Mutual Optimism in the Bargaining Model of War

Mark Fey*

Kristopher W. Ramsay†

December 19, 2016

Abstract

Mutual optimism is often considered an explanation for war, but recent research has challenged that claim in the simplest setting where choices are between fighting and settling but bargaining is not explicitly analyzed. How might we think about mutual optimism as an explanation for war when the settlements are the result of endogenous bargaining? Here we consider several alternative definitions of the occurrence of mutual optimism in crisis bargaining games. In each case, we find that mutual optimism is at best a partial explanation for war. But in no case is it necessary and sufficient, and some times it is neither. In fact, we show that war can occur even when both sides are almost completely pessimistic when there is uncertainty about the distribution of power.

*Professor, Department of Political Science, 109E Harkness Hall, University of Rochester, Rochester, NY 14627. phone: 585.275.5810, email: mark.fey@rochester.edu.

†Associate Professor, Department of Politics, 033 Corwin Hall, Princeton University, Princeton, NJ 08544. phone: 609.258.2960, email: kramsay@Princeton.edu.

The question of why states fight costly wars when less costly negotiated settlements are available is central to the politics between states. One long standing explanation, that has received much attention in conflict studies is associated with "mutual optimism" on the eve of conflict (Blainey 1988, Wagner 1994, Kim & Bueno de Mesquita 1995, Van Evera 1999). As most clearly expressed by Blainey (1988), the mutual optimism explanation states that when two countries have private estimates of their ability to prevail in a war, these estimates may preclude a peaceful settlement of the dispute. More specifically, if both countries are optimistic about their prospects in case of war—perhaps because both sides believe they are likely to prevail—then no proposed settlement will satisfy both sides. Since one side or the other rejects any proposed settlement, war results. As a consequence, "wars often occur when both sides are very optimistic about their chances of victory" (Leventoglu & Tarar 2008, 533).

Since at least Fearon (1995), the bargaining theory of war has offered another prominent explanation for war that also is intimately tied to uncertainty. In this theory, war is seen as a bargaining failure and the reason for this failure is uncertainty on the part of the proposer. Increasing the size of a demand leads to a better settlement for the proposer, but also increase the risk of bargaining failure and war. In many cases, the optimal risk-reward tradeoff, as it is known, involves running the risk of war.

Although the mutual optimism and bargaining failure explanations are closely related, the connection between them has not been fully explored. In this paper we consider the basic question of how does optimism that arises from private information about military capabilities and the probability of victory in war affects crisis bargaining behavior? One issue that complicates any attempt to answer this question is that it is somewhat unclear how to define the occurrence of "mutual optimism" in a bargaining model. We deal with this complication by considering several alternative definitions for mutual optimism.

Specifically, we consider four possible criteria for mutual optimism. First, motivated by the work of Slantchev & Tarar (2011), we examine the possibility of mutual optimism in a model in which uncertainty is one-sided. As we discuss later, ultimately we conclude that this is an unsatisfactory representation and instead posit that there must, in fact, be mutual uncertainty in order to admit the possibility of mutual optimism. Second, we consider defining mutual optimism as occurring when the private information of both sides is that they are strong rather than weak. Third, we refine this idea by identifying mutual optimism with situations in which the two sides have incompatible beliefs, in the sense that the two sides' estimates of the probability of winning sum to more than one. Finally, we consider mutual optimism as occurring when the private information of the two sides is such that there is no peaceful settlement that the two sides would both prefer to fighting.[1]

For each of these possible conceptualizations, we evaluate whether mutual optimism is necessary and/or sufficient for war in order to evaluate the mutual optimism explanation for conflict. In each case, we find that mutual optimism is not a complete explanation for war. For each definition, we give examples and general results that illustrate that either mutual optimism is not necessary (because war can occur in the absence of mutual optimism) or is not sufficient (because mutual optimism can occur in the absence of war).

But if mutual optimism is not a convincing explanation for war, what is the connection between uncertainty and conflict? One basic fact that holds in all crisis bargaining games is that stronger types are more likely to fight (Fey & Ramsay 2011). This is a reflection of the basic risk/reward tradeoff inherent in crisis bargaining. It may seem that this observation supports the mutual optimism argument, but it is important to recognize that this observation is a *relative* claim, while the mutual optimism argument is an *absolute* claim. To see the difference, note that incentives in crisis bargaining require that a somewhat weak type should be more likely to fight than a very weak type, while the mutual optimism argument would imply that neither weak type should fight at all (Banks 1990, Fey & Ramsay 2011).

More specifically, unlike the one-sided case in which the proposer does not possess private information, with mutual uncertainty the proposer's offers can be used to signal her strength. Thus, there is an incentive for strong proposers to make more demanding offers. Naturally, these offers are more likely to be rejected and so war is more likely occur between relatively stronger types. But again, this is a observation about relative optimism, rather than mutual optimism.

Indeed, this incentive for stronger types to makes offers that risk war is *always* present, regardless of the how optimistic these types are. To illustrate this point we show that war can occur even when both sides are almost completely pessimistic. In one such a situation, even though all types of country 1 have negative war payoffs, types that are slightly less pessimistic want to reveal this by making more demanding offers, even if this entails the possibility of war. A similar argument shows that the same risky behavior occurs when the best that country 1 can hope to do is win with a very small probability. So regardless of optimism or pessimism, relatively strong types of countries are willing to run the risk of war to signal this strength, and this is the fundamental cause of war in the bargaining model

[1]This definition is consistent with Wagner (1994) and Wittman (1979), who describe mutual optimism in this way. See also Werner & Yuen (2005).

with two-sided private information. Furthermore, this source of bargaining failure is absent from analysis with one-sided uncertainty.

Viewed most broadly, our work contributes to the theoretical literature on the causes of war, which is too large to survey here. More specifically, our work relates to the existing literature that investigates how uncertainty and optimism relate. For example, Wittman (1979) discusses the existence and non-existence of a bargaining range when there is uncertainty about the military balance and thus the probability of victory in war. This discussion is based on some bargaining logic, but it is not treated as a a formal bargaining problem. Similarly Wagner (2000) considers how private information about the probability of winning influences the incentives to fight. Fey & Ramsay (2007) consider a class of games in which the two sides have private information about the payoffs for fighting and settling and both sides must choose to fight in order for war to occur. Bargaining is precluded, though, as the terms of the peaceful settlement are not determined endogenously. They show that in such settings mutual optimism cannot lead to war, even with decision-makers who may face some types of bounded rationality. In a more recent article Wittman (2009) analyzes a game theoretic model of bargaining and war with two-sided uncertainty about the probability of victory, providing interesting and unexpected new results that relate to optimism. And, in response to Fey & Ramsay (2007), Slantchev & Tarar (2011) study a bargaining model with endogenous offers but with only one-sided uncertainty. They suggest that, according to their definition of mutual optimism, the occurrence of mutual optimism is both necessary and sufficient for war to occur. We move beyond this literature by considering two-sided uncertainty in a bargaining game and examining a variety of different conceptualizations of mutual optimism.

The paper proceeds in five sections. First we consider mutual optimism in a model with almost one-sided uncertainty to see if the simplification of using such a model where only one player has private information is a good (simple) approximation of an environment where both countries have private information about their military capabilities. The motivation for this exploration is recent results from Slantchev & Tarar (2011) that studies the effect of favorable priors on the probability of war in the bargaining model. We find that the answer is no. Even very small amounts of private information for the second country in the crisis bargaining game undoes their result that mutual optimism is necessary and sufficient for war.

Given this finding, we walk through different ways one might conceptualize mutual optimism in the bargaining model of war: mutual optimism is when only "strong types" fight each other, mutual optimism is a situation where private information of the two counties lead them to incompatible beliefs about the probability of victory in war, meaning that they sum to more than 1, and the situation where war payoffs, given private information, sum to more than 1 and, therefore, there is no peaceful settlement that both sides would prefer to fighting. In each case we show that mutual optimism is either not necessary for war, not sufficient for war, or both.

1 Mutual Optimism and Mutual Uncertainty

We begin our analysis by considering whether or not mutual optimism can be modeled in the context of one-sided uncertainty. Most notably, Slantchev & Tarar (2011) give an example of a take-it-or-leave-it bargaining game with one-sided incomplete information that claims to show that mutual optimism can be both necessary and sufficient for war. As the authors put it, "war occurs in the standard model *if, and only if, mutual optimism is present*, and hence the presence of mutual optimism is not just relevant, it in fact entirely determines the occurrence or nonoccurrence of war" (p. 139, emphasis in original). In their example, country 1 (the side that makes the offer) is uncertain about the strength of country 2, but country 2 has no uncertainty—it knows the true probability it will prevail in a war.

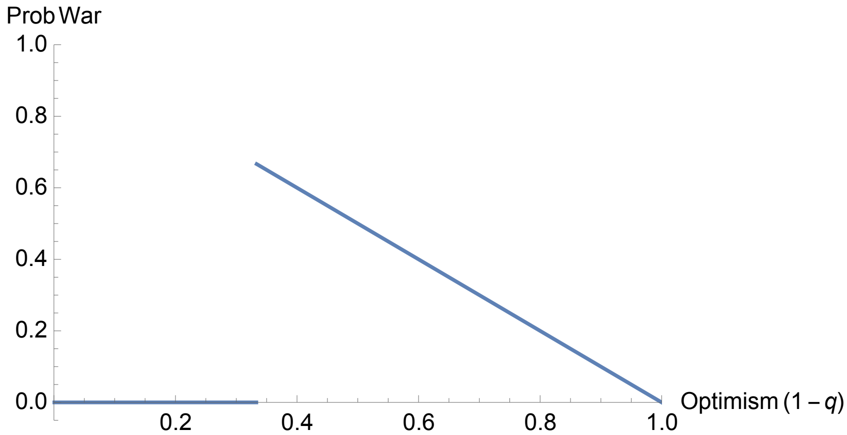
In this section, we present our view that a study of mutual optimism must permit both sides to have optimistic expectations drawn from private information and argue that the conclusion of Slantchev & Tarar (2011) that mutual optimism is both necessary and sufficient for war in their model is a very special and non-robust case given their definition of "mutual optimism." Moreover, we show that if we take their model with one-sided incomplete information and add an arbitrarily small amount of private information to the second side, the conclusion that mutual optimism is necessary and sufficient for war no longer holds. This fact is important because it means we cannot interpret the Slantchev & Tarar (2011) results as an approximation of a mutual optimism model with "almost" one-sided incomplete information.

In accordance with the substantive literature on mutual optimism, it seems clear to us that optimism or pessimism can only be the result of private information. How can a side be optimistic (or pessimistic) if it knows the true probability of winning in war? To us, this casts doubt on what can be learned about mutual optimism based on a model with private information on only one side. At most, such a model can only inform us as to the connection between unilateral optimism and war. Mutual optimism can only arise with mutual uncertainty.

Leaving this matter aside, there are additional difficulties with Slantchev & Tarar's (2011) claim that mutual optimism is necessary and sufficient for war in their example. Understanding these difficulties requires a brief description of their argument. First, the authors solve for the unique perfect Bayesian equilibrium of their example and show that if country 1's prior belief that country 2 is the strong type, denoted q, is greater than a threshold value k, then country 1 makes a generous offer that both types of country 2 accept, and if q is less than k, then country 1 makes a demanding offer that only the weak type accepts and the strong type rejects. Therefore, war occurs with positive probability if and only if $q \leq k$. Next, the authors discuss the meaning of optimism in their example. They posit that country 2 is optimistic if and only if it is the strong type. They then state that country 1 is optimistic when she is "sufficiently confident that she faces a weak opponent (when $q < k$)."

While it is perhaps defensible to view optimism as relating to how likely it is a side will prevail in war, which in this example depends on the strength of the opponent, Slantchev & Tarar (2011) provide no justification for why the threshold defining optimism should be at the value k. Indeed, there is no reason to use this value other than the fact that this is the threshold for making a demanding offer in equilibrium. This choice in the definition of optimism is problematic because it makes the claim that mutual optimism is necessary and sufficient for war essentially tautological. By *defining* optimism as occurring for precisely those values of q that lead to war, the conclusion that mutual optimism is necessary and sufficient for war is true by definition. In other words, because war occurs with positive probability if and only if $q < k$, it is not legitimate to simply define optimism as $q < k$ and claim this proves that war occurs if and only if mutual optimism occurs.

In addition to this, there is a second difficulty with the definition used by Slantchev & Tarar (2011). As q represents the prior probability that country 1 is facing a strong opponent, lower values of q reflect greater optimism on the part of country 1. As just noted, when q is less than the threshold value k, war occurs with positive probability. In fact, in this case, the probability of war is equal to q because country 1 makes an offer that the weak type peacefully accepts and the strong type rejects. But this creates a puzzling relationship between war and optimism, as defined by the authors. Specifically, as q becomes smaller, country 1 becomes more optimistic, but the probability of war goes down rather than up. At the extreme, as q goes to 1, country 1 is maximally optimistic (it believes with near certainty it faces a weak opponent) and yet the probability of war goes to zero. Clearly, this relationship runs counter to the claim that optimism is the explanation of war, at least using



Going left to right on the x-axis is the probability country 2 is weak $(1-q)$ and country 1 is becoming more optimistic. As country 1's belief that country 2 is weak goes up, whenever country 1 makes an offer that leads to war with positive probability, the probability of war goes down. In other words, as country 1 becomes more optimistic the probability of war goes down, if there is any risk of war at all.

Figure 1: Optimism and war with one-sided incomplete information.

the definition of Slantchev & Tarar (2011).

To visualize this point, Figure 1 shows the probability of war when the threshold $k = 2/3$. As q is the probability country 2 is strong, we can think of the prior that country 2 is weak, $1 - q$, as country 1's level of optimism. The figure represents the result with one-sided uncertainty. When the probability that country 2 is weak is small, there is no risk of war because country 1 makes an offer accepted by every type. If however, there is sufficient chance country 2 is weak, then country 1 makes an offer that risks war. This risk of war is maximized at the threshold value, when country 1 is the least "optimistic" type to make a risky offer, and decreases to zero when country 1 is the most optimistic. Thus, this figure illustrates the negative relationship that exists between optimism and the chance of war in this environment, instead of the expected positive relationship.

Forgetting for the moment the concerns just raised, can a bargaining model with one-sided uncertainty be viewed as an approximation to a model with mutual uncertainty? To analyze this possibility, we next modify the example of Slantchev & Tarar (2011) into a game in which both sides have private information in a way that insures our new game is still "close" to having one-sided incomplete information. We will show that the claim that war occurs if and only mutual optimism $(q < k)$ occurs fails to holds, even when the two-sided incomplete information is arbitrarily close to the original one-sided incomplete information example.

The model given in Slantchev & Tarar (2011) is a standard ultimatum offer game in which country 1 makes an offer that is either accepted by country 2 or rejected, leading to war. The cost of war to country i is denoted $c_i > 0$ and is commonly known. To define our game with two-sided incomplete information, we suppose that there is an arbitrarily small chance $\varepsilon > 0$ that the strength of country 1 varies slightly from that given in the one-sided incomplete information. In other words, with probability $1 - \varepsilon$, country 1 is a "standard" type and the parameters of the game are as in the example of Slantchev & Tarar (2011): if country 2 is strong then the probability that country 1 prevails in war is p_s and if country 2 is weak then this probability is $p_w > p_s$.[2] With probability ε, on the other hand, country 1 is a "variant" type. A variant type has a slightly different probability of winning a war. Specifically, when country 1 is a variant type, it prevails against a strong type of country 2 with probability $p_s + \gamma$ and against a weak type of country 2 with probability $p_w + \gamma$. We permit γ to be positive or negative, so the variant type may be slightly stronger or slightly weaker than the standard type of country 1. As this game has two-sided incomplete information, it has a large number of perfect Bayesian equilibria. Therefore, we focus on perfect Bayesian equilibria that satisfy the additional requirements of the D1 refinement (Cho & Kreps 1987).[3]

When we choose ε and γ to be close to zero, this game of two-sided incomplete information is "close" to the game of one-sided incomplete information given by Slantchev & Tarar (2011). We now prove that in such an arbitrarily close game, mutual optimism is not necessary and sufficient for war, even when using the definition of optimism given by Slantchev & Tarar (2011) that $q < k$. In other words, the claim of Slantchev & Tarar (2011) breaks down when we add an arbitrarily small amount of two-sided incomplete information.

Proposition 1 *For all sufficiently small values of ε and γ, there does not exist a perfect Bayesian equilibrium satisfying D1 in which mutual optimism is necessary and sufficient for war.*

Proof: Following Slantchev & Tarar (2011), we say country 2 is optimistic if and only if it is the strong type. We will show that for all sufficiently small values of ε and γ, there does not exist a perfect Bayesian equilibrium satisfying D1 such that war occurs if and only if country 2 is a strong type. In what follows we suppose that $\gamma > 0$. The case of $\gamma < 0$ follows from similar arguments.

[2]This notation differs from Slantchev & Tarar (2011) in that we define p to be the probability that country 1 wins instead of country 2.

[3]In the context of this bargaining game, the D1 refinement simply requires country 2 not believe that an unexpectedly strong demand is likely to come from the weaker type of country 1.

We proceed via a proof by contradiction. So suppose there exists such a perfect Bayesian equilibrium satisfying D1 such that war occurs if and only if country 2 is a strong type. We first note that this implies that both types of country 1 must be playing a pure strategy. To see this, note that our requirement that war occurs if and only if country 2 is a strong type means that, on the equilibrium path, the strong type of country 2 is rejecting all offers and the weak type is accepting all offers. But now suppose some type of country 1 is mixing over distinct offers x' and x''. Since both offers are accepted by the weak type and rejected by the strong type, whichever offer is higher gives country 1 a higher expected utility. But this violates the indifference condition for mixing. Therefore neither type of country 1 can be playing a mixed strategy.

From this, we know that the two types of country 1 are either playing a separating strategy or a pooling strategy. We can rule out a separating strategy by a similar argument as the above. Specifically, suppose the two types are making distinct offers. Since both offers are accepted by the weak type and rejected by the strong type, the type of country 1 making the small offer can gain by deviating and making the (higher) offer of the other type. So the only remaining possibility is a pooling strategy.

Before continuing the analysis, consider some offer $(x, 1-x)$ by country 1. It is clear that the strong type of country 2 will reject any offer $x > p_s + c_2 + \gamma$ and the weak type of country 2 will reject any offer $x > p_w + c_2 + \gamma$. Similarly, the strong type of country 1 will accept any offer $x < p_s + c_2$ and will reject any offer $x < p_w + c_2$.

Now suppose both types pool on the offer x^*. By assumption, the strong type of country 2 rejects this offer and the weak type accepts. This implies that $x^* \in [p_s + c_2, p_w + c_2 + \gamma]$. It is clear that $x^* \notin (p_s + c_2 + \gamma, p_w + c_2)$ because deviating to $x^* + \delta$ will be profitable for a sufficiently small positive δ. For the same reason we have $x^* \notin (p_w + c_2, p_w + c_2 + \gamma)$. Because $q < k$ we know that offering $p_s + c_2$ is not optimal in the one-sided game of incomplete information. For sufficiently small γ, then, it follows that it is not optimal to offer $x^* \in [p_s + c_2, p_s + c_2 + \gamma]$. The only two remaining possibilities, therefore, are $x^* = p_w + c_2$ and $x^* = p_w + c_2 + \gamma$.

Consider the offer $x^* = p_w + c_2 + \gamma$. In equilibrium, the weak type of country 2 accepts this offer and receives a payoff of $1 - x^*$. Rejecting this offer, however, gives the weak type of country 2 a payoff of

$$(1-\varepsilon)(1 - p_w - c_2) + \varepsilon(1 - p_w - \gamma - c_2) > 1 - x^*.$$

Therefore this offer cannot be a pooling equilibrium. The final possibility is $x^* = p_w + c_2$. But consider a deviation to $x' \in (p_w + c_2, p_w + c_2 + \gamma)$. As the deviant type has a higher war payoff, the D1 refinement requires that country 2 place probability one on such an offer coming from a deviant type. But this implies this offer will be accepted by the weak type. But then deviating to x' is profitable, so $x^* = p_w + c_2$ cannot be part of a pooling equilibrium. Therefore we have ruled out all possible strategies for country 1 and so no such perfect Bayesian equilibrium exists. ∎

So while Slantchev & Tarar (2011) claim that their example with one-sided uncertainty shows that mutual optimism can be necessary and sufficient for war, this result establishes that this claim is false in models with "almost one-sided" uncertainty. Specifically, the basis for the claim in Slantchev & Tarar (2011) is that with one-sided uncertainty, in equilibrium the strong type of country 2 always rejects the offer and fights while the weak type always accepts the offer. The proof of Proposition 1 shows that with almost one-sided uncertainty, there cannot be a perfect Bayesian equilibrium satisfying the D1 refinement in which this occurs. In fact, it is possible to show that in any perfect Bayesian equilibrium satisfying D1, the weak type of country 2 must reject the offer and fight with positive probability. Moreover, this probability goes to zero as the level of country 2's uncertainty goes to zero. Thus, the equilibrium in the model of Slantchev & Tarar (2011) is the limit of the equilibria satisfying D1 as country 2's uncertainty goes to zero. Importantly, however, mutual optimism is not necessary for war except in the limiting case and so the claim of Slantchev & Tarar (2011) only holds in the knife-edge case of exactly zero uncertainty for country 2.

It is also notable that a key step in the argument for Proposition 1 is that there cannot be a pooling equilibrium because the strong type of country 1 has an incentive to signal its strength by making a more demanding offer. We will return to this observation later and see that this behavior is present in all of the models we consider.

2 Mutual Optimism as Being Mutually Strong

Having argued that mutual uncertainty is a prerequisite for mutual optimism, we now proceed to consider definitions involving such mutual uncertainty. In the previous section, we considered a bargaining model with two-sided uncertainty in which the two sides each had two types, but the two sides had very uneven levels of uncertainty. Here we consider a similar model that also has mutual uncertainty, but that has equal levels of uncertainty on

	W	S
W	$(.5, .5)$	$(.5 - a, .5 + a)$
S	$(.5 + a, .5 - a)$	$(.5, .5)$

Figure 2: Probabilities of winning: (p_1, p_2)

both sides. In the context of this model, we consider a simple definition of mutual optimism as occurring when the two sides are mutually strong. We show that in this model, mutual optimism is not necessary for war. Thus, using the "mutually strong" definition, we see that mutual optimism is not a complete explanation for war.

As before, the basis for our bargaining model involves a take-it-or-leave-it offer by country 1, which country 2 either accepts or rejects. If an offer $(x, 1-x)$ is accepted by country 2, then countries 1 and 2 receive payoffs of x and $1 - x$, respectively. If country 2 rejects the offer, then war ensues and each country i receives its war payoff, which is given by $p_i - c$, where p_i is the probability that country i prevails in war and $c > 0$ is the cost of fighting a war.

There is two-sided uncertainty about the probability of winning. We suppose that each country is either *weak* or *strong*, so that the type space of country i is $T_i = \{W, S\}$. The probability that a given sides wins in war is a function of the strengths of both sides. We denote the probability that country 1 wins by $p_1(t_1, t_2)$ and therefore the probability that country 2 wins is given by $1 - p_1(t_1, t_2)$. We assume that this probability is symmetric in types. That is, $p_1(t_1, t_2) = 1 - p_1(t_2, t_1)$. This assumption encapsulates the idea that the probability of winning only depends on the relative strengths of the two sides and not on which one is labeled country 1 or country 2. Two conclusions follow from this symmetry assumption. First, if both sides are weak or both sides are strong, then the probability of either side winning the war is 1/2. That is $p_i(W, W) = p_i(S, S) = 1/2$ for $i = 1, 2$. Second, there exists a value $a > 0$ such that if a strong country faces a weak country in war, then the strong country wins with probability $1/2 + a$ (and the weak country wins with probability $1/2 - a$). Thus a reflects the advantage a strong country has over a weak adversary. These facts are summarized in Figure 2. Finally, in order to insure we have equal levels of uncertainty, we assume that each country believes it is equally likely that the other country is weak or strong.

As discussed above, we define mutual optimism to happen when the two sides are mutually strong. Thus, mutual optimism occurs at the type pair (S, S).

We now turn to the equilibria of this game. Because this a game with two-sided incomplete information with continuous action spaces, there are a number of perfect Bayesian equilibria. Our main result is that in all such equilibria, mutual optimism is not necessary for war. That is, although war occurs at the type pair (S, S), it also occurs at other type pairs.

Proposition 2 *Suppose $c + a < 1/2$ and $c < a/4$. In all perfect Bayesian equilibria of this model with symmetric uncertainty, mutual optimism is not necessary for war.*

Proof: In order to establish the proposition, we will show that in every perfect Bayesian equilibrium the Strong type of country 2 rejects the offers of both types of country 1.

In our proof, we will denote the strong type of country 1 by $1S$, the weak type of country 1 by $1W$, the strong type of country 2 by $2S$, and the weak type of country 2 by $2W$. To begin with, note that for $2S$, war gives a payoff of at least $1/2 - c$. Therefore $2S$ will reject any offer such that $1 - x < 1/2 - c$, which is equivalent to $x > 1/2 + c$. Likewise, for $2W$, war gives a payoff of at most $1/2 - c$. Therefore $2W$ will accept any offer such that $x < 1/2 + c$. One the other hand, $2W$ will reject any offer $x > 1/2 + a + c$. Because $c + a < 1/2$, this implies that both types of player 2 will reject an offer $x = 1$. For $1S$, offering $x = 1$ and having both types reject gives a payoff of $1/2(1/2 - c) + 1/2(1/2 + a - c) = 1/2 + a/2 - c$. If $1S$ offers $x \leq 1/2 - c$, then $2W$ will accept and the largest payoff possible is $1/2(x) + 1/2(1/2 - c) \leq 1/2 - c$. This is strictly lower than the payoff for $x = 1$, so $1S$ will not offer $x \leq 1/2 - c$ in equilibrium. Likewise, if $1S$ offers $x \in (1/2 - c, 1/2 + c)$, then $2W$ will accept and the largest payoff possible is $1/2(x) + 1/2(x) = x < 1/2 + c$. As $c < a/4$, this is strictly lower than the payoff for $x = 1$, so $1S$ will not make such an offer in equilibrium.

From this, we know that the only possible offer that $1S$ is willing to make that $2S$ is willing to accept is $x = 1/2 + c$. But note that $2S$ will accept this offer only if $1S$ is the only type making this offer. So could there be an equilibrium in which $1S$ offers $x = 1/2 + c$ with positive probability but $1W$ does not, and $2S$ accepts this offer? Given the strategies of country 1, clearly $2W$ will also accept $x = 1/2 + c$. From this, it is clear that $1W$ cannot be offering something less than $1/2 + c$ because deviating to $1/2 + c$ would be profitable. So thus $1W$ must offer something larger than $1/2 + c$. This offer must give a payoff of at least $1/2 + c$ even though $2S$ is rejecting it. But if this is the case, it is easy to check that $1S$ would want to deviate from $1/2 + c$ to this offer. So there can be no such equilibrium. Thus $2S$ rejects any offer that $1S$ makes.

Could $1W$ make an offer that $2S$ accepts? As $1S$ is not making this offer, $2S$ knows for sure that the offer comes from $1W$. Therefore, $2S$ will accept this offer only if $x \leq 1/2-a+c$. Clearly $1W$ will also accept such an offer. If $1W$ deviates to $1/2+c-\varepsilon$, then $2W$ accepts this offer and $1W$ receives a payoff of at least $1/2(1/2+c-\varepsilon)+1/2(1/2-a-c)$. For small enough ε, this is a profitable deviation. Therefore, $2S$ rejects all offers. ∎

The proof of the proposition shows that the strong type of country 2 rejects the offers of both types of country 1. Therefore when country 1 is weak and country 2 is strong, the outcome is fighting. But mutual optimism occurs only when both countries are the strong type. Therefore, mutual optimism is not necessary for war in this bargaining model.

As an aside, it is worth noting two technical features of this proposition. The first is the assumption that the cost of war c is not too high and the second is that the proposition holds for all perfect Bayesian equilibria of the model, not just those satisfying some refinement such as D1. These two features are related. The bounds on c allow us to simplify the proof as well as insuring that the conclusion holds for all equilibria. Alternatively, it can be shown through a more complicated argument that even with no restriction on c, the result continues to hold for all equilibria satisfying D1.

We can further strengthen our argument in this setting by considering this two-type symmetric model when the cost of war c is high enough that there is no mutual optimism present. In this case, as in Example 2, there is an equilibrium with a positive probability of war, even though mutual optimism is absent. Specifically, if $a < c$ in this model, there is no mutual optimism present but there exists a perfect Bayesian equilibrium that satisfies D1 in which the strong type of country 2 sometimes rejects the offer made by the strong type of country 1.[4] This makes our case that mutual optimism is not strongly connected to war onset even stronger.

3 Mutual optimism as incompatible beliefs

Our second alternative for conceptualizing mutual optimism is that it occurs when the two sides have incompatible beliefs. Specifically, we say that a type profile (t_1, t_2) has mutual optimism if the expected probability of winning a war for type t_1 of country 1 plus the expected probability of winning for type t_2 of country 2 is greater than one.[5] In other words,

[4]Details of this equilibrium are available from the authors on request.

[5]Some studying optimism and war have phrased the condition of mutual optimism as when both sides believe they are more likely to win, meaning that their assessment is that their probability of victory in war

the beliefs of the two sides about the expected outcome of the war are incompatible with each other.

As it happens, our analysis in the previous section can be applied to this conceptualization of mutual optimism as well. Of course, in the previous section, mutual optimism was defined as occurring when both sides were the strong type. But refer back to Figure 2 which gives the probabilities of winning for each type combination, since each type is equally likely, it is easy to see that the expected probability of winning is $1/2 - a/2$ for the W type and is $1/2 + a/2$ for the S type. From this it follows that the beliefs of the two sides are incompatible if and only if both sides are the strong type. Thus, for the symmetric two-type case considered in the previous section, mutual optimism defined as incompatible beliefs is the same as mutual optimism defined as being mutually strong.

Therefore, the proof of Proposition 2 carries over to the case of mutual optimism as incompatible beliefs. That is, in the symmetric two-type model under the same conditions as before, mutual optimism is not necessary for war in all perfect Bayesian equilibria. More specifically, we see that, in equilibrium, the strong type of country 2 chooses to fight against the weak type of country 1, even though these types have compatible beliefs.

4 Mutual optimism as the absence of a bargaining space

We now turn to our final conceptualization of mutual optimism. Here, we view mutual optimism as occurring if there is no peaceful agreement that both sides would prefer to fighting, based on their private information. Formally, this definition can be expressed by saying that mutual optimism holds for countries of types (t_1, t_2) if the sum of their ex ante expected payoffs from war, given their types, are greater than one.

To get a sense of what happens with this conceptualization, we begin with an example. In this example and the general analysis that follows, we move beyond the simple two type models of the previous sections and consider the case of a continuum of types of each country. In our example, suppose that the technology that determines the probability of victory for country 1 when the players types are t_1 and t_2 is

$$p(t_1, t_2) = \frac{1}{2} + \frac{t_1 - t_2}{4}$$

is greater than one-half. That is a special case of this definition, as the sum of these probabilities is greater than one.

and that each country's type is independently drawn from a uniform distribution on the interval $[0, 1]$. As a result, the probability of winning the war for country one ranges from 1/4 to 3/4. Suppose further that both countries' cost of war is 1/10.

We can then define country 1's assessment of their own war payoffs given their type as

$$\begin{aligned} W_1(t_1) &= \int_0^1 (\frac{1}{2} + \frac{t_1 - t_2}{4} - \frac{1}{10})\, dt_2 \\ &= \frac{11}{40} + \frac{t_1}{4} \end{aligned}$$

and for country 2,

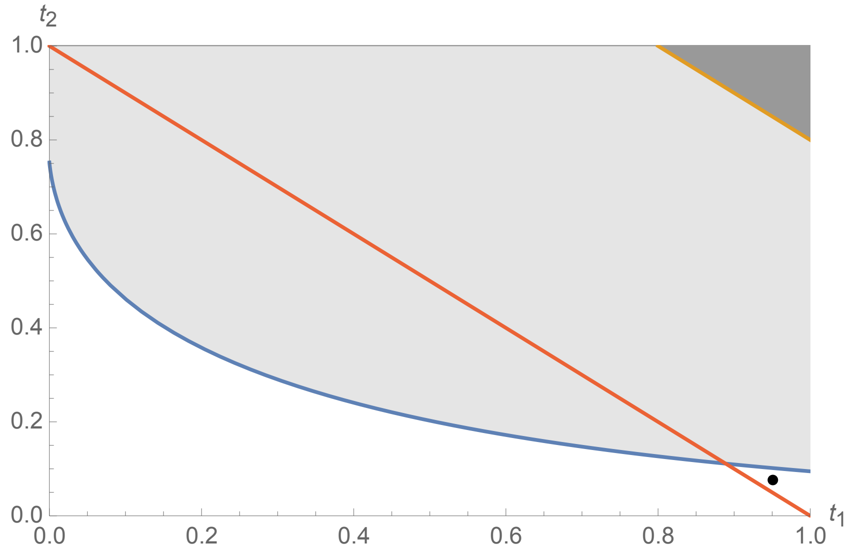
$$\begin{aligned} W_2(t_2) &= \int_0^1 (\frac{1}{2} + \frac{t_2 - t_1}{4} - \frac{1}{10})\, dt_1 \\ &= \frac{11}{40} + \frac{t_2}{4}. \end{aligned}$$

In this example, with types uniformly distributed and the symmetric function $p(t_1, t_2)$, we have mutual optimism whenever

$$\begin{aligned} W_1(t_1) + W_2(t_2) &= \frac{11}{20} + \frac{t_1 + t_2}{4} > 1 \\ t_1 &> \frac{9}{5} - t_2. \end{aligned}$$

As in the other sections, we solve for an perfect Bayesian equilibrium that satisfies the D1 refinement. The details of the solution are somewhat involved and so we present them in an appendix. The behavior of the two sides is easy to present visually, though. Figure 3 illustrates several features of this equilibrium. We can summarize the equilibrium outcome by a cut-line in the space of type pairs (t_1, t_2) such that the pairs on one side fight and those on the other settle. In the figure, the type pairs that fight are given by the grey region. The types that are mutually optimistic according to the condition above are given by the dark grey region. Finally, we also include the line given by $t_2 = 1 - t_1$, as it is clear that types above this line will have probabilities of winning that sum to more than one, which is the conceptualization of mutual optimism considered in the previous section.

This example allows us to immediately draw several conclusions about the relationship between mutual optimism and war. As is clear, there are is a large region in which players are not mutually optimistic, however defined, but they go to war nonetheless. Therefore we can conclude that mutual optimism is not necessary for war. In fact, if we use the definition



Grey region is the set of type pairs that go to war. The dark grey region are the type pairs that are mutually optimistic with respect to war payoffs. All points above the line $1 - t_1$ are those where the two sides' expected probabilities of victory sum to more than one, and therefore they are mutually optimistic with respect to the probability of victory in war. At the type pair $(.95, .075)$, marked by the black dot, countries are mutual optimistic regarding the probability of winning, but there is no war.

Figure 3: Type pairs that go to war.

of mutual optimism that the sum of victory probabilities, given their private information, are greater then one, then mutual optimism is also not sufficient. As shown in Figure 3, at the type pair $(.95, .075)$ there is just such incompatible beliefs regarding the probability of victory in war but there is also peace.

Of course, these conclusions are drawn from this one example. How general are these findings? Do they depend on the assumption of a continuum of types? To answer these questions in reference to conceptualizing mutual optimism as when war payoffs sum to more than one, in what follows we work with a very general formulation of information and payoffs. For example, we allow the two sides to have just two types, or some finite number of types, or a continuum of types. We also allow the distribution of these types to be arbitrary and permit these types to determine the probability of victory in a very general way.

Specifically, consider the following general information structure for our ultimatum game. This information structure includes both discrete types and continuum types. Formally, suppose country i's set of possible types is a compact set $T_i \subseteq \mathbb{R}$. Thus, T_i could be a collection of discrete points or a closed interval, for example. In order to ensure we have two-sided incomplete information, we assume that each T_i contains at least two elements. Country i's type t_i is independently drawn from a cumulative distribution function F with support T_i.[6] Note that we make no additional assumptions about the distributions F. As in the rest of the paper, the types of the two countries determine the probability that country 1 prevails in a war, which we denote $p(t_1, t_2)$. We assume that $p(t_1, t_2)$ is strictly increasing in t_1 and strictly decreasing in t_2, so that the type of a country is a measure of its strength in the case of war. For ease of notation, we sometimes write $p_1(t_1, t_2) = p(t_1, t_2)$ and $p_2(t_1, t_2) = 1 - p(t_1, t_2)$. Finally, we assume that the technology of war treats the two countries symmetrically so that $p_1(t_1, t_2) = p_2(t_2, t_1)$.

We define mutual optimism as discussed above. That is, mutual optimism occurs at a type pair (t_1, t_2) if the expected value for war for type t_1 of country 1 and the expected value for war for type t_2 of country 2 sum to more than one before considering bargaining strategies. Formally, we say mutual optimism holds at the type pair (t_1, t_2) if

$$\int_T (p(t_1, x) - c)\, dF(x) + \int_T (1 - p(y, t_2) - c)\, dF(y) > 1,$$

which, by symmetry, can be written as

$$\int_T p(t_1, x)\, dF(x) + \int_T p(t_2, y)\, dF(y) > 1 + 2c.$$

Again we will focus on perfect Bayesian equilibrium satistfing the D1 requirement for beliefs off the path of play. Recall this refinement requires that after observing an off-the-equilibrium-path offer, country 2 believes this offer comes from the type of country 1 that has the most to gain from deviating from the equilibrium to this offer. In order to formally describe the D1 refinement, we begin by providing notation for a perfect Bayesian equilibrium in this game. We let $x^*(t_1)$ be the equilibrium offer made by type t_1 of country 1.[7] Let μ_x be the equilibrium belief of country 2 about t_1 after receiving an offer x. Formally, then μ_x is a probability measure on T for every x. Then type t_2 of country 2 will accept an offer x if

$$1 - x \geq 1 - E_{\mu_x} p(t_1, t_2) - c$$
$$x \leq E_{\mu_x} p(t_1, t_2) + c.$$

[6]This implies that $F(t) > 0$ for all $t > \min T_i$ and $F(t) < 1$ for all $t < \max T_i$.

[7]It is possible to show that a perfect Bayesian equilibrium satisfying the D1 refinement cannot involve any type mixing.

For a fixed value of x, the left-hand side of this inequality is fixed and the right-hand side is strictly decreasing in t_2. Therefore the best response for country 2 can be characterized by a unique cutpoint $t_2^*(x)$ such that all types $t_2 < t_2^*(x)$ accept an offer $(x, 1-x)$ and all types $t_2 > t_2^*(x)$ reject an offer $(x, 1-x)$.[8] Note that $t_2^*(x) = \underline{t}$ implies that an offer $(x, 1-x)$ will be rejected with probability one and likewise $t_2^*(x) = \bar{t}$ implies that an offer $(x, 1-x)$ will be accepted with probability one. Finally, note that if $\underline{t} < t_2^*(x) < \bar{t}$, then $t_2^*(x)$ is the unique value of t such that

$$x = E_{\mu_x} p(t_1, t_2^*(x)) + c$$

For a given offer x, some types of country 2 may have a dominant strategy to accept this offer. Specifically, let $t_2^A(x)$ be the largest value of $t_2 \in T$ that satisfies $x \leq p(\underline{t}, t_2) + c$. If no such value of t_2 exists, set $t_2^A(x) = \underline{t}$. Likewise, some types of country 2 may have a dominant strategy to reject an offer x. Let $t_2^R(x)$ be the smallest value of $t_2 \in T$ that satisfies $x \geq p(\bar{t}, t_2) + c$. If no such value of t_2 exists, set $t_2^R(x) = \bar{t}$.

Finally, fix a perfect Bayesian equilibrium and let $U_1^*(t_1)$ be the equilibrium expected utility of type t_1 of country 1. For a given offer $(x, 1-x)$, denote an arbitrary mixed strategy profile for country 2 by $r_x : T \to [0, 1]$, where $r_x(t_2)$ is the probability that type t_2 of country 2 rejects the offer $(x, 1-x)$. Given such a mixed strategy, we can write the expected utility of type t_1 making the offer x as

$$U_1(x, t_1 \mid r_x) = \int_T r_x(t_2)(p(t_1, t_2) - c) + (1 - r_x(t_2))x \, dF(t_2).$$

We say a mixed strategy profile is undominated if $r_x(t_2) = 0$ for all $t_2 < t_2^A(x)$ and $r_x(t_2) = 1$ for all $t_2 > t_2^R(x)$. We can now state the D1 refinement as it applies to our setting. If the equilibrium satisfies the D1 refinement, then for every off the equilibrium path offer x, every undominated mixed strategy r_x, and every pair of types t_1 and t_1', if $U_1(x, t_1 \mid r_x) \geq U_1^*(t_1)$ implies $U_1(x, t_1' \mid r_x) > U_1^*(t_1')$, then t_1 is not in the support of μ_x.

We can now present our main result. It states that under the very general conditions given above, it is impossible for mutual optimism to be necessary and sufficient for war. Because of the general nature of our setting, we cannot say whether it will fail to be necessary or fail to be sufficient (or both), but we can establish that at least one of necessity and sufficiency must fail.

[8]Here we do not specify the action of type $t_2 = t_2^*(x)$. This action can be specified arbitrarily without affecting the equilibrium analysis.

Proposition 3 *In all perfect Bayesian equilibria of this model that satisfy D1, mutual optimism is either not necessary or not sufficient for war.*

Before offering a proof of the Proposition, we give a lemma that is proven in the appendix.

Lemma 1 *In all perfect Bayesian equilibria that satisfy D1, all types of country 1 (except possibly the weakest type $t_1 = \underline{t}$) make offers that are both accepted and rejected with positive probability.*

We now use this lemma to prove the Proposition.

Proof: We must show that mutual optimism is either not necessary or not sufficient for war in all equilibria satisfying D1. Given the definition of mutual optimism and the symmetry of the model, it is straightforward to show that mutual optimism is symmetric. That is, if mutual optimism holds at a type pair (a, b), then mutual optimism also holds at the type pair (b, a). So fix an arbitrary equilibrium satisfying D1 and consider the type $t_1 = \bar{t}$ for country 1. By the above result, this type is making an offer that is accepted with positive probability. Therefore, there exists a type $\tilde{t} > \underline{t}$ that accepts this offer. In other words, there is no war at the type pair $(\bar{t}, \tilde{t})$. There are two possibilities for mutual optimism at this type pair. If there is mutual optimism at $(\bar{t}, \tilde{t})$, then mutual optimism is not sufficient for war. On the other hand, if there is not mutual optimism at $(\bar{t}, \tilde{t})$, then there is not mutual optimism at $(\tilde{t}, \bar{t})$. But by the above result, the offer made by $\tilde{t}$ must be rejected by type $\bar{t}$ of country 2 and so there is war at this type pair. Therefore, mutual optimism is not necessary for war. We conclude then that mutual optimism is either not necessary or not sufficient for war. ∎

This proposition shows that, once again, mutual optimism is either not necessary or not sufficient for war. Thus the conclusions about mutual optimism that we have drawn in the earlier sections carry over to the conceptualization of mutual optimism as the lack of a bargaining range.

5 War and Uncertainty about Power

Our results up to this point show that mutual optimism is not a satisfactory explanation of war in bargaining models. But if mutual optimism does not explain the occurrence of war, what does? In this section we argue that the mere existence of two-sided uncertainty acts to prevent completely peaceful outcomes, except in the most extreme cases. Even with substantial mutual pessimism there is a chance of war.

As elsewhere in the paper, we suppose the two countries engage in an ultimatum game. Thus, country 1 makes an offer $(x, 1-x)$, where $x \in [0,1]$, which country 2 chooses to accept or reject. If country 2 accepts, the payoffs are $(x, 1-x)$. If country 2 rejects, both sides receive their war payoff, which is $p(t_1, t_2) - c_1$ for country 1 and $1 - p(t_1, t_2) - c_2$ for country 2, where $c_i > 0$ for $i = 1, 2$. We assume that $p(t_1, t_2)$ is strictly increasing in t_1 and strictly decreasing in t_2, so that the type of a country is a measure of its strength in the case of war. We make no other assumptions about this function. It will be useful to let $W_i(t_i)$ be the expected value of war for country i given the prior probability F_j, so that

$$W_i(t_i) = \int_{T_j} p_i(t_i, t_j) - c_i \, dF_j(t_j) = \int_{T_j} p_i(t_i, t_j) \, dF_j(t_j) - c_i.$$

Given the monotonicity assumptions about $p_i(t_i, t_j)$, it is clear that $W_i(t_i)$ is a strictly increasing function.

In what follows, we will need to identify the the smallest value of $W_i(t_i)$ that occurs with probability one, which we call $\bar{W}_i$. As F_i has full support, this is equal to $W_i(\bar{t}_i)$ if $\bar{t}_i$ is a mass point of F_i and $\lim_{t_i \to \bar{t}_i} W_i(t_i)$ if it is not.[9]

We can now describe the possible outcomes of equilibria to this bargaining game. We do so in the next theorem. As before, we restrict attention to perfect Bayesian equilibria that satisfy the D1 refinement.

Theorem 1 *Consider a ultimatum game with two-sided incomplete information as defined above. In every PBE of such a game that satisfies D1, either there is a positive probability that war occurs, or country 1 makes the offer $x = 1$, which is accepted with probability one.*

Proof: We will show that if there is a PBE that satisfies D1 in which the probability of war is zero, then it must be that all types of country 1 makes the offer $x = 1$. So fix a PBE satisfying D1 with a zero probability of war. We first show that all types of country 1 must be make the same offer in such an equilibrium. To see this, suppose that two distinct offers are made on the equilibrium path, either because different types are making different offers or some type is mixing in equilibrium. As the probability of war is zero, each offer is

[9] More formally, if $\bar{t}_i$ is an isolated point of F_i, then it is a mass point because F_i has full support. If not, then there is a sequence of elements of T_i converging to $\bar{t}_i$, and the limit of $W_i(t_i)$ exists because $W_i(t_i)$ is monotonic.

accepted with probability one. But then the larger offer gives a strictly higher payoff, so it cannot be optimal for some type to make the lower offer with positive probability. Thus, all types of country 1 must be making the same offer, which we denote x^*.

To proceed, note that as T_i is compact and contains at least two elements, it follows that $\underline{t}_i = \min T_i < \max T_i = \bar{t}_i$. As all types of country 1 are making the offer x^*, country 2's belief after seeing this pooling offer must equal its prior F_1. By assumption, the offer x^* is accepted with probability one, so this requires that $1 - x^* \geq W_2(t_2)$ holds with probability one. This is equivalent to $1 - x^* \geq \bar{W}_2$.

Next, we examine what the D1 refinement requires about an off-the-path offer $x' \in (x^*, 1]$. Because the equilibrium offer x^* is accepted with probability one, every type of country 1 has an equilibrium payoff $U^*(t_1) = x^*$. On the other hand, if the offer x' is rejected with probability $r \in [0, 1]$, then the utility to type t_1 of country 1 is $(1 - r)x' + rW_1(t_1)$. But then it is easy to see that the type of country 1 who potentially has the most to gain from this deviation (in the sense of the largest set of values of r that makes the deviation profitable) is $t_1 = \bar{t}_1$, because this type maximizes $W_1(t_1)$. So the D1 refinement requires that for every off-the-path offer $x' \in (x^*, 1]$, country 2 puts probability one on this offer coming from type $t_1 = \bar{t}_1$.

Finally, we will show that $x^* = 1$. To show this, suppose $x^* < 1$. By the above, after the off-the-path offer $x' \in (x^*, 1]$, type $\bar{t}_2$ has a war payoff of $1 - p(\bar{t}_1, \bar{t}_2) - c_2$. As F_1 has full support, this is strictly less than $\bar{W}_2$. This means that

$$1 - x^* \geq \bar{W}_2 > 1 - p(\bar{t}_1, \bar{t}_2) - c_2,$$

which implies that we can choose $x' > x^*$ such that $1 - x' > 1 - p(\bar{t}_1, \bar{t}_2) - c_2$. But this implies that accepting the offer x' is strictly better than fighting for every type of country 2. So this offer will be accepted with probability one, which means the offer x' is a profitable deviation for every type of country 1. This means there cannot be an equilibrium satisfying D1 with zero probability of war and $x^* < 1$. This completes the proof. ∎

This theorem shows that peaceful outcomes are only possible in the extreme case in which war is so disastrous for country 2 that even the strongest type prefers conceding the entirety of the disputed good rather than fight.

In short, this theorem says that there is always a positive probability of war except in the case in which country 2 is willing to give the entire pie to country 1 without a fight. We can think of this latter case as capitulation by country 2, in which case this result can be restated as saying that in ultimatum bargaining with two-sided incomplete information, there is some chance that war occurs unless country 2 capitulates—it is impossible to have a completely peaceful resolution of the dispute where both sides compromise.

This result stands in contrast to models with one-sided incomplete information. In this setting, for every value $x \in (0, 1)$ there are parameters in which the unique equilibrium involves both sides compromising on the division $(x, 1-x)$, with no chance of war. To see this, suppose that the war payoffs for country 1 and country 2 are $p - c_1$ and $1 - p - c_2$, respectively. Country 2 knows the true value of p and country 1's belief about the value of p is given by a continuous probability density $f(p)$ with an strictly increasing hazard rate $f(p)/(1 - F(p))$. With these assumptions it is not difficult to show that there is a unique equilibrium and, if $f(0) \geq 1/(c_1 + c_2)$, country 1 offers $x^* = c_2$ which is accepted with probability one.[10] This shows that with one-sided incomplete information, the full range of peaceful settlements are possible, while we have just shown that only the extreme case of $x = 1$ is possible as a peaceful settlement with two-sided incomplete information.

To further explore the cause of war in this setting, consider the following corollary.

Corollary 1 *Consider a ultimatum game with two-sided incomplete information as defined above. In every PBE of such a game that satisfies D1, a positive probability of war occurs if and only if* $\bar{W}_2 > 0$.

Proof: First, suppose $\bar{W}_2 > 0$. This implies that with positive probability country 2 will prefer to fight rather than accept an offer of $x = 1$. So it follows from Theorem 1 that every PBE satisfying D1 has a positive probability of war.

To show the converse, suppose $\bar{W}_2 < 0$. Then every offer $x \in [0, 1]$ will be accepted with probability one, so the unique PBE is for country 1 to offer $x = 1$ which is accepted by all types. If $\bar{W}_2 = 0$, then it could be that with positive probability, country 2 is a type that is indifferent between accepting $x = 1$ and fighting. But there cannot be a PBE in which war occurs with positive probability, because of the standard argument that country 1 could offer $1 - \varepsilon$ instead, which will be accepted for sure, and this is a profitable deviation for sufficiently small ε. This proves the result. ∎

What does this mean for mutual optimism? Consider a situation in which country 1 is completely pessimistic about the outcome of a war and country 2 is almost completely pessimistic. Formally, suppose that all the types of country 1 have $W_1(t_1) < 0$ and thus

[10] Details can be found in Fey (2015).

all types would prefer to give up the entire resource than fight. In particular, note that $W_1(t_1)$ can be arbitrarily negative for all t_1. For country 2, suppose an arbitrarily small set of types of country 2 have an ex ante payoff to war that is positive. That is, $W_2(\hat{t}_2) > 0$ for $\hat{t}_2 \in [\bar{t}_2 - \varepsilon, \bar{t}_2]$ for some $\varepsilon > 0$. Then by Corollary 1 there is a positive probability of war, but country 1 is never optimistic relative to their war payoff and all but this small group of country 2 types would rather give up the entire pie than go to war, yet war occurs! Indeed, this can occur when country 1's cost of war is arbitrarily large!

Alternatively, if we want to define mutual optimism as the situation where the two countries think they are likely to win, we can see that this is also not necessary. Suppose now that the probability of country 1 winning satisfies $p(t_1, t_2) < \varepsilon$ for some arbitrarily small number $\varepsilon > 0$ and for all t_1 and t_2. Clearly, in this situation country 1 never thinks they are likely to win, but it is easy to find a range of costs for country 2 such that the expected war payoff of the strongest type is positive, especially since every type thinks it will win with a probability very close to 1. By Corollary 1, then, there are some types of country 1 who make offers that lead to war even though it is not the case that both sides think they will prevail in war.

Why does this happen? The fundamental elements in a situation with two-sided incomplete information about the probability of victory is that the private information of each decision-maker influences the relative values of war and peace for both players and, as a result, "stronger" proposers have an incentive to signal their strength through making higher demands. Furthermore, there is an incentive for slightly weaker types to mimic these stronger types by also making large demands. Thus, in the bargaining model, types that are not particularly optimistic about either their chances of winning or their final war payoff may end up fighting because the demands are too large even for adversaries that are not particularly optimistic themselves.

This is what we mean when we argue mutual optimism and war are not tightly connected in rationalist model of war. Sometimes war occurs with mutual optimism and sometimes war occurs without it. Sometimes there is mutual optimism yet no war happens at all. The one basic fact that holds in all crisis bargaining games is that stronger types are more likely to fight (Fey & Ramsay 2007). This is a reflection of the basic risk/reward tradeoff inherent in crisis bargaining. As we have said, it is important to remember, though, that this fact is a *relative* claim, while the mutual optimism argument is an *absolute* claim. Thus, as in our examples above, a very pessimistic type can choose to fight because it is slightly stronger than other types, while the mutual optimism argument would say that such a type would not choose to fight.

6 Conclusion

Uniting our results are some common themes that are worth highlighting. First, as emphasized by Blainey (1988) and Wittman (1979), our models focus on uncertainty about balance of power. More specifically, we analyze conflict games in which the two sides have privation information about the likelihood of prevailing in war. There are two important aspects of this focus. First, we focus on two-sided incomplete information. In our view, this is necessary in order to study *mutual* optimism. If only one side is uncertain, there can only be unilateral optimism. Second, because both sides have private information about something they both care about, namely the outcome of war, we have interdependence in the two sides' values of choosing war. This is distinct from models with uncertainty about "privately valued" elements of utility, such as costs, and makes our models more complicated but also strategically more interesting. Indeed, if uncertainty is solely about costs, as is common in much of the literature on crisis bargaining, then there will *always* be some peaceful agreement that *both* sides prefer to war. This suggests that mutual optimism as commonly understood cannot arise in such models.

References

Banks, Jeffrey S. 1990. "Equilibrium behavior in crisis bargaining games." *American Journal of Political Science* pp. 599–614.

Blainey, Geoffrey. 1988. *The Causes of War*. New York, NY: The Free Press.

Cho, In-Koo & David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *The Quarterly Journal of Economics* 102(2):179–221.

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

Fey, Mark. 2015. "Learning to Fight: Information, Uncertainty, and the Risk of War." working paper.

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

Fey, Mark & Kristopher W. Ramsay. 2011. "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict." *American Journal of Political Science* 55(1):149–169.

Kim, Woosang & Bruce Bueno de Mesquita. 1995. "How Perceptions Influence the Risk of War." *International Studies Quarterly* 39(1):51–65.

Leventoglu, Bahar & Ahmer Tarar. 2008. "Does Private Information Lead to Delay or War in Crisis Bargaining?" *International Studies Quarterly* 52(3):533–553.

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

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

Wagner, R. Harrison. 1994. "Peace, War and the Balance of Power." *American Political Science Review* 88(3):593–607.

Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44(3):469–484.

Werner, Suzanne & Amy Yuen. 2005. "Making and Keeping Peace." *International Organization* 58(2):261–292.

Wittman, Donald. 1979. "How Wars End." *Journal of Conflict Resolution* 23(4):743–763.

Wittman, Donald. 2009. "Bargaining in the Shadow of War: When Is a Peaceful Resolution Most Likely?" *American Journal of Political Science* 55(5):588–602.

A Equilibrium Details for Example

In this appendix, we provide the details of the equilibrium for the example given in Section 4. Recall that in the example,

$$p(t_1, t_2) = \frac{1}{2} + \frac{t_1 - t_2}{4}$$

and that each country's type is independently drawn from a uniform distribution on $[0, 1]$. Recall also that both countries' cost of war is $1/10$.

It is then easy to calculate the war payoff for side i, conditional on t_i, to be

$$\begin{aligned} W_i(t_i) &= \int_0^1 (\frac{1}{2} + \frac{t_i - t_j}{4} - \frac{1}{10})\, dt_j \\ &= \frac{11}{40} + \frac{t_i}{4}. \end{aligned}$$

In addition, recall that we then have mutual optimism whenever

$$\begin{aligned} W_1(t_1) + W_2(t_2) &= \frac{11}{20} + \frac{t_1 + t_2}{4} > 1 \\ t_1 &> \frac{9}{5} - t_2. \end{aligned}$$

We will find a perfect Bayesian equilibrium satisfying the D1 refinement. It is well known that such an equilibrium must be fully separating, so that every type of country 1 makes a distinct offer. We denote the offer made by type t_1 as $x(t_1)$. Because country 2 knows its own type and can perfectly infer the type of country 1 from its offer, type t_2 of country 2 will accept the equilibrium offer $x(t_1)$ if country 1's offer satisfies

$$1 - x(t_1) \geq 1 - (\frac{1}{2} + \frac{t_1 - t_2}{4}) - \frac{1}{10}$$

and the type $\hat{t}_2$ that is indifferent is given by

$$\hat{t}_2 = 2 + \frac{4}{10} + t_1 - 4x(t_1).$$

All higher types of country 2 will reject and all lower types will accept.

To characterize the equilibrium, we use the incentive compatibility condition given by

Fey & Ramsay (2011). In this example the incentive compatibility condition is given by

$$\frac{dU_1(t_1)}{dt_1} = \frac{1}{4}(1 - \hat{t}_2).$$

By direct calculation, we have

$$U_1(t_1) = \hat{t}_2 x(t_1) + \int_{\hat{t}_2}^{1} p(t_1, t_2) - \frac{1}{10} dt_2$$

and after substitution and differentiating, we have

$$\frac{dU_1(t_1)}{dt_1} = x(t_1)(1 - 4x'(t_1)) + x'(t_1)(2 - \frac{4}{10} + t_1) - (\frac{1}{4} + \frac{t_1}{4} - \frac{1}{10}).$$

Setting these two derivatives equal gives us a differential equation. The correct boundary condition to use is to fix the offer of the lowest type to be their ultimatum offer in the one-sided incomplete information game ($x^*(0) = .4$). This insures that the lowest type's participation constraint is satisfied and gives us the following solution:

$$x^*(t_1) = \frac{3}{5} + \frac{t_1}{4} + \frac{1}{5}\text{ProductLog}(-.367879e^{-1.25t_1}).$$

With this expression for the equilibrium offer of country 1, we can calculate the values of $\hat{t}_2$, which in turn allows us to generate the regions given in the figure in the text.

B Proof of Lemma

Lemma 1 *In all perfect Bayesian equilibria that satisfy D1, all types of country 1 (except possibly the weakest type $t_1 = \underline{t}$) make offers that are both accepted and rejected with positive probability.*

Proof: Fix an equilibrium that satisfies D1. By Proposition 4 in Fey & Ramsay (2011), there exists $t^p \in T$ and $t^w \in T$ with $t^p \leq t^w$ such that all types $t_1 < t^p$ make offers that are accepted with probability 1 and all types $t_1 > t^w$ make offers that are rejected with probability 1. Let $W(t_1)$ be the utility of type t_1 of country 1 if its offer is rejected with probability 1. That is,

$$W(t_1) = \int_T p(t_1, t_2)\, dF(t_2) - c.$$

We first show that in this model, $t^p = \underline{t}$. For a proof by contradiction, suppose that $t^p > \underline{t}$. This implies that there exists some $\tilde{x}$ such that $x^*(t_1) = \tilde{x}$ for all $t_1 \in [\underline{t}, t^p)$ and $x^*(t_1) \neq \tilde{x}$ for all $t_1 > t^p$. Therefore, it follows that

$$\tilde{x} \leq E[p(t_1, \bar{t}) + c \mid t_1 < t^p].$$

If we let $x^p = p(t^p, \bar{t}) + c$, then because $p(t_1, t_2)$ is strictly increasing in t_1, we have that $\tilde{x} < x^p$. Now consider the offer $(x^p, 1 - x^p)$. If this offer is on the equilibrium path, then for all $t_1 < t^p$, t_1 is not in the support of μ_{x^p}. On the other hand, if this offer is off the equilibrium path, then take types $t_1 < t_1' < t_p$ and an undominated mixed strategy r_{x^p} such that $U_1(x^p, t_1 \mid r_{x^p}) \geq U_1^*(t_1) = \tilde{x}$. This inequality simplifies to

$$\int_T r_{x^p}(t_2)(p(t_1, t_2) - c - \tilde{x}) + (1 - r_{x^p}(t_2))(x^p - \tilde{x})\, dF(t_2) \geq 0.$$

Because p is strictly increasing in t_1, it follows that

$$\int_T r_{x^p}(t_2)(p(t_1', t_2) - c - \tilde{x}) + (1 - r_{x^p}(t_2))(x^p - \tilde{x})\, dF(t_2) > 0.$$

Note that this is true even if $r_{x^p} = 0$ almost everywhere because $x^p > \tilde{x}$. But this strict inequality can be written as $U_1(x^p, t_1' \mid r_{x^p}) > U_1^*(t_1')$ and so the D1 refinement requires that t_1 is not in the support of μ_{x^p}. As this argument holds for any $t_1 < t^p$, it must be that $t_1 < t^p$ implies t_1 is not in the support of μ_{x^p}. Using this, we now argue that the offer $(x^p, 1 - x^p)$ will be accepted by all types of country 2. As $p(t_1, t_2)$ is strictly decreasing in t_2, we have $p(t^p, \bar{t}) \leq p(t^p, t_2)$ for all $t_2 \in T$. In addition, as the support of μ_{x^p} does not contain the interval $[\underline{t}, t^p)$, it must be that $p(t^p, t_2) \leq E_{\mu_{x^p}} p(t_1, t_2)$. We thus have, for all $t_2 \in T$,

$$x^p = p(t^p, \bar{t}) + c \leq E_{\mu_{x^p}} p(t_1, t_2) + c,$$

and so all types of country 2 will accept the offer $(x^p, 1 - x^p)$. But this means that for any type $t_1 < t^p$, deviating from $x^*(t_1) = \tilde{x}$ to offering x^p will be a profitable deviation. This contradiction proves our claim that $t^p = \underline{t}$.

Next, we consider types of country 1 that are making offers that are rejected with probability 1. From above, there exists $t^w \in T$ such that all types $t_1 > t^w$ are making such offers. We first show that $t^w = \bar{t}$. For a proof by contradiction, suppose that $t^w < \bar{t}$. We begin by selecting $\hat{t} \in (t^w, \bar{t}]$ such that $p(\hat{t}, \underline{t}) - p(t^w, \underline{t}) < c$. This is possible by the continuity of $p(t_1, t_2)$. In addition, we let $x^0 = p(t^w, \underline{t})$ and so we have $x^0 > p(\hat{t}, \underline{t}) - c$. Type $\hat{t}$ of country 1 is making an equilibrium offer that is rejected with probability one. Therefore, the equilibrium payoff of this type is $U_1^*(\hat{t}) = W(\hat{t})$. We claim that this implies that the offer $(x^0, 1 - x^0)$ must be rejected with probability one in this equilibrium. To see this, observe that the equilibrium payoff $W(\hat{t})$ must be at least as big as the payoff of making the offer x^0, so we have

$$\int_{\underline{t}}^{\bar{t}} (p(\hat{t}, t_2) - c)\, dF(t_2) \geq \int_{\underline{t}}^{t_2^*(x^0)} x^0\, dF(t_2) + \int_{t_2^*(x^0)}^{\bar{t}} (p(\hat{t}, t_2) - c)\, dF(t_2)$$
$$\int_{\underline{t}}^{t_2^*(x^0)} p(\hat{t}, t_2) - c\, dF(t_2) \geq \int_{\underline{t}}^{t_2^*(x^0)} x^0\, dF(t_2).$$

However, we have $x^0 > p(\hat{t}, \underline{t}) - c$ which implies that $x^0 > p(\hat{t}, t_2) - c$ for all $t_2 \in T$. From this it is clear that the only way for the equilibrium condition to hold is for $t_2^*(x^0) = \underline{t}$, which means that the offer x^0 is rejected with probability one.

We now use the fact that the offer x^0 is rejected with probability one to derive a contradiction. There are two cases, as above. First, suppose that the offer x^0 is on the equilibrium path. Then because it is rejected with probability one, it cannot be played by any type $t_1 < t^w$. Therefore, the support of μ_{x^0} must be contained in the interval $[t^w, \bar{t}]$. From this it follows that $E_{\mu_{x^0}} p(t_1, \underline{t}) \geq p(t^w, \underline{t}) = x^0$. But this implies that $E_{\mu_{x^0}} p(t_1, \underline{t}) + c > x^0$ and therefore $t_2^*(x^0) > \underline{t}$, which contradicts our result that the offer x^0 is rejected with probability one. The second case is that x^0 is off the equilibrium path. Before we give the D1 refinement for this case, we first show that for all $t_1 < t^w$, $U_1^*(t_1) > W(t_1)$. Because x^0 is rejected with probability one and this cannot be a profitable deviation for type t_1, we know that $U_1^*(t_1) \geq W(t_1)$ for all $t_1 < t^w$. So suppose that there exists a type t_1 such that $U_1^*(t_1) = W(t_1)$ and consider a type $t_1' \in (t_1, t^w)$. If we let $x' = x^*(t_1')$, then from the fact that $U_1^*(t_1') \geq W(t_1')$ we have

$$\int_{\underline{t}}^{t_2^*(x')} x'\, dF(t_2) + \int_{t_2^*(x')}^{\bar{t}} (p(t_1', t_2) - c)\, dF(t_2) \geq \int_T (p(t_1', t_2) - c)\, dF(t_2)$$
$$\int_{\underline{t}}^{t_2^*(x')} x'\, dF(t_2) \geq \int_{\underline{t}}^{t_2^*(x')} (p(t_1', t_2) - c)\, dF(t_2).$$

But now consider the payoff if type t_1 deviates to the offer x'. This is given by $\int_{\underline{t}}^{t_2^*(x')} x'\, dF(t_2)+$

$\int_{t_2^*(x')}^{\bar{t}} (p(t_1, t_2) - c)\, dF(t_2)$. From the above, we have

$$\begin{aligned}
&\int_{\underline{t}}^{t_2^*(x')} x'\, dF(t_2) + \int_{t_2^*(x')}^{\bar{t}} (p(t_1, t_2) - c)\, dF(t_2) \\
&\qquad \geq \int_{\underline{t}}^{t_2^*(x')} (p(t_1', t_2) - c)\, dF(t_2) + \int_{t_2^*(x')}^{\bar{t}} (p(t_1, t_2) - c)\, dF(t_2) \\
&\qquad > \int_{\underline{t}}^{t_2^*(x')} (p(t_1, t_2) - c)\, dF(t_2) + \int_{t_2^*(x')}^{\bar{t}} (p(t_1, t_2) - c)\, dF(t_2) = W(t_1),
\end{aligned}$$

where the last inequality comes from the fact that $p(t_1, t_2)$ is increasing in t_1 and that $t_2^*(x') > \underline{t}$. But this implies that deviating to x' is a profitable deviation for type t_1, which is a contradiction. This establishes that for all $t_1 < t^w$, $U_1^*(t_1) > W(t_1)$. We now return to the D1 refinement. Pick an arbitrary type $t_1 < t^w$ and an undominated mixed strategy r_{x^0} such that $U_1(x^0, t_1 \mid r_{x^0}) \geq U_1^*(t_1)$. Because $U_1^*(t_1) > W(t_1)$, the mixed strategy r_{x^0} must involve the offer $(x^0, 1 - x^0)$ being accepted with positive probability. But for the type $\hat{t}$, we know from the earlier argument that $x^0 > p(\hat{t}, t_2) - c$ for all $t_2 \in T$ and so any such r_{x^0} results in a strictly higher payoff than $W(\hat{t})$. In other words, $U_1(x^0, \hat{t} \mid r_{x^0}) > U_1^*(\hat{t})$ for all such r_{x^0} and so the D1 refinement requires that t_1 is not in the support of μ_{x^0}. But this holds for all $t_1 < t^w$ and so the support of μ_{x^0} must be contained in the interval $[t^w, \bar{t}]$. As in the first case, this means that $t_2^*(x^0) > \underline{t}$, which contradicts our result that the offer x^0 is rejected with probability one.

We thus have shown that it is not the case that $t^w < \bar{t}$. This means that no type $t_1 < \bar{t}$ is making an equilibrium offer that is rejected with probability one. But what about the type $t_1 = \bar{t}$? In fact, the argument that we have just given applies to the case in which this type is making an offer that is rejected with probability one. In sum, then, we have established that *all* types of country 1 make offers in equilibrium that are accepted with positive probability and all types of country 1, except possibly for $t_1 = \underline{t}$, make offers in equilibrium that are rejected with positive probability. ∎