, xv, 58). Geoffrey Parker puts the problem this way: "to understand what seemed rational to early modern statesmen and strategists, their mental universe and assumptions must enter the equation" (Parker, 1998, xviii). We should understand many of Philip's policies, including spectacular failures like the Spanish Armada, not as the irrational impulses of a zealot, but as the more-or-less rational policies of a well-informed, meticulous, and thoughtful statesman who began from very different premises.

Different worldviews can lead otherwise-rational leaders to different (but still rational) beliefs. Blainey points out that these beliefs can include the likely outcomes of war. He observes that, in 1914, both German and British leaders expected to win a short war. Their expectations do not seem to have been driven by private information. Rather, their incompatible beliefs seem to have been driven by their underlying views of the world: Britain, naturally, thought that economics would determine the pace and victor of the war, and so a rapid economic collapse of its enemies would result in a swift British victory; the technophilic Germans thought that military technology would be decisive, and so likewise expected a swift German victory (Blainey, 1973, 40). Neither belief was necessarily irrational. Later, Smith and Stam (2004) took Blainey's insight further and applied it to Austria and Prussia in the 1870s: their leaders observed the same facts (notably, the American Civil War), yet they reached different conclusions because they had different underlying theories about how the world, and war, worked.

We will model worldviews more broadly than Smith and Stam and more narrowly than Goldstein and Keohane (1993). Since this paper is only one step toward a rational/ideational synthesis, we will confine our analysis to the ways worldviews might shape leaders' expectations about victory in war. Of course, worldviews implicate far more than estimates of battlefield victory. Yet they shape even these quotidian beliefs in dramatic ways, and their effects bear on the likelihood of war or peace.

Two dramatic examples may help illustrate the point. In medieval Europe, religious beliefs about divine justice shaped monarchs' expectations about the outcomes of war. Indeed, James Turner Johnson (2015, 62-63) notes that wars between Christians were especially difficult to prevent *because* of their shared beliefs about the ways divine justice would influence a war's outcome. (Note that Johnson anticipates our main claim in this paper.) In the modern world, Soviet teleology shaped their premiers' beliefs, including their strategic assessments about the likely outcomes of major wars and insurrections. Clearly, these beliefs are not the product of private information acquired about the true fate of the universe; just as clearly, these beliefs will shape the way otherwise-rational actors will interpret new information, including mundane information like news stories and casualty reports.

1.3 Modeling Worldviews

In what follows, we model different initial worldviews using heterogeneous prior beliefs (Galperti, 2019). Specifically, the actors disagree about the probability each wins a decisive conflict, and this disagreement is common knowledge. Following Harsanyi (1967), it is standard to assume players share a common-prior belief, and so differences in beliefs only arise through private information. However, this assumption has been shown to have a number of theoretical and empirical limitations (Morris, 1995; Yildiz, 2011).[9] Correspondingly, an extensive literature has relaxed the common-priors assumption across a variety of contexts.[10]

In our model, the divergence in prior beliefs is explicitly *not* about differences in private information. As outlined above, examples might include religious, ideological, or philosophical beliefs/intuitions that structure the way new information is interpreted. It is these different kinds of underlying beliefs that we capture with non-common priors. They allow us to model learning when there is no mechanism (within the model) tying the initial beliefs to the true state of the world. In short, they allow us to model divergent beliefs which *precede* any sort of private information.[11]

Like Hirsch (2016) and Smith and Stam (2004), however, we are not discarding rationality, nor even Bayesian learning. Even if players start with commonly-known differences in initial expectations, that doesn't mean they cannot learn.[12] Several articles have taken this approach to studying learning and strategic action in the face of differing initial worldviews (Che and Kartik, 2009; Van den Steen, 2004, 2009), and we do the same. We also show that, even given these widely-divergent worldviews, leaders will still gradually converge to similar beliefs about their relative power.

Heterogeneous prior beliefs are particularly important for studying conflict. Different groups may hold different worldviews, and these worldviews may shape their beliefs about the efficacy of violence and their own likelihood of success. As well, leaders' own worldviews

[9] Aumann (1976) shows that people with common-prior beliefs cannot publicly agree to disagree, and suggests this is evidence against the assumption given "... there are in fact people who respect each other's opinions and nevertheless disagree heartily about subjective probabilities."

[10] A non-exhaustive list of applications includes: financial trading (Morris, 1996); persuasion (Che and Kartik, 2009; Alonso and Câmara, 2016); political hierarchies (Hirsch, 2016); bargaining between political parties (Simsek and Yildiz, 2016); pre-trial negotiations (Watanabe, 2006; Vasserman and Yildiz, 2019); and electoral competition (Izzo, 2023).

[11] In our simple model, we model different worldviews simply as different priors over a single parameter. More complex models might allow states to have more complex worldviews, for instance, different beliefs about underlying data-generating functions. The appendix includes an example of fanatical peace under such complex beliefs.

[12] In economics, an established literature studies the "endogenous uncertainty" that arises when actors have heterogenous priors in a market. This agenda gained traction, in part, because of a growing recognition that actors could interpret the same information differently (Kurz, 1996).

may create biases that lead them into conflicts with overconfidence. This overconfidence could arise by chance, but it might also be endogenous: as we discuss in the conclusion, in our model citizens and states would be incentivized to choose leaders who are overly optimistic about their probability of winning. Finally, while states may be able to draw analogies to previous cases (Khong, 1992), in general new crises are novel situations. As such, there is no reason to assume that states' or leaders' beliefs have already converged at the start of the conflict.[13] Of course, during the conflict both states will update their beliefs, and in the long-run at least their beliefs about their respective odds of success will converge (though their underlying worldviews might not); our model reflects this dynamic.[14]

2 Contribution to the Literature

Despite the importance of worldviews to conflict, to our knowledge only one other crisis bargaining model assumes that actors might fundamentally disagree but still learn. Smith and Stam (2004) assume that states have noncommon priors—that, beneath everything, they disagree about the basic possibilities which govern the world.[15] This assumption allows them to study how different worldviews shape the way beliefs converge during wartime and peacetime. They show how, during peace, actors who witness the same information can draw opposite conclusions. The gap between these beliefs will only close as war compels a more accurate view of reality.[16] In our model, we are less concerned about the origins of actors' different worldviews and more concerned about the way they reshape actors' incentives to bargain. We will thus pay less attention than Smith and Stam (2004) to the postwar implications of long- versus short-wars, and more to the prospects for immediate rather than delayed settlements. Our goal is to compare the effects of different prior beliefs on the duration of war, not to characterize the process of learning within that war, or the effect of that learning process on future conflicts.

[13] Morrow (2014) argues that international law may play a vital role in helping states coordinate their strategic expectations about intrawar behavior, but the laws of war may be less helpful in leading expectations about victory to converge.

[14] In fact, the entire field of Bayesian statistics is based on the fact that, with enough data, the likelihood eventually swamps the prior in the presence of data.

[15] Note that Powell (2004*a*) and Conde and Obiols (2006), among others, label Smith and Stam's players as 'irrational" because of their noncommon priors.

[16] An important difference between Smith and Stam (2004) and other informational models of war is the role of fighting: in the typical approach, exemplified in Slantchev (2003), information is primarily conveyed by states' willingness to fight, not by the outcomes of battles themselves; by contrast, in Smith and Stam's approach, states already know their rivals' resolve and self-assessments, and so information is primarily conveyed by objective battlefield feedback. To keep things simple, our approach here tracks Smith and Stam, but we believe future research may produce even more dramatic "fanatical peace" outcomes when studying the information conveyed by states' willingness to fight.

In the broader bargaining literature, Muhamet Yildiz links increasing optimism to more efficient negotiations. Yildiz (2003) finds that sufficiently-optimistic beliefs can enable players to reach an immediate agreement, but his model does not allow for non-trivial learning while bargaining. In subsequent work, Yildiz (2004) allows non-trivial learning when players disagree about their recognition probabilities. Similar to our result here, he shows that players either immediately agree because initial optimism is sufficiently large, or they delay agreement until optimism is sufficiently small. Unlike Yildiz (2004), the learning process we study allows disagreement to widen or narrow after battles; as a result, the time to an agreement is stochastic rather than fixed in our model, and along the equilibrium path extreme optimism can induce an agreement even after fighting begins. Furthermore, we model uncertainty over players' outside options (not recognition probabilities), and we characterize how features of our specific environment, such as the costs of war, influence the duration of conflict.

As previously mentioned, differences in beliefs about power are typically modeled by assuming states have private information about their capabilities. The closest papers in this tradition to ours are those that incorporate this private information into a model of bargaining while fighting (Powell, 2004*b*; Fearon and Jin, 2021). In these models, the proposing state initially makes a low offer that then becomes more generous over time.[17] Thus, the strongest types of the state receiving offers fights the longest before settling.[18] This contrasts with our argument, where the conditions for fanatical peace are easiest to sustain when the state receiving offers believes it is strong and there is significant heterogeneity in beliefs.[19]

Alternatively, a state may have private information about which action it will take if it wins a conflict, rather than about its probability of victory. In this setting, a state that takes a more antagonistic action may be thought of as more fanatic. However, as in the case of uncertainty over power, when there is bargaining while fighting conflict is most likely against the more "fanatical" type (Spaniel and Bils, 2018).[20]

Because our model has no private information, states learn differently than in most models of mutual optimism. There, states learn both from a battle's outcome (as in our model) and from the behavior of other players, since their behavior reveals something about their private

[17]Similar results hold in the broader literature on bargaining with interdependent values (Deneckere and Liang, 2006).

[18]Furthermore, conflict in these models can last a long time (Baliga and Sjöström, 2023).

[19]Likewise, this distinguishes our results from complete information models of conflict which show that greater heterogeneity in states' discount factors leads to conflict (Chadefaux, 2011; Spaniel, Bils and Judd, 2020).

[20]With this type of uncertainty, communication can help avoid conflict (Bils and Spaniel, 2017; Haynes and Yoder, 2020). This is a substantively different mechanism from fanatical peace. Moreover, conflict is avoided against the type that has preferences closer to the proposing state and the type with more divergent preferences (Bils and Spaniel, 2017).

information. Our model lacks this second kind of learning. Since players' prior beliefs are commonly-known, each player is able to predict perfectly how the other will behave, and so they learn nothing from observing it.

This difference especially distinguishes our model from Fey and Ramsay (2007). In their model, an opponent's willingness to fight compels an actor to reconsider his own willingness to do so. (If your opponent has private information that makes him willing to fight, and you observe that willingness, then you can infer something about his private information.) According to Fey and Ramsay, iterating this logic will eventually eliminate any mutual optimism derived from private information. In our model, players learn nothing from each other's willingness to fight, and so mutual optimism (which is derived from their divergent worldviews, not from private information) can persist across many rounds of interaction.[21]

By considering worldviews rather than private information, we can understand how conflicts between civilizations might sometimes be less likely than clashes inside them. The objectivity of war drives beliefs to converge, but this process takes time. In particular, the more worldviews diverge, the longer it will take for beliefs to converge. It is much easier to change someone's mind when they agree with you on fundamentals. As a tool to change someone's mind, war is therefore most useful when fighting against people you mostly agree with.

3 Model

To illustrate the logic of this mechanism, we introduce a model of bargaining in which states may learn about their relative power by fighting. Two countries, A and B, are in a dispute over a good of value 1 and interact over T periods, with $T \in \mathbb{N}_1$.[22]

Bargaining and fighting. At the start of each period $t \in \{1, 2..., T\}$, state A proposes a division of the good, denoted $x \in [0, 1]$. If state B accepts, then the game ends with state A receiving x of the good and state B obtaining $1 - x$. On the other hand, if state B rejects, then the states fight a battle. With probability $k \in (0, 1)$ the battle is decisive. In this case, the victor of the decisive battle receives the entire good and the game ends. With probability

[21] Slantchev and Tarar (2011) fundamentally agree with Fey and Ramsay (2007) that mutual optimism comes from private information, but disagree about how to model the strategic interaction. They present a bargaining model with one-sided incomplete information that has configurations of priors and realization of types that lead to war. Since there is only one-sided incomplete information, there are not two priors to compare, common or otherwise.

[22] In the Appendix, we show that our main results also hold if there is an infinite horizon ($T = \infty$). In this case, the game ends only after a decisive battle or the states reach an agreement.

$1-k$ the battle is not decisive, the states observe the victor, and the game moves to the next period.[23]

The victor of the battle depends on $\omega \in \{A, B\}$. If the battle is decisive, then the victor is simply state ω. If the battle is not decisive, then state ω wins the battle with probability $\sigma \in (1/2, 1)$. Thus, state ω is more likely to be the victor, but the battlefield outcome does not perfectly reflect who would win in a decisive battle. We denote the victor of a non-decisive battle in period t by $s_t \in \{A, B\}$, hence, $Pr(s_t = A|\omega = A) = Pr(s_t = B|\omega = B) = \sigma$.

Information and beliefs. States are uncertain about the true value of ω. Importantly, states do not share the same prior belief about ω. At the start of the game, state A's prior that it will be victorious in a decisive battle is given by p_1^A, i.e., it believes that $Pr(\omega = A) = p_1^A$. Instead, state B's prior belief is that $\omega = A$ with probability p_1^B. We assume that $p_1^B < p_1^A$, thus, each state is relatively more optimistic about its prospects in a decisive battle. Because the outcome, s_t, of a non-decisive battle is informative about the eventual winner in a decisive conflict, over time the states update their beliefs about ω. We assume states use Bayes's rule to update their beliefs. Define p_t^i as the probability that player i assigns to the event $\omega = A$ at the start of period t.

Payoffs. If the states agree to a division of the good $(x, 1-x)$, then state A receives a payoff of x and state B a payoff of $1-x$. If there is a decisive battle, then the victor claims the entire good and receives a payoff of 1, while the loser gets a payoff of 0. Furthermore, in each period that the states fight, state i incurs a cost $c_i > 0$, and we assume $c_A + c_B < 1$. Thus, if the states settle in period $s \leq T$ after fighting in every previous period, then state A's receives the payoff $x - \sum_{t=1}^{s-1} c_A$ and state B the payoff $1 - x - \sum_{t=1}^{s-1} c_B$.

If the states reach period T and still do not come to an agreement, then the game ends and assigns a payoff of $p_T^A - \frac{c_A}{k}$ to state A and a payoff of state $1 - p_T^B - \frac{c_B}{k}$ to player B. That is, the states get their expected payoffs from fighting until one state is victorious. This is consistent with the standard crisis bargaining set-up in which failure to reach an agreement results in war modeled as a costly lottery. Thus, if the states fight in every period then state A's payoff is $p_T^A - \frac{c_A}{k} - \sum_{t=1}^{T-1} c_A$, and state B's payoff is $1 - p_T^B - \frac{c_B}{k} - \sum_{t=1}^{T-1} c_B$.

Hence, if, for example, the states fight for the first two periods and then settle in period 3 with division $(x, 1-x)$, then state A's expected utility is:

$$k \cdot p_1^A - c_A + (1-k)\Big(k \cdot \mathbb{E}_A[p_2^A] - c_A + (1-k) \cdot x\Big).$$

[23]Thus, as in Wagner (2000), the states face uncertainty about whether the game will reach the terminal period.

Fanaticism and updating. A state's belief p^i is crucial for its expectations about battlefield outcomes and thus for equilibrium behavior. Additionally, because our states have heterogeneous beliefs, the relative values of p^A and p^B also play an important role. Here we introduce terminology to distinguish between the two concepts.

We refer to p_t^i as state i's *fanaticism* at time t. Specifically, state A is more fanatical if p^A is closer to 1 and state B is more fanatical if p^B is closer to 0. Our framework is flexible enough to capture different forms of fanaticism. For example, state A may be highly fanatical, while state B is more uncertain, e.g., $p^A = .9$ and $p^B = .6$. Alternatively, both states could be fanatics, e.g., $p^A = 1$ and $p^B = 0$, or both states relatively uncertain, e.g., $p^A = .6$ and $p^B = .4$. Although our analysis does not require specifying an objective prior belief, if one exists then the extent of a state's fanaticism can change depending on the context. To illustrate, suppose $p^A = .9$ and $p^B = .3$, but the assessment of an objective observer is that $Pr(\omega = A) = .5$. In this case, we might consider both states fanatical, as the beliefs of both diverge significantly from the objective assessment. On the other hand, if the objective probability is $Pr(\omega = A) = .85$, then state A is now only somewhat fanatical, despite being highly confident in its probability of success.[24]

Beyond the values of p^A and p^B in absolute terms, a key quantity in the model is the difference in the states' beliefs. Specifically, at time t *optimism* is measured by $p_t^A - p_t^B$. That is, as the players' beliefs become more divergent optimism increases, since each is relatively more convinced of its success in the event of a decisive conflict. Similarly, decreasing $p_t^A - p_t^B$ decreases optimism, and so there is less disagreement about the probability each state is victorious in a battle. Notice that, fixing p_t^A, increasing p_t^B decreases B's fanaticism and lowers optimism. Thus, there is a tight link between the two concepts as defined here.

Although the states hold different prior beliefs, they each use Bayes's rule to update their beliefs about ω after observing battlefield outcomes. By Bayes's rule, state i's belief that $\omega = A$ at the start of period t, given state A has previously won m battles, is given by:

$$p_t^i(m) = \frac{p_1^i \sigma^m (1-\sigma)^{t-m-1}}{\pi_t^i(m)},$$

where $\pi_t^i(m) = p_1^i \sigma^m (1-\sigma)^{t-m-1} + (1-p_1^i)\sigma^{t-m-1}(1-\sigma)^m$ is the probability player i assigns to A winning a battle in period t. As in the standard case, state i does not expect its own belief to change, e.g., $\mathbb{E}_i[p_{t+1}^i] = \pi_t^i(m)p_{t+1}^i(m+1) + (1-\pi_t^i(m))p_{t+1}^i(m) = p_t^i(m)$. However, because the states hold different prior beliefs, state i does expect state j's belief

[24]For the concept of fanaticism to be sensible with an objective prior, we suppose here that beliefs are on opposite sides of the objective belief. However, note that the model is more general and allows for any $p^A, p^B \in [0,1]$ with $p^A > p^B$.

to change. In particular, i expects j's belief to move closer to its own after each battle, e.g., $\mathbb{E}_A[p^B_{t+1}] = \pi^A_t(m)p^B_{t+1}(m+1) + (1-\pi^A_t(m))p^B_{t+1}(m) > p^B_t(m)$.

As the states observe battlefield outcomes, they update their beliefs. The speed at which state i's belief updates depends on the informativeness of a battle, σ, and on p^i_t. If p^i_t is very high or very low, then state i is already highly certain it knows the state and one battle does not shift its belief very much. Therefore, while a battle shifts both states' beliefs in the same direction, optimism may increase or decrease depending on which state's belief updates faster. However, in the long run optimism will decrease because as the number of battles grows large the states eventually learn whether $\omega = A$ or B, thus $|p^A_t - p^B_t| \to 0$.

Figure 1 shows a sample path of beliefs changing over time as the states observe the outcomes of battles.

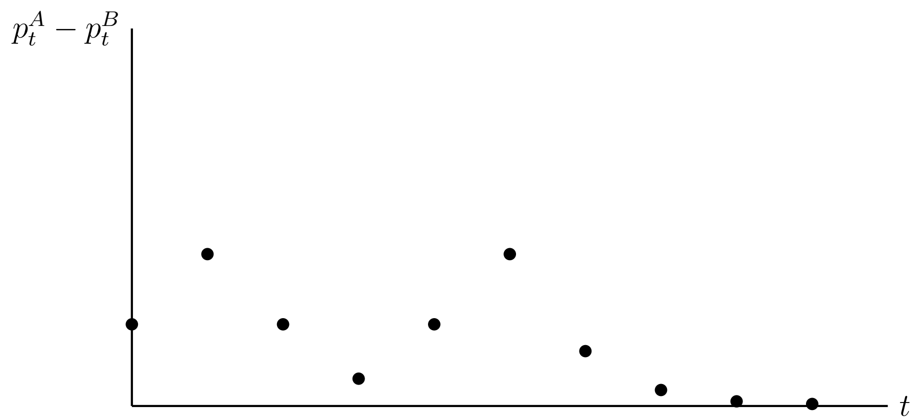

Figure 1: Sample path of beliefs over 9 battles beginning with 3 victories for state A followed by 6 victories for B ($p^A_1 = .94, p^B_1 = .73, \sigma = .8$.)

4 Results

We now study equilibrium behavior in our model. Our solution concept is subgame perfect equilibrium (SPE). Further technical details and proofs are available in the Appendix.

State A may want to fight to obtain a better settlement in the future, because A expects fighting to increase B's belief that $\omega = A$. State A's incentives to fight depend directly on the costs of fighting, c_A, and how quickly it expects fighting to end, parameterized by k. Additionally, its incentives depend on how quickly fighting will move state B's beliefs towards it own. Consequently, the speed of learning, which depends on σ and B's current belief p^B_t, plays an important role in A's decisions.

Our first result characterizes the effect of optimism in equilibrium, when the probability that a single battle is decisive is high.

Lemma 1. *Assume $k > c_A + c_B$. If optimism is sufficiently large then the states fight.*

Lemma 1 demonstrates that the standard intuition about the relationship between optimism and conflict holds in our model when the probability a battle is decisive is large relative to the costs of fighting. Under this condition, in equilibrium the states fight in any period t such that $p_t^A - p_t^B$ is sufficiently large. To see why, imagine that a single battle is decisive, then learning never occurs and the incentive to settle rather than teach the opponent does not exist.

This result, however, relies on the states being confident that a conflict will end quickly. In particular, even if they continue fighting every period until a winner is decided, the expected costs of all-out war remain low. In most cases, when it comes to convincing once opponent through war, the expected costs matter.

Our results demonstrate that the conclusion of Lemma 1 and, thus, the standard intuition, no longer holds when battles are costly relative to the rate of learning. To see how this works, consider the following simplified example.

Example 1. *Assume there are two periods, so $T = 2$, that $p_1^A = 1$, and $k < c_A + c_B$. There exists $\overline{\sigma} \in (1/2, 1)$ such that in the first period:*

1. *If $\sigma < \overline{\sigma}$ then the states settle.*
2. *If $\sigma > \overline{\sigma}$ then there exists $\underline{p}$ and $\overline{p}$, with $0 < \underline{p} < \overline{p} < 1$, such that the states fight if and only if $p_1^B \in (\underline{p}, \overline{p})$.*

We establish this equilibrium in the appendix, but the intuition for why it is true is clear. When the informativeness of a first period battle is low, $\sigma < \overline{\sigma}$, state B updates its beliefs slowly. Consequently, the gains to state A from fighting to shift state B's beliefs are outweighed by the costs of fighting and the states come to a peaceful agreement, regardless of the amount of optimism.

When $\sigma > \overline{\sigma}$ fighting can alter state B's beliefs fast enough for it to be worth it to state A to teach state B through fighting. State A expects the battle to shift B's beliefs towards 1 and improve the settlement it can get from state B in the second period. The caveat to this logic, however, is that when B is highly fanatical, i.e., p_1^B is close to 0, then B updates its beliefs slowly, even if battles are very informative. This implies a non-monotonic effect of optimism on conflict, depicted in Figure 2. When A and B do not have very different beliefs, $p_B^1 > \overline{p}$, optimism is low and the states come to a peaceful agreement for the usual bargaining logic. Moderate levels of optimism, $p_1^B \in (\underline{p}, \overline{p})$, lead to conflict because state A believes fighting will significantly shift B's belief towards 1, allowing A to obtain a better settlement in the second period. Finally, a high degree of optimism due to B's fanaticism, $p_B^1 < \underline{p}$, also leads to peace because fighting is costly and A does not expect to significantly alter B's belief.

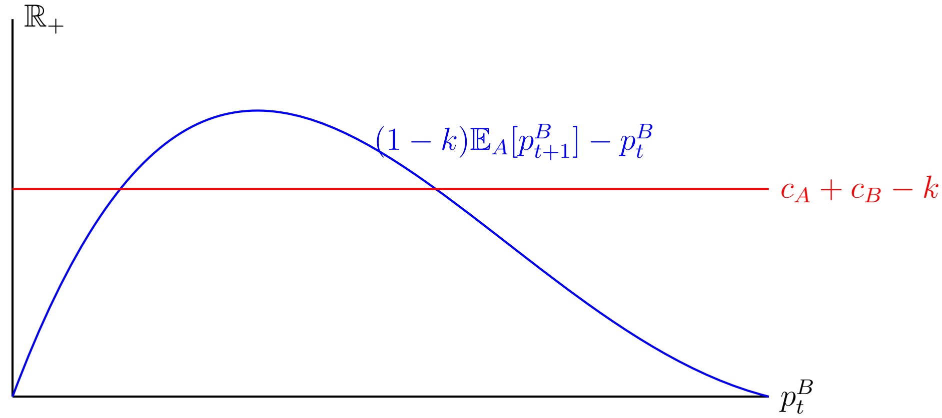

Figure 2: Expected one-round change in B's beliefs (blue) and the cost of delaying a bargain (red)

We now consider the general case, allowing $p_1^A < 1$ and any $T \in \mathbb{N}_1$. In this case, it is difficult to precisely characterize which levels of optimism lead to conflict versus peace. Nevertheless, Proposition 1 shows that fanatical peace still occurs in equilibrium.

Proposition 1. *Assume $c_A + c_B > k$.*

1. *In every period t there exists $\underline{p}_t$ and $\overline{p}_t$, with $0 < \underline{p}_t < \overline{p}_t < p_t^A$, such that in every SPE if $p_t^B \in [0, \underline{p}_t]$ or $p_t^B \in [\overline{p}_t, p_t^A]$ then the states come to a peaceful agreement.*

2. *Assume $c_A + c_B$ is sufficiently small and σ is sufficiently large. For every $p_t^A > p^* + c_A + c_B$ there exists an open interval $\mathcal{I}_t \subseteq [\underline{p}, \overline{p}]$ such that in every SPE if $p_t^B \in \mathcal{I}_t$ then the states fight.*

If optimism is sufficiently low or sufficiently high, then war does not occur for the same logic as discussed after Example 1. Note that the optimism low case includes the standard model where there is no mutual optimism and states have the same beliefs ($p_A = p_B$). Additionally, Proposition 1 shows there exist conditions under which state A prefers to fight to teach its opponent when optimism is moderate. Thus, increasing optimism can move the states from a region of conflict to a region of peace. Overall, this implies a non-monotonic relationship between optimism and war.

The following example demonstrates that the non-monotonic relationship between optimism and war can be even more nuanced than the characterization given in Example 1.

Example 2. *Assume $T = 3$, $\sigma = .8$, $p_1^A = .6$, $c_A = .055$, $c_B = .065$, and $k = .6$. In the first period, the states fight if $p_1^B \in (.026, .047) \cup (.0837, .131)$ and settle otherwise.*

When $p_1^A < 1$ both states update their beliefs over time. Moreover, a longer time horizon implies that changing B's fanaticism, p_1^B, alters both how responsive B's belief is to battlefield outcomes and whether A settles in the future following each battlefield outcome. In this example, A always settles if it fights and loses but it may fight or settle in period 2 following a victory, depending on p_1^B. Example 2 shows that this can lead to the set of beliefs for which the states fight being non-convex, i.e., increasing optimism can lead from fighting to settling back to fighting. Figure 3 demonstrates this effect by comparing A's payoff from settling to fighting as a function of p_1^B.

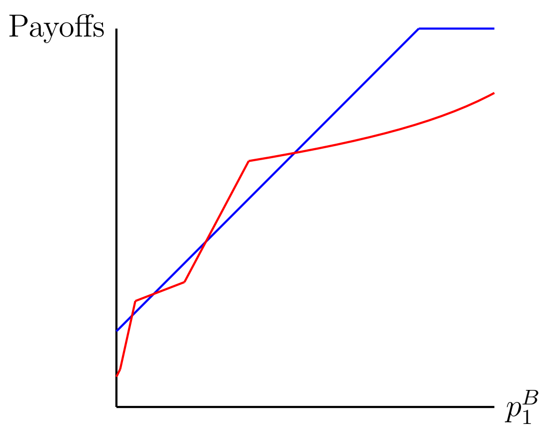

Figure 3: Continuation payoffs in period 1 from settling (blue) and fighting (red) as a function of p_1^B. Parameters are $T = 3$, $\sigma = .8$, $p_1^A = .6$, $c_A = .055$, $c_B = .065$, $k = .6$.

We can now see how fighting changes based on the model primitives. We start by analyzing the informativeness of battles.

Proposition 2. *The maximum number of periods that A expects to fight in equilibrium is maximized at an interior σ.*

As discussed above, when battles are not sufficiently informative, it is not worthwhile to state A to teach state B through fighting. On the other hand, if σ is sufficiently close to 1 then state A expects state B's belief to jump close to p_A after only one battle, mitigating the need for multiple rounds of fighting.

Our next proposition places an upper bound on how long state A expects to fight and studies how this bound changes in the probability of a decisive battle and the costs of fighting.

Proposition 3. *The maximum number of battles that state A expects to fight is bound above by:*

$$\left\lfloor \frac{\ln\left(\frac{c_A+c_B-k}{c_A+c_B}\right)}{\ln(1-k)} \right\rfloor.$$

1. *Increasing c_A or c_B decreases the upper bound.*

2. *If $k \to c_A + c_B$ the upper bound goes to infinity.*

Intuitively, increasing c_A makes state A less willing to fight and increasing c_B makes state B willing to concede more of the good. Thus, both of these effects push towards decreasing the maximum amount of fighting that can occur. When the probability that the battle is decisive is large, the upper bound on the number of battles that could be observed in equilibrium becomes large. When it is likely that the battle is decisive, state A is incentivized to fight due to its optimism about the outcome. Indeed, by Lemma 1, once k exceeds c_A+c_B state A is willing to fight even if B is maximally optimistic, $p_t^B = 0$.

Our next two propositions consider implications of our model. The first analyzes how states bargain after fighting for a long time.

Proposition 4. *Assume T is sufficiently large. There exists $t^* < \infty$ such that if $t = t^*$ then $p_t^B \in [\overline{p}_t, p_t^A]$ and the states settle.*

Although optimism may increase or decrease over time, in the long run the beliefs of both A and B converge to the truth. In turn, this implies that eventually the states must settle. Thus, settlements after long periods of fighting are due to the states converging to similar beliefs about the world, rather than settling because they agree to disagree.

We can also characterize how the expected number of bargaining periods, parameterized by the horizon of the game, influences the circumstances that tie divergent beliefs to war.

Proposition 5. *Assume $k < c_A + c_B$. The set of p_1^B that lead to conflict in the game of length T is a subset of the set of p_1^B that lead to conflict for the game of length $T+1$.*

Proposition 5 shows that allowing for more bargaining periods can lead to war. A longer bargaining time horizon gives A more opportunities to teach B and shift its belief which, in turn, incentivizes conflict. Thus, the closer the states get to all-out war, the more likely they are to settle. Additionally, Proposition 5 suggests that fighting to teach an opponent is another mechanism by which a longer time horizon may incentivize conflict, as in Fearon (1998), rather than encourage cooperation (as in, e.g., Axelrod, 1984).

5 Robustness of the Fanatical Peace

In this section we consider the robustness of our fanatical peace result to two extensions. First, we allow state B to have private information about its prior belief. Second, we consider states with heterogeneous beliefs about how long the conflict will last, rather than over power. Additionally, in the appendix, we extend the previous results to the case where there is an infinite horizon.

5.1 Incomplete Information

Our main model assumes the players have the same information about their environment. In particular, each player knows the other player's prior belief about ω. Doing so allowed us to isolate the new dynamics introduced by non-common prior beliefs. We now relax this assumption, building on the parameter setting from Example 1. Specifically, assume player A still has a commonly known prior $p_A = 1$, while player B has prior belief $Pr(\omega = 1) = \theta$, where $\theta \in \{p_1^B, 1\}$. State B knows θ, while player A has a commonly known prior belief that $Pr(\theta = p_1^B) = \mu$. Thus, state A is uncertain whether state B shares its same belief, $\theta = 1$, or is a fanatic, $\theta = p_1^B < 1$. To characterize equilibrium behavior, we focus on perfect Bayesian equilibrium.[25]

In this setting, the non-fanatical type of state B may be incentivized to act fanatical in order to obtain a better settlement. However, Proposition 6 demonstrates that fanatical peace can still emerge in equilibrium.

Proposition 6. *Assume $T = 2$, $p_1^A = 1$, $c_A + c_B > k$, and $\sigma > \overline{\sigma}$.*

1. *If $p_1^B > \overline{p}$ then in every equilibrium the probability of war is at most μ. For μ sufficiently high the probability is* 0.

[25]Our approach here incorporates the type of definition of optimism as a player's subjective assessment that her opponent is weak, as in Slantchev and Tarar (2011); increasing optimism increases confidence in one's strength relative to an adversary.

2. If $p_1^B \in (\underline{p}, \overline{p})$ then in every equilibrium the probability of war is at least μ.

3. If $p_1^B < \underline{p}$ then in every equilibrium the probability of war is at most μ. For μ sufficiently high the probability is 0.

If $p_1^B > \overline{p}$ or $p_1^B < \underline{p}$ then under complete information state A is willing to settle with both types of state B. Although the $\theta = 1$ type is incentivized to mimic the fanatical type, fighting is more costly to the $\theta = 1$ type because it believes that a decisive war is more likely to lead to defeat. As such, state A can always make an offer that screens the types and faces the standard risk-return tradeoff. It can make the offer that appeases the fanatical type, or make the offer that is better for itself but is rejected by the fanatical type. In either case, the non-fanatical type accepts. Thus, if A's prior belief μ about the "belief" type of B is sufficiently high, then A will make the offer that also appeases the fanatical type as before, and conflict is always avoided.

When $p_1^B \in (\underline{p}, \overline{p})$ uncertainty over the degree of optimism significantly alters state A's incentives from the standard model. Even with complete information state A never wants to settle with the fanatical type. As such, it does not face the usual tradeoff. State A has no incentive to make an offer that appeases the fanatical type and thus (at most) makes an offer that only appeases the non-fanatical type, regardless of its prior belief μ. As such, A always fights the fanatical type and the probability of war is at least μ. This difference in incentives leads to a fanatical peace. That is, the probability of war is maximized when the beliefs of the fanatical type of state B moderately diverge from state A.

5.2 Disagreement about the Decisiveness of Conflict

We have thus far considered how known disagreements about the probability of victory influence conflict. This modeling choice gives a very clear interpretation for how states may have optimistic beliefs relative to each other. However, states may also disagree about other aspects of the environment. Again, looking to Blainey (1973), optimism may not result from different beliefs about the likelihood of victory, but rather delusions that the war will be short. In this section, we explore this possibility and assume states disagree about how quickly the conflict will end.

Let $T = 2$ and assume states have a common prior belief $p \in (0, 1)$ that state A will be victorious in the event of a decisive battle. On the other hand, the states have heterogeneous beliefs about k, the probability a battle is decisive. Specifically, the states are uncertain about the value of k, which can be high or low, $k \in \{\underline{k}, \overline{k}\}$ with $0 < \underline{k} < \overline{k} < 1$. State i has a prior belief that $Pr(k = \underline{k}) = q_1^i$. We allow for $q_1^A \neq q_1^B$, as such, states may disagree about the expected speed at which conflict concludes in a decisive victor. For example, if $q_1^A > q_1^B$

then state A puts higher probability on $k = \underline{k}$ than state B, and thus expects a conflict to last longer before a decisive victory. Additionally, to simplify the statements of the results we assume $c_A = c_B$, and c_B is sufficiently small that A does not obtain the entire pie in the first period.

In this setting, states still learn from battlefield outcomes and update their beliefs using Bayes's rule. If the battle is not decisive they update in favor of $k = \underline{k}$, whereas if the battle is decisive they update towards $k = \overline{k}$. However, states only bargain again if the conflict is not decisive. Thus, conditional on bargaining in the second period they always believe conflict is now more likely to last a long time. Consequently, over time states become more pessimistic that the conflict will end quickly, and strategic considerations are driven by this one-sided updating. This differs from our main model where disagreement is over p, and if states bargain again after a conflict they may have updated their beliefs in favor of A or B. Let $k_t^i = q_t^i \underline{k} + (1 - q_t^i)\overline{k}$ be player i's expected value for k given belief q_t^i in period t. Because states only bargain in the second period after a non-decisive battle we abuse notation and define q_2^i as player i's updated belief conditional on bargaining again, i.e., $q_2^i = Pr(k = \underline{k}|\text{non-decisive battle})$. Given this notation, $k_1^i > k_2^i$.

In the second period countries always settle because differences in k act simply as differences in the cost of war. As noted above, if they do bargain again they must update towards $\underline{k}$. Thus, both actors face higher costs of war in the second period. However, because state A holds the bargaining power, it always benefits from this updating as it can extract more from B. This leads to a tradeoff for A in first period: settle and avoid the cost of fighting, or fight and obtain a better bargain in the second period.

Our first result provides conditions under which the states always come to an agreement, despite their differences in beliefs.

Lemma 2. *There exists $k_* > 0$, $k^* < 1$, and $\overline{q}^A < 1$ such that: If $\underline{k} \geq k_*$, or $\overline{k} \leq k^*$, or $q_1^A \leq \max\{q_1^B, \overline{q}^A\}$, then the states settle peacefully in the first period.*

Lemma 2 shows that the states always settle peacefully if $\overline{k}$ is high or if $\underline{k}$ low. In turn, a necessary condition for fighting is that $\underline{k}$ is low and $\overline{k}$ is high. This necessary condition for fighting has two implications. First, a large difference in the possible values of k ensures that a non-decisive battle shifts the states' beliefs significantly towards $\underline{k}$. This plays a similar role to the assumption that σ is sufficiently large in baseline model. Second, a large difference between $\underline{k}$ and $\overline{k}$ implies there is significant gain to A for moving B's beliefs. Taken together, for fighting to ever be worthwhile it must be that a non-decisive battle significantly shifts state B's belief and the effect of this shift is non-trivial.

Lemma 2 also demonstrates that a requirement for fighting to occur in the first period is that q_1^A is not too low. If $q_1^A = q_1^B$ then there is no disagreement and states can find an agreement, as usual. Additionally, decreasing q_1^A increases k_1^A and makes A believe that a battle is likely to be conclusive, and thus it is unlikely to benefit from trying to shift B's beliefs.

Proposition 7 now gives conditions under which the emergence of war is non-monotonic in the difference in beliefs for the states.

Proposition 7. *Assume* $\underline{k} < k_*, k^* < \overline{k}$, *and* $q_1^A < \overline{q}^A$. *There exists* $0 < \underline{q}_B < \overline{q}_B < 1$ *such that if* $q_1^B \in (\underline{q}_B, \overline{q}_B)$ *then states fight in the first period; otherwise, the states settle peacefully.*

For this non-monotonicity to occur requires $\underline{k}$ low, $\overline{k}$ high, and q_1^A low. In this case, if q_1^B is also low, then actors settle because there is little disagreement. For intermediate values of q_1^B the states fight. Here, State A is willing to incur the costs of war and risk of conclusive battle to try to shift B's beliefs. Finally, for q_1^B high state B is stubborn and it does not update its beliefs very much even following a non-decisive battle. In this case, it is also not worth it for A to fight a B with very different beliefs. We thus have a range in which fanatical peace emerges, so that proceeding from an intermediate amount of disagreement to a large amount of disagreement moves the states from war to peace.

We now reconnect our investigation into optimism over k to the conditions for fanatical peace in the baseline model. Notice that a scope condition for Proposition 1 to hold is that $k < c_A + c_B$. Thus, a requirement for fanatical peace is that states are optimistic they can win a conflict, but believe victory could take a long time. This suggests that fanatical peace may be difficult to sustain if states are optimistic about both their probability of victory *and* how quickly they can succeed. However, notice from this extension that after a non-decisive conflict the states revise their expectation about k downward. Thus, even if optimism initially leads to conflict, if the conflict is not decisive then the expectation of k decreases, which can move the parameters back into a region of fanatical peace.

6 Empirical Implications

Some enemies are teachable; others are not. If our model is accurate, then we should see at least two kinds of events: i) we should see wars begun in order to change an enemy's mind; and ii) we should see wars avoided because it would be too costly to change an enemy's beliefs. The latter we have called "fanatical peace." Since our paper is primarily interested in (ii), we will dwell on it more extensively in the examples that follow; but we begin with (i) to establish the plausibility of the causal mechanism. Importantly, we are not arguing (and our model does not show) that divergent beliefs always decrease the likelihood of war.

Rather, we claim that they *sometimes* reduce the likelihood of war, and that these cases are interesting and unexplained by the current literature.

An actor might deliberately initiate conflict to humble an opponent, i.e., to correct that opponent's self-confidence. Leaders initiate these short, limited conflicts to educate the other side; they relent when the other side updates its beliefs.[26]

The European colonial empires undertook such conflicts regularly, both as outright wars and as smaller operations. The Second Afghan War offers an example. As Britain grew stronger, and as Russo-British relations frayed, the British Viceroy sought greater influence in Afghanistan.[27] Afghanistan's Shere Ali Khan resisted. He sought more or less to maintain the status quo of the past two decades, a *modus vivendi* in which Afghanistan leaned on British security guarantees without actually choosing between its British and Russian neighbors. This policy became unacceptable to Britain.[28] Importantly, the Brits regarded the Afghan rulers as eminently "teachable" (Hanna, 1899, 23)—a telling description—and they sought to correct Afghan rulers' perceptions by only a small increment. They sought to demonstrate that Britain was now more powerful than before, and so the Afghans would have to accept British envoys into their territory. This demand, while emphatic, was relatively modest, and it would not depart significantly (in British eyes) from previous Anglo-Afghan relations.[29] It seems that the war was begun entirely for the logic we study in the model: to compel a rational opponent to update his beliefs. The war's initial phase, which seems to have been almost entirely about educating the enemy, lasted less than seven months and concluded with the Treaty of Gandamak. It obtained for Britain even more than it had hoped from the conflict, including a decisive shift in the Afghan government's perceptions of strategic reality. Unfortunately for the British, while their limited war successfully changed the minds of the "teachable" Afghan leadership class, they had not reckoned on the Afghan people, and so a second and unexpected phase of the war followed soon after.

We now turn to examples of fanatical peace. A "fanatical peace" occurs when two nations avoid conflict *because* each knows the other is too radical to be persuaded. As dogs that don't bark, these cases are difficult to identify. We can adopt a few strategies to try to

[26]Because the initiator, too, can operate under mistaken beliefs, these expeditions occasionally backfire, as in the Mayaguez Incident, which seems to have begun as a limited operation to alter the other side's perceptions.

[27]One contemporary account observes, "His [the Viceroy's] difficulty was this—that the British government wanted to get something from the Ameer, whereas the Ameer did not want to get anything from the British Government" (Campbell, 1879, 455).

[28]MacGregor (1908). Britain chose to act after Ali permitted a Russian Mission in Kabul. See also Hanna (1899, 185).

[29]Shortly before the war, the Kabul Envoy remarked, "we have long been on terms of friendship, and the Amir now clings to the skirt of the British Government, and till his hand be cut off, he will not relax his hold" (Hanna, 1899, 167).

identify them.

First, we can recognize that the costs of war are endogenous, not given. This fact is easily overlooked in the conflict literature, especially in one-shot bargaining models of war. Because costs are endogenous, statements like "war would be too costly" can conceal more than they explain: *why* would the war be as costly as projected? The expected costs of conflict are not only a function of military technology; they are also a function of beliefs. The more polarized are enemies' beliefs, the more costly would be a war. So, when we attribute peace to the large prospective costs of war, it is important to ask to what extent these costs reflect technology or beliefs.

The Korean peninsula offers an interesting example. There are many scholarly treatments of the Korean War, yet these treatments rarely go beyond 1953.[30] They seem to take peace for granted, as if peace after war were the natural course of things. Yet this is not obvious. The Korean War ended indecisively, with few of the outstanding issues resolved. (Indeed, as trivia buffs like to recall, the war is technically still ongoing.) We might conclude that the costs of war have kept the ROK and DPRK at peace, yet this raises the question: why are the costs sufficient to prevent renewed hostilities? To the extent that the costs of war kept the Korean peninsula at peace, these costs do *not* primarily reflect new military technology nor the risk of nuclear war, since China did not acquire nuclear weapons for over a decade and the DPRK until the Bush 43 administration. Rather, the prospective costs of war seem primarily to reflect the intransigent worldviews of the antagonists, especially in the 1953-1964 period. Because these beliefs are so very far apart, they make war more costly—and, in doing so, may make peace easier to sustain. Indeed, according to some accounts, the South Koreans were so convinced of their own odds of victory that they preferred to continue the war in 1953: Syngman Rhee "publicly criticized the truce [the Armistice ending the Korean War] in the hope that the war could continue against the northern republic" (Lee, 2013, 184). Similarly, it may be worth noting that Eisenhower "proposed no specific exit plan" to end the conflict; rather, his only strategy seems to have been "to escalate the war as a means to end it" (Reiter, 2009, 88).

As the two nations' ideologies have drawn still further apart since 1953, especially with the democratization of South Korea in the 1980s, the prospective costs of war have likewise increased, as well. And it is not hard to find a surfeit of bellicose rhetoric from Korean leaders about their likelihoods of prevailing in a second Korean War: just this year, Kim Jong Un talked of "annihilating" combined US and ROK forces; in response, Shin Won-sik,

[30] Importantly, we are not offering fanatical peace as a sole or even primary explanation for the Korean War's termination. We defer to scholars like Reiter (2009) and Stanley (2009), who emphasize commitment problems and domestic politics. But we do think even these seminal accounts of the war leave the subsequent peace underexplained, and we believe the logic of fanatical peace might buttress their accounts.

the ROK Defense Minister, replied that South Korea, in the event of conflict, would replace the government of North Korea and terminate the Kim dynasty (Kim, 2024). It is not moderated expectations of victory that have kept the Koreas at peace, as the conventional logic of mutual optimism would suggest.

While the costs of war are a function of beliefs, they are also a function of the likelihood of decisive victory. As a decisive victory becomes less likely (in the model, as k declines), the prospective costs of war will increase. Therefore, we can look for cases where, as states downgrade their estimates about k (about the likelihood of a decisive victory), they likewise become more willing to settle a war and avoid future ones—even though their ideological differences remain just as pronounced.

This dynamic seems to underpin the perennial tensions between Morocco and Algeria. In addition to a geopolitical rivalry, since Algerian independence the two governments have typically been guided by antithetical worldviews: Morocco, by a more-or-less pro-Western, traditionalist monarchy; Algeria, by socialist, pro-Soviet, and at times Marxist radicals. Their visions for their own nations and for the region are deeply opposed, and this opposition has often been exacerbated by nationalist, expansionist ideologies like Greater Morocco. In 1974, Algeria expelled all Moroccan nationals, and there have been several MIDs over the past six decades, any one of which might easily have erupted into war. Yet, despite these tensions, the two have not fought a war since the Sand War of 1963-64.[31] Their uneasy modus vivendi seems to reflect, in part, the fanatical peace we theorize here.

The Sand War war began with a confluence of greed (disputed territory was rich in natural resources), diversionary politics (Morocco faced unrest at home), and ideology (Moroccan nationalism, including Istiqlal, pushed for war). It seems that Morocco initially hoped a newly-independent Algeria would be open to renegotiating the border, given positive Algerian rhetoric and Moroccan aid during Algeria's struggle for independence. These hopes were dashed.[32] Facing an intransigent Algeria that had only just achieved independence, and that trailed Moroccan military spending, Morocco believed it could quickly force the Algerians to recognize the hard realities of their situation and, consequently, revise the border. While Morocco enjoyed significantly more victories than Algeria during the Sand War, it nonetheless found itself less able than it had expected to quickly force a decision over the territory, especially with the possibility of foreign intervention looming. The war concluded in a bitter peace that restored the status quo ante. Indeed, the war seems to have demonstrated to both sides that a decisive outcome was unlikely and that further conflict would be "a lose-lose

[31]While Algeria backed rebel groups in the Western Sahara War, it did not intervene directly.

[32]As one scholar noted shortly after the war, "Morocco's hope that an independent Algerian government would prove more understanding than the French appears to have had no concrete basis beyond a too literal interpretation of mutual professions of North African brotherhood" (Zartman, 1965, 163).

outcome."[33]

Importantly, this peace did *not* proceed from a reduction in ideological tensions, nor did it foster a rapprochement between Algeria and Morocco. Quite the reverse: the war and subsequent peace entrenched the belligerents' ideological opposition and erased any fellow-feeling that might have lingered from Algeria's struggle for independence. This antipathy has lasted to the present day, and in 2024 the two countries seem more fanatically opposed than ever—yet still, fortunately, at peace.[34] Repurposing Aron's famous quip, one scholar summarizes the situation as "peace impossible, war improbable" (Lefèvre, 2016, 738).

The outbreak, tempo, and settlement of the Sand War thus track the logic in proposition 7 above. An initial, moderate disagreement about the likely length of war results in conflict. The war reveals to Morocco that its ability to achieve a decisive outcome is lower than expected; moreover, far from narrowing the gulf between the belligerents, if anything the war widens it (as Figure 2 shows can happen). As their estimates about the decisiveness of war decreased, the antagonists moved into a parameter region where their intense disagreement—a disagreement stemming from their antithetical worldviews—could achieve a fanatical peace. That fanatical peace has endured to the present day.

Finally, we can look for cases when opponents back down from the brink of war because they suddenly perceive the gap between their beliefs to be larger than previously thought. Whereas the conventional logic of mutual optimism would suggest this sudden, yawning gap should make war more likely, the logic of fanatical peace suggests the opposite. As we discussed in the introduction, the net effect of these two logics is difficult to predict. Yet we can say that mutual optimism struggles to rationalize cases where peace results from a sudden revelation of *dis*agreement between the players; we need the logic of fanatical peace to explain such a case. The Oregon Dispute offers an example.[35]

The British Empire and the rapidly-expanding United States avoided war, which seemed probable,[36] by negotiating a last-minute settlement. This settlement became feasible only when both sides recognized the extreme beliefs of the other; so long as at least one side believed the other's beliefs to be relatively corrigible, war seemed likely. Each side was convinced its own legal case was airtight; in terms of material position, each also believed it

[33] Zoubir (2020). Zoubir goes on to quote a "senior" Algerian official who said that war, in the current era, would "compromise the future."

[34] A recent piece in the popular press observes, "never have the two regimes been so far apart" (Evans, 2024, 41). Scholarship on the region paints the same picture: Moroccan and Algerian leaders have invested so much in demonizing the other that their worldviews are now worlds apart (Zoubir, 2020).

[35] For the connection between this dispute and Anglo-American trade relations, see James and Lake (1989).

[36] The diplomatic historian Thomas Bailey describes two notable flareups. In 1841, in response to an American bill that would have constructed forts in the territory, Palmerston, on the floor of the Commons, said "it would be a declaration of war." In 1843, Senator Benton (MO) suggested, approvingly, that armed settlers "will annihilate the Hudson's Bay Company" (Bailey, 1980, 224).

held the upper hand.

It is important to recognize that, despite their shared Anglo heritage, the antagonists' worldviews at this time were often far apart. Britain strongly associated the United States with slavery, even though Oregon would not be a slaveholding state, and with belligerent nationalism. Both associations were not unjustified in this era of Manifest Destiny, which peaked with James Polk's "extreme nationalism" (Bailey, 1980, 226). The gap between these nations would become especially prominent when the Tory government—its foreign policy headed by the milquetoast Aberdeen—fell to the more excitable, and more radical, Whigs.

In the 1844 election, Polk (then a candidate) made sweeping claims to all of Oregon. Far from moderating these demands after his victory, he doubled down on them in his Inaugural Address. He also declined British offers of arbitration. Both sides escalated the dispute, and the British began naval preparations for war. Yet, the issue was abruptly settled when the Tory government fell, to be replaced imminently by the even-more belligerent Whigs. Just three days before receiving the British proposal—before the Tories fell—Polk had been "'certain' it ought to be rejected" (Bailey, 1980, 233). Facing a more entrenched foe, one he had no hope of persuading, Polk abandoned his hope of reshaping British perceptions. A hastily-drafted treaty was signed a matter of days before the Whigs took power.

These examples cover the two major implications of our model—that some wars should be fought to change beliefs, and that other wars should be avoided because beliefs are or become too polarized to change. Both predictions follow immediately from the model. If we are willing to step a bit beyond the model, we can also conjecture a third implication. If war with a radicalized enemy is often not worthwhile, we might expect some canny leaders to try to make themselves and their nations "unteachable." It does not seem a stretch to suggest that during WWII the Japanese leadership, before Hiroshima, undertook just such a project, as with its infamous propaganda campaign, "The Glorious Death of One Hundred Million." But it must be left to future work to model these perverse incentives.

7 Conclusion

"Once customs are established and prejudices have become deeply rooted," mused Jean-Jacques Rousseau, "it is a dangerous and vain undertaking to want to reform them." Turned on its head, this truism has a surprising implication: conflict may be more likely when disagreements are modest. That is, there is a non-monotonic relationship between the polarization of leaders' beliefs and the probability they choose war over peace.

While wars are not always more likely between factions than between fanatics, understanding when divergence in worldviews leads to war is of significant historical and theoretical interest. We can reject the idea that trenchant disagreement always increases the risk of conflict. Finally, we hope that this paper has shown how worldviews can influence rational action, and we hope it spurs others to further integrate ideational and rational theory.

Our results suggest several directions for further research. In our model, we treat actors as rational apart from their prior worldviews. This modeling choice reflects our desire to study different kinds of information within a rationalist framework; we were less interested in studying the effects of bounded rationality. But we might have mixed rationality with nonrationality still further. For instance, we might have asked, when two actors fundamentally disagree, can that disagreement bias them, or make them irrationally stubborn, when confronted with new evidence? To the extent that such irrationality would further reduce the "teachability" of those with whom one strongly disagrees, we conjecture that it would further expand the conditions for a "fanatical peace."

Second, our results also raise the question, do citizens or leadership coalitions have an incentive to delegate bargaining to agents more fanatical than themselves? While a full analysis of the agency problem for crisis bargaining is beyond the scope of this paper, we can suggest how the selection incentives are affected by the beliefs of potential agents and each state's bargaining power. In line with Johnson and Fowler (2011), the structure of bargaining may reward overconfidence in certain situations. In particular, actors with comparatively little bargaining power may prefer to nominate agents with radical beliefs. Paradoxically, doing so might also reduce conflict.

Finally, we chose to model optimism as the gap between players' posterior expectations about power (or the decisiveness of conflict in the extension), but it could be operationalized in other ways. Many would be only cosmetic changes to the game; however, two are worth noting. First, if we allowed players to disagree about the way s_t is drawn, then we might define increasing optimism as the extent to which positive signals reinforce players' subjective expectations about ω; in this case, deriving conditions for "fanatical peace" would likely differ from those we deduce; nonetheless, fanatical peace could still occur in such a world, and we offer an example in the appendix. Second, we might create uncertainty over multiple variables, e.g. uncertainty about the other player's resolve as well as about ω, like Debs (2022); but it becomes a bit unwieldy to model changes in players' beliefs over these variables with heterogenous priors, especially since one is a function of the other, and since players' behavior is conditioned on their posterior beliefs about other players' posterior beliefs.

References

Adler, Emanuel. 1992. "The emergence of cooperation: national epistemic communities and the international evolution of the idea of nuclear arms control." *International organization* 46(1):101–145.

Allan, Bentley B. 2018. *Scientific cosmology and international orders*. Vol. 147 Cambridge University Press.

Alonso, Ricardo and Odilon Câmara. 2016. "Bayesian persuasion with heterogeneous priors." *Journal of Economic Theory* 165:672–706.

Arena, Philip and Scott Wolford. 2012. "Arms, Intelligence, and War." *International Studies Quarterly* 56(2):351–365.
URL: *http://dx.doi.org/10.1111/j.1468-2478.2012.00724.x*

Aumann, Robert J. 1976. "Agreeing to Disagree." *The Annals of Statistics* 4(6):1236–1239.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. BasicBooks.

Bailey, Thomas A. 1980. *A Diplomatic History of the American People*. 10 ed. Englewood Cliffs, NJ: Prentice-Hall.

Baliga, Sandeep and Tomas Sjöström. 2023. Long wars. Technical report Working Paper.

Bas, Muhammet and Robert Schub. 2016. "Mutual Optimism as a Cause of Conflict: Secret Alliances and Conflict Onset." *International Studies Quarterly* 60(3):552–564.

Bils, Peter and William Spaniel. 2017. "Policy bargaining and militarized conflict." *Journal of Theoretical Politics* 29(4):647–678.

Black, Jeremy. 1983. "The theory of the balance of power in the first half of the eighteenth century: a note on sources." *Review of International Studies* 9(1):55–61.

Blainey, Geoffrey. 1973. *The Causes of War*. Macmillan.

Brand, Charles M. 1962. "The Byzantines and Saladin, 1185-1192: Opponents of the Third Crusade." *Speculum* 37(2):167–181.

Campbell, George. 1879. *The Afghan Frontier*. London: E. Stanford.

Chadefaux, Thomas. 2011. "Bargaining over power: when do shifts in power lead to war?" *International Theory* 3(2):228–253.

Chamley, Christophe. 2004. *Rational herds: Economic models of social learning.* Cambridge University Press.

Che, Yeon-Koo and Navin Kartik. 2009. "Opinions as incentives." *Journal of Political Economy* 117(5):815–860.

Conde, Amparo M. Mármol and Clara Ponsati Obiols. 2006. Bargaining Multiple Issues with Leximin Preferences. Economic Working Papers at Centro de Estudios Andaluces E2006/05 Centro de Estudios Andaluces.

Davenport, Christian and Scott Gates. 2014. "Four Reasons Why Interstate Conflict Scholars Don't Read Intrastate Work and Why They Are Wrong, Part 1.".
URL: *https://politicalviolenceataglance.org/2014/01/21/four-reasons-why-interstate-conflict-scholars-dont-read-intrastate-work-and-why-they-are-wrong-part-1/*

Debs, Alexandre. 2022. "Mutual Optimism and War, and the Strategic Tensions of the July Crisis." *American Journal of Political Science* 66(2):271–284.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12569*

Debs, Alexandre and Jessica Chen Weiss. 2016. "Circumstances, Domestic Audiences, and Reputational Incentives in International Crisis Bargaining." *Journal of Conflict Resolution* 60(3):403–433.
URL: *https://doi.org/10.1177/0022002714542874*

Deneckere, Raymond and Meng-Yu Liang. 2006. "Bargaining with interdependent values." *Econometrica* 74(5):1309–1364.

Duelfer, Charles A. 2011. "Chronic Misperception and International Conflict: The U.S.-Iraq Experience.".

Evans, Martin. 2024. "THE SAND WAR." *History Today* 74(2):28 – 41.
URL: *http://ezproxy.baylor.edu/login?url=https://search.ebscohost.com/login.aspx?direct=trueamp;db= liveamp;scope=site*

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

Fearon, James D. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52(2):pp. 269–305.

Fearon, James and Xin Jin. 2021. Armed Conflict Bargaining. Technical report mimeo, Stanford University.

Fey, Mark and Kristopher W. Ramsay. 2007. "Mutual Optimism and War." *American Journal of Political Science* 51(4):738–754.

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* 46(4):819–837.

Fox, Jonathan. 2002. "Ethnic Minorities and the Clash of Civilizations: A Quantitative Analysis of Huntington's Thesis." *British Journal of Political Science* 32(3):415–434.

Galperti, Simone. 2019. "Persuasion: The Art of Changing Worldviews." *American Economic Review* 109(3):996–1031.

Goldstein, Judith and Robert O. Keohane, eds. 1993. *Ideas and Foreign Policy: Beliefs, Institutions, and Political Change*. Cornell University Press.

Haas, M.L. 2022. *Frenemies: When Ideological Enemies Ally*. Cornell Studies in Security Affairs Cornell University Press.

Hanna, H.B. 1899. *The Second Afghan War, 1878-79-80: its causes, its conduct, and its consequences*. Vol. 1 Westminster: Archibald Constable Co.

Harsanyi, John C. 1967. "Games with incomplete information played by "Bayesian" players, I–III Part I. The basic model." *Management science* 14(3):159–182.

Haynes, Kyle and Brandon K. Yoder. 2020. "Offsetting Uncertainty: Reassurance with Two-Sided Incomplete Information." *American Journal of Political Science* 64(1):38–51.

Hirsch, Alexander V. 2016. "Experimentation and persuasion in political organizations." *American Political Science Review* 110(1):68–84.

Holmes, Marcus. 2015. "Believing This and Alieving That: Theorizing Affect and Intuitions in International Politics." *International Studies Quarterly* 59(4):706–720.

Holmes, Marcus and David Traven. 2015. "Acting Rationally Without Really Thinking: The Logic of Rational Intuitionism for International Relations Theory." *International Studies Review* 17(3):414–440.
URL: *https://doi.org/10.1111/misr.12228*

Horowitz, Michael C. and Allan C. Stam. 2014. "How Prior Military Experience Influences the Future Militarized Behavior of Leaders." *International Organization* 68(3):527–559.

Huntington, Samuel P. 1993. “If Not CivilizaCivil, What? Paradigms of the Post-Cold War World.”.

Izzo, Federica. 2023. “Ideology for the Future.” *American Political Science Review* 117(3):1089–1104.

James, Scott C. and David A. Lake. 1989. “The second face of hegemony: Britain’s repeal of the Corn Laws and the American Walker Tariff of 1846.” *International Organization* 43(1):129.

Jervis, Robert. 1976. *Perception and Misperception in International Politics.* Princeton University Press.

Johnson, Dominic D.P. 2004. *Overconfidence and War.* Harvard University Press.

Johnson, Dominic D.P. 2020. *Strategic Instincts: the Adaptive Advantages of Cognitive Biases in International Politics.* Princeton University Press.

Johnson, Dominic DP and Dominic Tierney. 2011. “The Rubicon theory of war: how the path to conflict reaches the point of no return.” *International Security* 36(1):7–40.

Johnson, Dominic DP and James H Fowler. 2011. “The evolution of overconfidence.” *Nature* 477(7364):317–320.

Johnson, Dominic D.P, Rose McDermott, Emily S Barrett, Jonathan Cowden, Richard Wrangham, Matthew H McIntyre and Stephen Peter Rosen. 2006. “Overconfidence in wargames: experimental evidence on expectations, aggression, gender and testosterone.” *Proceedings of the Royal Society of London B: Biological Sciences* 273(1600):2513–2520.

Johnson, James Turner. 2015. *Ideology, reason, and the limitation of war: Religious and secular concepts, 1200-1740.* Princeton University Press.

Joseph, Michael F. 2021. “A Little Bit of Cheap Talk Is a Dangerous Thing: States Can Communicate Intentions Persuasively and Raise the Risk of War.” *The Journal of Politics* 83(1):166–181.
URL: *https://doi.org/10.1086/709145*

Jost, John T., Mahzarin R. Banaji and Brian A. Nosek. 2004. “A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo.” *Political Psychology* 25(6):881–919.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9221.2004.00402.x*

Kahneman, Daniel and Jonathan Renshon. 2007. "Why Hawks Win." *Foreign Policy* (158):pp. 34–38.

Kant, Immanuel. 1983. *Perpetual Peace and Other Essays.* Hackett Publish Company.

Katzenstein, Peter J. 2022. *Worldviews in World Politics.* Cambridge Studies in International Relations Cambridge University Press p. 1–70.

Khong, Yuen Foong. 1992. *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965.* Princeton University Press.

Kim, Hyung-Jin. 2024. "North Korea's Kim says military should 'thoroughly annihilate' US and South Korea if provoked.".
URL: *https://apnews.com/article/north-korea-kim-us-missiles-nuclear-6c8834f71ac43bb9d0addc404fe00f18*

Kurz, Mordecai. 1996. "Rational Beliefs and Endogenous Uncertainty." *Economic Theory* 8(3):383–397.
URL: *http://www.jstor.org/stable/25054980*

Lee, Steven. 2013. "The Korean Armistice and the End of Peace: The US-UN Coalition and the Dynamics of War-Making in Korea, 1953–76." *Journal of Korean Studies* 18(2):183–224.

Lefèvre, Raphaël. 2016. "Morocco, Algeria and the Maghreb's cold war." *Journal of North African Studies* 21(5):735 – 740.

Levy, Jack S. and William R. Thompson. 2010. *Causes of War.* Wiley-Blackwell.

Little, Andrew T. and Thomas Zeitzoff. 2017. "A Bargaining Theory of Conflict with Evolutionary Preferences." *International Organization* 71(3):523557.

MacGregor, Charles. 1908. *The Second Afghan War, 1878-80: Abridged Official Account, produced in the Intelligence branch army headquarters, India.* London: John Murray.

Mercer, Jonathan. 2013. "Emotion and Strategy in the Korean War." *International Organization* 67(2):221–252.

Morris, Stephen. 1995. "The common prior assumption in economic theory." *Economics & Philosophy* 11(2):227–253.

Morris, Stephen. 1996. "Speculative investor behavior and learning." *The Quarterly Journal of Economics* 111(4):1111–1133.

Morrow, James D. 2014. *Order within anarchy: The laws of war as an international institution.* Cambridge University Press.

Parker, Geoffrey. 1998. *The grand strategy of Philip II.* Yale University Press.

Powell, Robert. 2004*a*. "Bargaining and Learning While Fighting." *American Journal of Political Science* 48(2):344–361.

Powell, Robert. 2004*b*. "Bargaining and learning while fighting." *American Journal of Political Science* 48(2):344–361.

Reiter, Dan. 2009. *How Wars End.* Princeton, NJ: Princeton University Press.

Said, Edward W. 2001. "The Clash of Ignorance.".

Sheffer, Lior, Peter John Loewen, Stuart Soroka, Stefaan Walgrave and Tamir Sheafer. 2018. "Nonrepresentative Representatives: An Experimental Study of the Decision Making of Elected Politicians." *American Political Science Review* 112(2):302321.

Simsek, Alp and Muhamet Yildiz. 2016. Durability, deadline, and election effects in bargaining. Technical report National Bureau of Economic Research.

Slantchev, Branislav L. 2003. "The Principle of Convergence in Wartime Negotiations." *The American Political Science Review* 97(4):621–632.

Slantchev, Branislav L. and Ahmer Tarar. 2011. "Mutual Optimism as a Rationalist Explanation of War." *American Journal of Political Science* 55(1):135–148.

Smith, Alastair and Allan C Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48(6):783–813.

Spaniel, William and Peter Bils. 2018. "Slow to learn: Bargaining, uncertainty, and the calculus of conquest." *Journal of Conflict Resolution* 62(4):774–796.

Spaniel, William, Peter Bils and Gleason Judd. 2020. "Stall Wars: When Do States Fight to Hold onto the Status Quo?" *International Studies Quarterly* 64(3):523–529.

Stanley, Elizabeth A. 2009. *Paths to peace: Domestic coalition shifts, war termination and the Korean War.* Stanford University Press.

Tickner, J. Ann. 2014. 147On Taking Religious Worldviews Seriously. In *A Feminist Voyage through International Relations.* Oxford University Press.
URL: *https://doi.org/10.1093/acprof:oso/9780199951246.003.0010*

Tickner, Judith Ann. 2009. *Power, Interdependence, and Non-State Actors in World Politics.* Princeton University Press chapter On taking religious worldviews seriously, pp. 223–239.

Van den Steen, Eric. 2004. "Rational Overoptimism (and Other Biases)." *American Economic Review* 94(4):1141–1151.

Van den Steen, Eric. 2009. "Authority versus Persuasion." *American Economic Review* 99(2):448–453.

Van Evera, Stephen. 1999. *Causes of War: Power and the Roots of Conflict.* Cornell University Press.

Vasserman, Shoshana and Muhamet Yildiz. 2019. "Pretrial negotiations under optimism." *The RAND Journal of Economics* 50(2):359–390.

Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44(3):pp. 469–484.
URL: *http://www.jstor.org/stable/2669259*

Watanabe, Yasutora. 2006. "Learning and bargaining in dispute resolution: Theory and evidence from medical malpractice litigation." *Working paper* .

Wendt, Alexander. 1999. *Social Theory of International Politics.* Cambridge University Press.

Wittman, Donald. 1979. "How a war ends: A rational model approach." *Journal of Conflict Resolution* 23(4):743–763.

Yildiz, Muhamet. 2003. "Bargaining without a Common Prior - An Immediate Agreement Theorem." *Econometrica* 71(3):793–811.

Yildiz, Muhamet. 2004. "Waiting to Persuade." *The Quarterly Journal of Economics* 119(1):223–248.

Yildiz, Muhamet. 2011. "Bargaining with Optimism." *Annu. Rev. Econ.* 3(1):451–478.

Zartman, I William. 1965. "The politics of boundaries in North and West Africa." *The Journal of Modern African Studies* 3(2):155–173.

Zoubir, Yahia H. 2020. *Shocks and Rivalries in the Middle East and North Africa.* Georgetown University Press chapter The Algerian-Moroccan Rivalry: Constructing the Imagined Enemy, pp. 179–200.

A Online Appendix

A.1 Proofs of Main Results

In this section we prove the main results of the paper. We begin by outlining the organization of the proofs. Note that, throughout, we allow for an infinite horizon (with the exception of Lemma 5 which explicitly studies the finite case.)

We first prove that a player has the same continuation payoff in any equilibrium (Lemma A.1). The proof of our main result (Proposition 1) is then structured around three arguments:

1. If players' beliefs are close to each other then they must settle for the standard arguments.
2. If p_t^B is close to 0 then the states must settle because state B does not update quick enough to make fighting worth it to A.
3. There are conditions for which beliefs are sufficiently malleable that a region of fighting exists.

We first prove a number of preliminary lemmas that we use in the proofs of Proposition 1. Lemmas A.2, A.6, and A.7 provide characterization of the expected change in A's belief about p^B. Lemmas A.3 and A.4 establish bounds on A's continuation payoffs. Lemma A.5 shows if B's initial belief p_0^B is sufficiently low then its updated belief remains low even after A wins many battles. We next use the characterization from these results to prove Propositions 2-5. Finally, we conclude with the proofs for Examples 1 and 2.

Lemma A.1. *Given (m, t, i), there exists a unique $V_t^i(m) \in [-\frac{c_i}{k}, 1]$ such that, in any subgame perfect equilibrium, the continuation value of player i at (m, t) is $V_t^i(m)$. Furthermore,*

$$V_t^A(m) = \max\Big\{\min\{1, 1 - V_t^B(m), kp_t^A(m)\} - c_A + (1-k)\mathbb{E}^A[V_{t+1}^A|m, t]\Big\},$$
$$V_t^B(m) = \max\Big\{1 - p_t^B(m) - \frac{c_B}{k}, 0\Big\}.$$

In any subgame perfect equilibrium, State B accepts x if and only if $x \leq 1 - V_t^B(m)$. If $V_t^A(m) + V_t^B(m) \leq 1$ then the states come to a peaceful agreement. Otherwise, they fight.

Proof. To obtain uniqueness of continuation values we adapt the proof of Theorem 1 from Yildiz (2003) to our setting. Let $\underline{V}_t^i(h^t)$ and $\overline{V}_t^i(h^t)$ be the least and greatest payoff that player i can expect to get at the beginning of period t following history h^t given the set of remaining strategy profiles is Σ^*.

Given any history h^t, rejecting any offer x such that

$$\begin{aligned}1-x >& k(1-p_t^B(m)) - c_B + (1-k)\mathbb{E}^B[\overline{V}_{t+1}^B|m,h^t] \\ &= k(1-p_t^B(m)) - c_B + (1-k)\Big(\pi_t^B\overline{V}_{t+1}^B(m+1|h^t) + (1-\pi_t^B)\overline{V}_{t+1}^B(m|h^t)\Big)\end{aligned}$$

is conditionally dominated for B by accepting $1-x$, since $\overline{V}_{t+1}^B(m|h^t)$ is the most B expects to get in Σ^*. Hence, given any $\epsilon > 0$ if A makes an offer with $1-x = k(1-p_t^B(m)) - c_B + (1-k)\mathbb{E}^B[\overline{V}_{t+1}^B|m,h^t] + \epsilon$ it will be accepted. Therefore,

$$\begin{aligned}\underline{V}_t^A(m|h^t) \geq \max\Big\{&1-k+kp_t^B(m) + c_B - (1-k)\mathbb{E}^B[\overline{V}_{t+1}^B|m,h^t], \\ &kp_t^A(m) - c_A + (1-k)\mathbb{E}^A[\underline{V}_{t+1}^A|m,h^t]\Big\}.\end{aligned}$$

Additionally, since $\underline{V}_t^B$ is the least B can expect to get in Σ^*, for B accepting x such that $1-x < k(1-p_t^B(m)) + (1-k)\mathbb{E}^B[\underline{V}_{t+1}^B|m,h^t] - c_B$ must be conditionally dominated by rejecting. Thus,

$$\begin{aligned}\overline{V}_t^A \leq \max\Big\{&kp_t^B(m) + c_B + (1-k)(1-\mathbb{E}^B[\underline{V}_{t+1}^B|m,h^t]), \\ &kp_t^A(m) - c_A + (1-k)\mathbb{E}^A[\overline{V}_{t+1}^A|m,h^t]\Big\}.\end{aligned}$$

Combining the above inequalities yields $\overline{V}_t^A - \underline{V}_t^A \leq [1-k]\Big[\mathbb{E}^A[\overline{V}_{t+1}^A|m,h^t] - \mathbb{E}^A[\underline{V}_{t+1}^A|m,h^t]\Big]$.

If $T=\infty$ then, for any $\epsilon > 0$, there exists $t+t'$ such that $(1-k)^{t'} < \epsilon$, thus, $\overline{V}_t^A - \underline{V}_t^A \leq [1-k]^{t'}\Big[\overline{V}_{t+t'}^A - \underline{V}_{t+t'}^A\Big] < \epsilon$. On the other hand, if $T < \infty$ then $\overline{V}_T^A - \underline{V}_T^A = 0$.

Finally, because A has proposal power in every period, we must have $V_t^B(h^t) = 1 - p_t^B(h^t) - \frac{c_B}{k}$. Assume not, so there exists a SPE in which $V_t^B(h^t) > 1 - p_t^B(m) - \frac{c_B}{k}$ for some history h^t. Thus, there must exist some period $t' > t$ and history $h^{t'}$ at which the states settle and A offers x such that $1-x > 1-p_t^B(h^t) - \frac{c_B}{k}$. However, at $h^{t'}$ A could deviate and instead offer x' such that $1-x' = V_{t'}^B(m')$, which B must accept in equilibrium, a contradiction.

□

Lemma 1. *Assume $k > c_A + c_B$. If optimism is sufficiently large then the states fight.*

Proof. At $p_t^B = 0$ state A's payoff from settling is $\frac{c_B}{k}$ and its payoff from fighting is bound below by its payoff from fighting every period, which is given by $p_t^A - \frac{c_A}{k}$, thus, $k > c_A + c_B$ implies that A fights when $p_t^B = 0$. □

Lemma A.2. *The term $(1-k)\mathbb{E}_A[p_{t+1}^B] - p_t^B$ is single-peaked in p_t^B.*

Proof. Differentiating yields

$$\frac{\partial \mathbb{E}_A[p_{t+1}^B]}{\partial p_1^B} = \frac{\sigma(1-\sigma)\Big(p_t^B(2\sigma-1)^2 + p_t^A(1-p_t^B)(2\sigma-1)^2 + (1-\sigma)\sigma\Big)}{(\pi_t^B)^2(1-\pi_t^B)^2}$$

Thus, $(1-k)\mathbb{E}_A[p_{t+1}^B] - p_t^B$ is increasing in p_t^B if and only if

$$(1-k)\frac{\sigma(1-\sigma)\Big(p_t^B(2\sigma-1)^2 + p_t^A(1-p_t^B)(2\sigma-1)^2 + (1-\sigma)\sigma\Big)}{(\pi_t^B)^2(1-\pi_t^B)^2} - 1 \geq 0 \qquad (1)$$

If $k > \frac{p_t^A(1-4\sigma(1-\sigma))}{\sigma(1-\sigma)+p_t^A(1-4\sigma(1-\sigma))}$ then equation 1 never holds. If $k < \frac{p_t^A(1-4\sigma(1-\sigma))}{\sigma(1-\sigma)+p_t^A(1-4\sigma(1-\sigma))}$ then there exists a cutoff $\tilde{p}$ such that inequality 1 holds for $p_t^B < \tilde{p}$ and does not for $p_t^B > \tilde{p}$.[37] □

Lemma A.3. $V_{t+1}^A(m+1) \geq V_t^A(m)$.

Proof. We prove the lemma by induction. First, let $T = 2$. In this case, $V_1^A(0) = \max\{p_1^B + \frac{c_B}{k}, kp_1^A - c_A + (1-k)(p_1^A - \frac{c_A}{k})\}$ and $V_2^A(1) = \max\{p_2^B(1) + \frac{c_B}{k}, p_2^A(1) - \frac{c_A}{k}\}$. Clearly, $V_1^A(0) \leq V_2^A(1)$. Now assume that this inequality holds for $T = \bar{t}$. We show that it holds for $T = \bar{t}+1$. Given any history h_1, consider the subgame of length $\bar{t}$ starting at time $t = 2$. By the induction assumption, the inequality holds once we are in the subgame starting at $t = \bar{t}$. Therefore, $V_{t+1}^A(m+1) \geq V_t^A(m)$ for all $t \geq 2$ and m. Thus, we just need to show $V_0^A(0) \leq V_1(1)$. This holds if

$$\max\Big\{p_2^B(1) + \frac{c_B}{k}, kp_2^A(1) - c_A + (1-k)\mathbb{E}[V_3^A|m=1]\Big\} - \max\Big\{p_1^B + \frac{c_B}{k}, kp_1^A - c_A + (1-k)\mathbb{E}[V_2^A|m=0]\Big\} \geq 0.$$

We have $p_1^A(1) \geq p_0^A$ and by the induction assumption $\mathbb{E}[V_3^A|m=1] \geq \mathbb{E}[V_2^A|m=0]$. Since there is a probability k that fighting is decisive, payoffs are continuous in the limit and so the inequality holds at $T = \infty$. □

Lemma A.4. $\mathbb{E}_A[V_{t+1}^A|m] < p_A^t(m) + \frac{c_B}{k}$.

Proof. We prove it by induction on T. First, let $T = 2$. In this case, at time $t = 1$ we have $\mathbb{E}[V_2^A] = \pi_t^A \max\{p_2^A(1) - \frac{c_A}{k}, p_2^B(1) + \frac{c_B}{k}\} + (1-\pi_t^A)\max\{p_2^A(0) - \frac{c_A}{k}, p_2^B(0) + \frac{c_B}{k}\} < p_1^A + \frac{c_B}{k}$, since $p_t^B(m) < p_t^A(m)$ for all t and m.

Now assume the inequality holds for a game of length $T = \bar{t}$. We show that it holds for $T = \bar{t}+1$. Given any history h_1, consider the subgame of length $\bar{t}$ starting at time $t = 2$. By the induction assumption, the inequality holds once we are in the subgame

[37] The associated Mathematica file derives this explicitly.

starting at $t = 2$. Thus, $\mathbb{E}_A[V_{t+1}^A] < p_t^A(m) + \frac{c_B}{k}$ for all $t \geq 2$. Thus, we just need to show $\mathbb{E}_A[V_2^A] < p_1^A + \frac{c_B}{k}$. By the induction assumption we have $p_2^A(1) + \frac{c_B}{k} > \mathbb{E}_A[V_3^A|m = 1]$ and $p_2^A(0) + \frac{c_B}{k} > \mathbb{E}_A[V_3^A|m = 0]$.

Thus, in the first period,

$$\begin{aligned}\mathbb{E}_A[V_2^A] &= \pi_t^A \max\left\{p_2^B(1) + \frac{c_B}{k}, kp_2^A(1) - c_A + (1-k)\mathbb{E}_A[V_3^A|m = 1]\right\} \\ &+ (1-\pi_t^A)\max\left\{p_2^B(0) + \frac{c_B}{k}, kp_2^A(0) - c_A + (1-k)\mathbb{E}_A[V_3^A|m = 0]\right\} \\ &< \pi_t^A \max\left\{p_2^B(1) + \frac{c_B}{k}, kp_2^A(1) - c_A + (1-k)(p_2^A(1) + \frac{c_B}{k})\right\} \\ &+ (1-\pi_t^A)\max\left\{p_2^B(0) + \frac{c_B}{k}, kp_2^A(0) - c_A + (1-k)(p_2^A(0) + \frac{c_B}{k})\right\} \\ &\leq p_1^A + \frac{c_B}{k}.\end{aligned}$$

Since there is a probability k that fighting is decisive, payoffs are continuous in the limit and so the inequality holds at $T = \infty$. □

Define $\phi^n(p^i) = \frac{\sigma^n p^i}{\sigma^n p^i + (1-\sigma)^n(1-p^i)}$. That is, $\phi^n(p^i)$ gives player i's belief that $\omega = A$ after the states fight n rounds and A wins every battle, given initial belief p^i.

Lemma A.5. *Given any $p' > 0$ and $n < \infty$, $\phi^n(p) < p'$ for all p sufficiently small.*

Proof. Fix n. We have that $\phi^n(p)$ is strictly increasing in p. Additionally, $\lim_{p\to 0}\phi^n(p) = 0$. Since $\phi^n(p)$ is continuous in p, for any $p' > 0$ it is the case that $\phi^n(p) < p'$, for all p sufficiently small. □

Lemma A.6. *If $p_t^B(m)$ is sufficiently small and $p_t^B(m) + \frac{c_B}{k} < kp_t^A(m) - c_B + (1-k)\mathbb{E}_A[V_{t+1}^A|m, t]$ then $p_{t+1}^B(m+1) + \frac{c_B}{k} < kp_{t+1}^A(m+1) - c_B + (1-k)\mathbb{E}[V_{t+2}^A|m+1, t+1]$.*

Proof. By Lemma A.5 for any (m, t) there exists a $p_t^B(m)$ sufficiently small such that $p_{t+1}^B(m+1) - p_t^B(m) < k\Big(p_{t+1}^A(m+1) - p_t^A(m)\Big)$. For the result to hold we need

$$\begin{aligned}&\left[p_{t+1}^B(m+1) + \frac{c_B}{k}\right] - \left[p_t^B(m) + \frac{c_B}{k}\right] \\ &< \Big[kp_{t+1}^A(m+1) - c_B + (1-k)\mathbb{E}[V_{t+2}^A|m+1, t+1]\Big] - \Big[kp_t^A(m) - c_B + (1-k)\mathbb{E}_A[V_{t+1}^A|m, t]\Big].\end{aligned}$$

This simplifies to

$$p_{t+1}^B(m+1) - p_t^B(m) < k\Big(p_{t+1}^A(m+1) - p_t^A(m)\Big) + (1-k)\Big(\mathbb{E}[V_{t+2}^A|m+1, t+1] - \mathbb{E}_A[V_{t+1}^A|m, t]\Big),$$

which holds by Lemma A.3 and the assumption that $p_t^B(m)$ is sufficiently small. □

Lemma A.7. *Let p^* be a maximizer of $\mathbb{E}_A[p^B_{t+1}] - p^B_t$, where $\mathbb{E}_A[p^B_{t+1}] = \phi^1(p^B_t)$. We have $p^* \in (0, p^A_t)$.*

Proof.

$$\frac{\partial}{\partial p^B_t}\Big\{\mathbb{E}_A[p^B_{t+1}] - p^B_t\Big\} = \pi^A_t \frac{\sigma(1-\sigma)}{(1-\sigma+p^B_t(2\sigma-1))^2} + (1-\pi^A_t)\frac{\sigma(1-\sigma)}{(\sigma - p^B_t(2\sigma-1))^2} - 1$$
$$\lim_{p^B_t \to p^A_t} \frac{\partial}{\partial p^B_t} = \sigma(1-\sigma)\Big[\frac{1}{\pi^A_t} + \frac{1}{1-\pi^A_t}\Big] - 1 < 0,$$
$$\lim_{p^B_t \to 0} \pi^A_t \frac{\sigma}{1-\sigma} + (1-\pi^A_t)\frac{1-\sigma}{\sigma} - 1 > 0,$$

where the inequalities hold by $p^A_t \in (0,1)$ and $\sigma \in (1/2, 1)$. □

Proposition 1. *Assume $c_A + c_B > k$.*

1. *In every period t there exists $\underline{p}_t$ and $\overline{p}_t$, with $0 < \underline{p}_t < \overline{p}_t < p^A_t$, such that in every SPE if $p^B_t \in [0, \underline{p}_t]$ or $p^B_t \in [\overline{p}_t, p^A_t]$ then the states come to a peaceful agreement.*

2. *Assume $c_A + c_B$ is sufficiently small and σ is sufficiently large. For every $p^A_t > p^* + c_A + c_B$ there exists an open interval $\mathcal{I}_t \subseteq [\underline{p}, \overline{p}]$ such that in every SPE if $p^B_t \in \mathcal{I}_t$ then the states fight.*

Proof. First, we show that if $p^B_t(m)$ is sufficiently close to $p^A_t(m)$ then the states settle in every SPE. By Lemma A.4 State A's expected utility for fighting is bound above by fighting today and having B's posterior move to $p^A_t(m)$. Therefore, settling must be preferred to fighting when $p^B_t(m) + \frac{c_B}{k} \geq kp^t_A(m) - c_A + (1-k)(p^A_t(m) + \frac{c_B}{k})$, which rearranges to $c_A + c_B > p^A_t(m) - p^B_t(m)$, which always holds for $p^B_t(m)$ sufficiently close to $p^A_t(m)$.

Second, we argue that if p^B_t is sufficiently close to 0 then there must be agreement. Assume not. Then by Lemma A.6, A's continuation value is bound above by the payoff for fighting and winning every battle until $p^B_{t+1}(m+1) - p^B_t(m) > k\Big(p^A_{t+1}(m+1) - p^A_t(m)\Big)$, denote the number of battles for this by $N(p^B_t)$. However, by Lemma A.5, for any finite number of rounds n, we can find p^B_t sufficiently small such that $n < N(p^B_t)$. In particular, $\lim_{p^B_t \to 0} N(p^B_t) = \infty$ and A's continuation value for fighting goes to $p^A_t - \frac{c_A}{k}$ and payoff for settling goes to $\frac{c_B}{k}$. Thus, in the limit, A must strictly prefer to settle by assumption that $k < c_A + c_B$.

Finally, a sufficient condition for state A to fight is that A prefers to fight this period and then settle in the next period, regardless of the outcome, over settling this period. This holds if and only if

$$kp_t^A(m) - c_A + (1-k)\Big(\pi_t^A p_{t+1}^B(m+1) + (1-\pi_t^A)p_{t+1}^B(m) + \frac{c_B}{k}\Big) > p_t^B(m) + \frac{c_B}{k}$$
$$\Leftrightarrow k\Big(p_t^A(m) - \mathbb{E}_A[p_{t+1}^B]\Big) + \mathbb{E}[p_{t+1}^B] - p_t^B > c_A + c_B. \tag{2}$$

By the envelope theorem we have

$$\frac{\partial}{\partial\sigma}\Big\{\mathbb{E}_A[p^*] - p^*\Big\} = \frac{(1-p^*)p^*(2\sigma-1)(p_t^A - p^*)}{(p^*+\sigma-2p^*\sigma-1)^2(p^*+\sigma-2p^*\sigma)^2} > 0.$$

As $\sigma \to 1$ we have $\mathbb{E}_A[p^*] = p_t^A$. Thus, if $p_t^A - p^* > c_A + c_B$ then for all σ sufficiently close to 1 inequality (2) holds. □

Proposition 2. *The maximum number of periods that A expects to fight in equilibrium is maximized at an interior σ.*

Proof. For σ sufficiently small learning is slow and the states always settle. If σ is sufficiently large then states will fight at most once. Thus, the maximum number of periods of fighting can always be achieved by an interior value of σ. □

Proposition 3. *The maximum number of battles that state A expects to fight is bound above by:*

$$\left\lfloor \frac{\ln\left(\frac{c_A+c_B-k}{c_A+c_B}\right)}{\ln(1-k)} \right\rfloor.$$

1. *Increasing c_A or c_B decreases the upper bound.*

2. *If $k \to c_A + c_B$ the upper bound goes to infinity.*

Proof. A's payoff from fighting for n periods is bound above by its payoff for fighting for n periods and settling for 1. Thus, settling is always preferred to fighting for n periods when

$$\frac{1-(1-k)^n}{k}(k-c_A) + (1-k)^n(1+\frac{c_B}{k}) < p^B + \frac{c_B}{k}$$
$$\Leftrightarrow n > \frac{\ln\left(\frac{c_A+c_B-k(1-p^B)}{c_A+c_B}\right)}{\ln(1-k)}.$$

The RHS of the inequality is decreasing in p^B. Thus, it is maximized at $p^B = 0$, which yields the bound in the lemma and implies that state A always prefers to settle if rather than fight for n periods, with $n > \left\lfloor \frac{\ln\left(\frac{c_A+c_B-k}{c_A+c_B}\right)}{\ln(1-k)} \right\rfloor$. Therefore, state A never expects to fight for more than $\left\lfloor \frac{\ln\left(\frac{c_A+c_B-k}{c_A+c_B}\right)}{\ln(1-k)} \right\rfloor$ periods in equilibrium.

The proofs of parts 1 and 2 follow immediately by differentiating and by taking the limit as $k \to c_A + c_B$. $\square$

Proposition 4. *Assume T is sufficiently large. There exists $t^* < \infty$ such that if $t = t^*$ then $p_t^B \in [\overline{p}_t, p_t^A]$ and the states settle.*

Proof. Consider the sequence of player i's beliefs $(p_1^i, p_2^i...)$. Recall, by the standard properties of Bayesian learning, we have for each player i that $\mathbb{E}_i[p_{t+1}^i | p_1^i, ..., p_t^i] = p_t^i$. Then, by Lemma 2.1 in Chamley (2004), if $\omega = A$ then p^i converges to 1, otherwise, if $\omega = B$ then p^i converges to 0. Thus, there exists a t^* such that if $t > t^*$ then $p_t^A - p_t^B < c_A + c_B$ and the states settle. $\square$

Proposition 5. *Assume T is finite and $k < c_A + c_B$. The set of p_1^B that lead to conflict in the game of length T is a subset of the set of p_1^B that lead to conflict for the game of length $T + 1$.*

Proof. In either game the payoff from settling in period t is $p_t^B + \frac{c_B}{k}$. Let the payoff from fighting in game T be $kp_t^A - c_A + (1-k)\mathbb{E}_A^T[V_{t+1}^A]$ and the game in $T+1$ be $kp_t^A - c_A + (1-k)\mathbb{E}_A^{T+1}[V_{t+1}^A]$. Thus, the set of p_t^B such that the states fight in game T is a subset of those in game $T+1$ if and only if

$$\mathbb{E}_A^T[V_{t+1}^A] \leq \mathbb{E}_A^{T+1}[V_{t+1}^A].$$

Consider period T. $\mathbb{E}_A^{T+1}[V_{t+1}^A] = p_t^A(m) - \frac{c_A}{k} \leq \pi_t^A \max\{p_{t+1}^A(m+1) - \frac{c_A}{k}, p_{t+1}^B(m+1) + \frac{c_B}{k}\} + (1-\pi_t^A)\max\{p_{t+1}^A(m) - \frac{c_A}{k}, p_{t+1}^B(m) + \frac{c_B}{k}\}$. Assume the desired inequality holds for period $\bar{t} + 1 < T$ onward. Consider period $\bar{t}$. We have

$$\begin{aligned}
&\mathbb{E}_A^T[V_{\bar{t}}^A] = \\
&\qquad \pi_{\bar{t}}^A \max\{kp_{\bar{t}+1}^A(m+1) - c_A + (1-k)\mathbb{E}_A^T[V_{\bar{t}+1}^A | m+1], p_{\bar{t}+1}^B(m+1) + \frac{c_B}{k}\} \\
&\qquad + (1-\pi_{\bar{t}}^A)\max\{kp_{\bar{t}+1}^A(m) - c_A + (1-k)\mathbb{E}_A^T[V_{\bar{t}+1}^A | m], p_{\bar{t}+1}^B(m) + \frac{c_B}{k}\} \\
&\qquad \leq \pi_{\bar{t}}^A \max\{kp_{\bar{t}+1}^A(m+1) - c_A + (1-k)\mathbb{E}_A^{T+1}[V_{\bar{t}+1}^A | m+1], p_{\bar{t}+1}^B(m+1) + \frac{c_B}{k}\} \\
&\qquad + (1-\pi_{\bar{t}}^A)\max\{kp_{\bar{t}+1}^A(m) - c_A + (1-k)\mathbb{E}_A^{T+1}[V_{\bar{t}+1}^A | m], p_{\bar{t}+1}^B(m) + \frac{c_B}{k}\} \\
&= \mathbb{E}_A^{T+1}[V_{\bar{t}}^A],
\end{aligned}$$

where the inequality holds by the induction assumption. $\square$

Proof for Example 1.

Proof. In period 2 the states settle if $\min\{1, p_2^B + \frac{c_B}{k}\} \geq p_2^A - \frac{c_A}{k}$ which always holds by $c_A + c_B > k$. Now consider the first period. Since the states always settle in the second period, state A makes a peaceful offer in the first period if

$$\min\{p_1^B + \frac{c_B}{k}\} \geq k - c_A + (1-k)\Big(\mathbb{E}_A[p_2^B] + \frac{c_B}{k}\Big)$$

If $p_1^B \geq 1 - \frac{c_B}{k}$ then the states settle. Next, assume $p_1^B < 1 - \frac{c_B}{k}$, we rearrange the above as

$$c_A + c_B - k \geq (1-k)\mathbb{E}_A[p_2^B] - p_1^B \tag{3}$$

If $p_1^B = 0$ then $(1-k)\mathbb{E}_A[p_2^B] - p_1^B = 0$ and inequality 3 holds. Likewise, if $p_1^B = 1$ then $(1-k)\mathbb{E}_A[p_2^B] - p_1^B = -k$ and inequality 3 still holds. By Lemma A.2 $(1-k)\mathbb{E}_A[p_2^B] - p_1^B$ is single-peaked in p_1^B which implies the first part of the result.

To complete the proof of part 2, let p^* be the maximizer of $(1-k)\mathbb{E}_A[p_2^B] - p_1^B$. By the envelope theorem we have that

$$\frac{\partial}{\partial p_1^B}[(1-k)\mathbb{E}_A[p_2^B(p^*)] - p^*] = \frac{(1-p^*)^2 p^*(2\sigma - 1)}{(\pi_1^B(p^*))^2(1-\pi_1^B(p^*))^2} > 0.$$

Finally, for $p_1^B > 0$ if $\sigma \to 1$ then the RHS of 3 goes to $1 - k - p_1^B$ which is strictly larger than $c_A + c_B - k$ for p_1^B sufficiently small by assumption that $c_A + c_B < 1$. $\square$

Proof for Example 2.

Proof. See Mathematica file. $\square$

A.2 Proofs for Robustness Extensions

Incomplete Information

Let $\tilde{\mu}$ be state A's belief that state B is the $\theta = p_B$ type at the start of the second period if B rejected A's offer in the first period.

In the second period, by assumption that $c_A + c_B > k$, there exists a peaceful settlement with each type of state B that both states prefer over fighting. Thus, state A faces the standard crisis bargaining tradeoff. It can offer $x = 1$, which is only accepted by the $\theta = 1$ type and yields expected utility $\tilde{\mu}(1 - \frac{c_A}{k}) + (1 - \tilde{\mu})$, or it can offer $x = p_2^B + \frac{c_B}{k}$, which is accepted by both types and thus yields payoff $p_2^B + \frac{c_B}{k}$. As such, state A makes the risky offer if $\tilde{\mu} < \frac{k(1-p_2^B)-c_B}{c_A}$, and otherwise makes the safe offer.

Therefore, the $\theta = p_B$ type's expected utility from rejecting state A's offer in the first period is $k(1-p_1^B) + (1-k)(1-\mathbb{E}_B[p_2^B] - c_B) - c_B = 1 - p_B - c_B$. Thus, in the first period it must always accept any offer such that $x \leq p_B + \frac{c_B}{k}$. The $\theta = 1$ type's expected utility from rejecting state A's offer in the first period depends on the second period offer. Recall that $\theta = 1$ type expects to always lose the battle. Thus, if after rejecting $\tilde{\mu} > \frac{k(1-p_2^B(1))-c_B}{c_A}$ then it's expected utility is $-c_B$, while if $\tilde{\mu} < \frac{k(1-p_2^B(1))-c_B}{c_A}$ then its expected utility is $(1-k)(1-p_2^B(1) - \frac{c_B}{k}) - c_B$, where $p_2^B(1)$ is the probability the $\theta = p_1^B$ type believes that $\omega = A$ after losing one battle. Thus, in the first period, the $\theta = 1$ type must always accept any offer x such that $x \leq p_2^B(1) + \frac{c_B}{k} + k(1 - p_2^B(1))$

Let $p_1^B \in (\underline{p}, \overline{p})$, we show that in any equilibrium the probability of war is at least μ. For a contradiction, assume there is an equilibrium in which the $\theta = p_1^B$ type accepts A's offer in the first period. This implies the $\theta = 1$ type accepts as well. Thus, A's payoff is at best $p_1^B + \frac{c_B}{k}$. However, if A deviates to $x = p_2^B(1) + \frac{c_B}{k} + k(1 - p_2^B(1))$ then the $\theta = 1$ type must accept and the $\theta = p_B$ type rejects. Consequently, in period 2 state A believes that B is the $\theta = 1$ type. This improves A's payoff from the $\theta = 1$ type accepting and, by $p \in (\underline{p}, \overline{p})$, A prefers its expected payoff from fighting over $p_1^B + \frac{c_B}{k}$. Thus, this is a profitable deviation and contradicts that this is an equilibrium.

Next, consider $p \notin [\underline{p}, \overline{p}]$. We show that in every equilibrium the probability of war is at least μ. Towards a contradiction, assume there is an equilibrium in which the $\theta = 1$ type rejects A's first period offer, x_1^*. This implies that the $\theta = p_1^B$ type rejects in equilibrium as well. Thus, A has posterior belief $\tilde{\mu} = \mu$ in the second period.

First, assume that in the second period state A makes the safe offer after winning a battle. Then A's equilibrium payoff is given by $k - c_A + (1-k)(p_2^B + \frac{c_B}{k})$. The $\theta = 1$ type's expected utility from rejecting is $(1-k)(1 - p_2^B - \frac{c_B}{k}) - c_B$. Thus, if A deviates and offers $x = p_2^B + \frac{c_B}{k} + k(1 - p_2^B)$ then B accepts. This yields A a payoff $\mu_0(k - c_A + (1-k)(p_2^B + \frac{c_B}{k})) + (1-\mu_0)(p_2^B + \frac{c_B}{k} + k(1-p_2^B))$ which is strictly greater than $k - c_A + (1-k)(p_2^B + \frac{c_B}{k})$ by $c_A > 0$.

Second, assume that in the second period state A makes the risky offer after winning. The $\theta = 1$ type's expected utility from rejecting is $-c_B$. However, this implies that the $\theta = 1$ type would accept any offer x_1^*, which contradicts that the $\theta = 1$ type rejects in equilibrium.

Finally, for $p \notin [\underline{p}, \overline{p}]$ we show that the probability of war in every equilibrium must be 0 for μ sufficiently. Assume not, so at least one type rejects. Thus, at least the $\theta = p_1^B$ type must reject A's offer, and A's equilibrium payoff is bound above by assuming that $\theta = 1$ is willing to accept $x = 1$ in the first period. Thus, A's equilibrium expected utility in the first period is at most $\mu(k+(1-k)(\mathbb{E}_A[p_2^B]+\frac{c_B}{k}-c_A)+(1-\mu)$. However, for μ sufficiently close to 1 this term is $\approx k+(1-k)(\mathbb{E}_A[p_2^B]+\frac{c_B}{k}-c_A$, which is strictly less than $p_1^B+\frac{c_B}{k}$ by $p \notin [\underline{p},\overline{p}]$. Since the $\theta = p_1^B$ type always accepts $p_1^B+\frac{c_B}{k}$ state A has a profitable deviation, which contradicts that there is an equilibrium with a positive probability of war.

Disagreement about the Decisiveness of Conflict

In the last period state B accepts any offer x such that $1-x \geq 1-p-\frac{c_B}{k_2^B}$. Thus, the best acceptable offer that A can make is $x = p+\frac{c_B}{k_2^B}$. State A prefers to make this offer over fighting if and only if $p+\frac{c_B}{k_2^B} > p-\frac{c_A}{k_2^A}$, which always holds.

Therefore, in the first period, in equilibrium State B accepts an offer x if and only if

$$\begin{aligned}
&1-x \geq k_1^B\Big(1-p\Big)+(1-k_1^B)\Big(1-p-\frac{c_B}{k_2^B}\Big)-c_B\\
&\Leftrightarrow p+(1-k_1^B)\frac{c_B}{k_2^B}+c_B \geq x.
\end{aligned}$$

Thus, in equilibrium, state A either makes the best acceptable offer, $x = \min\{1, p+(1-k_1^B)\frac{c_B}{k_2^B}+c_B\}$, or makes an offer that leads to war. Clearly, State A offers $x=1$ and settles if $p+(1-k_1^B)\frac{c_B}{k_2^B}+c_B \geq 1$. Moving forward, assume $p+(1-k_1^B)\frac{c_B}{k_2^B}+c_B < 1$, which always holds for c_B sufficiently small. State A prefers to fight if and only if

$$\begin{aligned}
&k_1^A p+(1-k_1^A)\Big(p+\frac{c_B}{k_2^B}\Big)-c_A > p+(1-k_1^B)\frac{c_B}{k_2^B}+c_B\\
&\Leftrightarrow \frac{k_1^B-k_1^A}{k_2^B} > \frac{c_A+c_B}{c_B} = 2,
\end{aligned}$$

where $\frac{c_A+c_B}{c_B}=2$ follows from our simplification that $c_A=c_B$. The conditions to obtain fighting versus peace then follow straightforwardly from analyzing the inequality $\frac{k_1^B-k_1^A}{k_2^B} > 2$. Specifically, we obtain

$$\begin{aligned}
&k_* = 1/3\\
&k^* = \sqrt{2\underline{k}(1-\underline{k})}\\
&\overline{q}_A = \frac{3\overline{k}+2\underline{k}-3}{\overline{k}-\underline{k}}+\sqrt{8\frac{1-\underline{k}-\overline{k}+\underline{k}\overline{k}}{(\overline{k}-\underline{k})^2}}.
\end{aligned}$$

The expressions for $\underline{q}_B$ and $\overline{q}_B$ can similarly be obtained, but are lengthy and not insightful, so we omit them. Further, notice this inequality does not depend on c_B, thus we can take it as small as required for $p + (1 - k_1^B)\frac{c_B}{k_2^B} + c_B < 1$ to hold. In particular,

$$c_B < \frac{(1-\overline{k})\overline{k} - (1-\overline{k}-\underline{k})(\overline{k}-\underline{k})q_1^B}{1-\overline{k}+(1+\underline{k}-\overline{k})(\overline{k}-\underline{k})q_1^B+(\overline{k}-\underline{k})^2(q_1^B)^2}.$$

A.3 Example of Fanatical Peace with More Complex Worldviews

In the model, we have simplified "worldviews" to a single prior belief about ω. Of course, by worldview, scholars ordinarily mean something more expansive—at the very least, they mean something that shapes the way new information is received and incorporated. The following example is heavily-stylized and very simple. It is intended to show how fanatical peace might arise with more complex information structures.

Suppose there are two worldviews, Marxism and Capitalism. We'll assume that player A (who has the bargaining power) is always a Capitalist, but B can be either a Capitalist or a Marxist. Capitalists believe that $P(s_t = \omega) = .75$. Marxists, though, believe that $P(s_t = A; \omega = A) = .9$ but $P(s_t = A; \omega = B) = .4$, i.e. that the Capitalist imperialists often win battles even when they're destined to lose the war. We'll confine the game to just two periods, $T = 2$. Suppose $c = .025$ and $p_1^A = .8$. It is straightforward to see that, when A is facing another Capitalist, it fights under more circumstances than when it is facing a Marxist. For instance, if $p_1^B = .75$, then A's utility from negotiating an immediate settlement is .8; if B is a Capitalist, then A's expected utility to fighting at least one battle is .805; if B is a Marxist, then A's expected utility to fighting at least one battle is .796. We thus get a fanatical peace: peace is more feasible (in some circumstances) between actors of opposed cultural/ideological beliefs than between ones of similar beliefs.

Again, this is a heavily-stylized example. Presumably, real actors would attach some probability to the truth of one worldview (here, capitalism) and some to another (Marxism), and these estimates would also update in light of new information. That must be left, though, to future work.