Mutual Optimism and Choosing to Fight

Mark Fey*

Kristopher W. Ramsay†

May 8, 2011

Abstract

In this paper we argue that mutual optimism is not a valid explanation for war. Specifically, we ask if it is the case that war occurs if and only if mutual optimism occurs and if it is the case that the presence of mutual optimism is necessary for a positive probability of war. We show that in models in which either side can choose to fight, the answer to both of these questions is no. Thus, in a wide range of models, including the standard bargaining model of war, mutual optimism is neither always necessary or sufficient for war. War can occur without mutual optimism and mutual optimism can occur without war.

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

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

1 Introduction

The question of why states fight costly wars when less costly negotiated settlements are available is central to the politics between states. One potential explanation that has received much attention in the literature 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 is 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" (Leventoğlu & Tarar 2008, 533).

In considering what causes war, one difficulty that arises is that there are, in fact, many possible causes of war. Fearon (1995), for example, describes three rationalist explanations for war. In order to concentrate on a specific cause of war, then, we must somehow distinguish it from other possible causes. Given this, how might the mutual optimism explanation be validated? One answer is given by Fey & Ramsay (2007). Their approach is to create a set of environments in which mutual optimism should produce war and, conversely, a lack of mutual optimism should produce peace. If war occurs in these environments, the mutual optimism explanation for war gains credence. If war does not occur in these environments, then the mutual optimism explanation is suspect. Fey and Ramsay then show that in every instance of these environments, war does not occur. They thus conclude that mutual optimism is not a valid rationalist explanation for war.

Fey & Ramsay (2007) construct these environments by imposing two main assumptions. The first is that war can only occur when both sides agree to fight and the second is that the settlement procedure used if war does not occur is always available to either side, regardless of the actions taken previously. How do these assumptions work to insure that mutual optimism should accompany war? The answer is twofold. First, if mutual optimism is present, then both sides should be willing to agree to fight and forgo the peaceful settlement that is available to each side. In other words, even with these two assumptions mutual optimism should be sufficient for war. Second, if mutual optimism is absent, then at least one side will prefer the peaceful settlement and thus refuse to agree to fight. In this way, these two assumptions insure that mutual optimism should also be necessary for war. Thus, the environments considered by Fey & Ramsay (2007) (those that satisfy these two assumptions) are not meant to be realistic models of conflict. Rather, they are meant to be a kind of "thought experiment" which are constructed to test the logical validity of the mutual optimism explanation of war; these environments are ones in which mutual optimism *should* be necessary and sufficient for war. Thus, the main finding of Fey & Ramsay (2007) that in these environments mutual optimism does not lead to war casts doubt on mutual optimism as a rationalist explanation of war.

In this paper, we consider another way of examining the validity of mutual optimism as an explanation of war. We move away from the "thought experiment" environments considered in Fey & Ramsay (2007) and instead focus on environments that could be considered more realistic descriptions of international conflict. Instead of assuming that either side can impose peace, in this paper we consider crisis behavior under a "unilateral war" assumption that permits either side to start a war if it is dissatisfied with the peaceful alternative. The first part of the paper examines models of unilateral war in which the peaceful alternative is some fixed alternative, such as the status quo. The remainder of the paper considers bargaining models in which both decision to go to war and the terms of the peaceful settlement are determined endogenously by the choices of the two sides. In this way, the class of models considered in this paper is perhaps more substantively compelling than the models examined in Fey & Ramsay (2007).

Focusing on this class of models also allows us to address the criticisms of the assumptions and arguments of Fey & Ramsay (2007) made in a recent article by Slantchev & Tarar (2011). The latter article examines a model that is outside of the class considered by Fey & Ramsay (2007) and in which only one side has private information and conclude that, in this model, mutual optimism is both necessary and sufficient for war to occur. 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). While there are several aspects of the argument in Slantchev & Tarar (2011) that we find problematic and discuss in more detail in section 5, the fundamental deficiency with this claim is the misleading definition of mutual optimism given in the article. In their usage of the term "mutual optimism," the authors apply different definitions to the two actors in the model. Specifically, while one state (state D) is said to be optimistic if it is likely to prevail in war, the other state's optimism is not defined relative to its war prospects. Instead, state S is optimistic about whether or not its offer will be accepted. Part of the problem with this definition of mutual optimism is that in the model described by Slantchev & Tarar (2011), only one side has private information. Our view is that mutual optimism requires both sides to have private information, otherwise only unilateral optimism is possible.

Thus, while we agree with Slantchev & Tarar (2011) that we must establish whether mutual optimism is necessary and/or sufficient for war in order to evaluate the mutual optimism explanation, it is clear we must first delineate a valid definition of mutual optimism in models of conflict with two-sided incomplete information. In order to construct this definition, we first specify what it means for a single country to be optimistic about its prospects in war. When there is a fixed peaceful alternative, our definition of optimism for a given side is straightforward: a country is optimistic if, based on its private information, it views war as preferable to peaceful agreement. Given this, it is natural then to define mutual optimism as occurring if both countries prefer fighting to peaceful agreement, based on their individual private information about fighting. These definitions seem to us to capture the idea of mutual optimism described informally by Blainey (1988) and Wagner (1994), in addition to having been been used in earlier work on mutual optimism (Fey & Ramsay 2007). When the terms of peaceful agreement are themselves the product of bargaining, however, it is not clear how to make sense of these definitions. Take unilateral optimism, for example. It is quite likely that we can find *some* peaceful agreement that a given side prefers to war, so should we conclude optimism is all but impossible in a bargaining model? While we find it difficult to provide a reasonable definition of unilateral optimism in this case, we draw on the existing literature to inform a definition of mutual optimism in crisis bargaining. Specifically, we draw on Wagner (1994) and Wittman (1979), who describe mutual optimism as occurring occurring when there is no negotiated settlement that both sides prefer to war. In particular, in our bargaining models, we define mutual optimism as occurring when the expected values of fighting for the two sides sum to more than the total value of the prize under dispute, as this is equivalent to saying no peaceful division exists that both sides would accept. Put another way, we suppose that mutual optimism exists in a bargaining model if, given the private information of the two sides, the bargaining range is empty.[1] This definition of mutual optimism for bargaining games is logically stronger than the definition for fixed peaceful settlements. This reflects the fact that allowing settlements to be determined endogenously is a weaker assumption than requiring them to be fixed exogenously.

We introduce the main ideas of the paper in Section 2 by way of two examples. Both

[1]This is explicitly the model in Werner & Yuen (2005).

examples involve a simple crisis game between two countries in which both sides in turn have the option of fighting to change some fixed status quo. Both sides are uncertain as to the strength of the other side—mutual optimism occurs if they both believe they can do better by war than through settlement. In our first example, we show that there can be states of the world in which mutual optimism is present but war does not occur. That is, mutual optimism need not be sufficient for war. Also in this example, we show that war occurs in some states of the world in which mutual optimism is absent. This shows that mutual optimism need not be necessary for war. Immediately, then, this simple example shows how mutual optimism can fail to satisfy two basic requirements of causality. We build on this in the second example and show that, surprisingly, war can occur even when both sides are pessimistic. That is, we show that even if neither side is unilaterally optimistic, an equilibrium exists in which war occurs. This further calls into question the link between mutual optimism and war.

While these examples are quite suggestive, it is best that general claims be based on broad general models. To this end, in Sections 3 and 4 we describe a general class of models with unilateral war—either side can reject a peaceful settlement and choose war instead. Our first general result answers the question: what leads to war in these environments? We show that the existence of a type of one country that is unilaterally optimistic about the outcome of a war implies that the overall equilibrium probability of war is positive. So while it is not the case that countries with optimistic expectations regarding war always choose to fight, nor is it the case that countries with pessimistic expectations never fight, a single side with optimistic expectations is sufficient to guarantee a non-zero risk of war. We then turn to considering some general results relating mutual optimism to war when any country can unilaterally choose to fight. Our second result is that in every such model, if war occurs in equilibrium, there must be states of the world in which war occurs but mutual optimism is not present. In other words, mutual optimism is never necessary for war. These two results illustrate that we can draw a broad conclusion in a general class of models that mutual optimism is not a satisfactory explanation of war between rational opponents. But what if players aren't fully rational? We address this question by showing that in a setting where players are limited in their information processing, our general result that mutual optimism is never necessary for war continues to hold.

In Section 5, we drop the assumption that the terms of the peaceful settlement are fixed and instead allow the two sides to bargain over a peaceful settlement. This allows us to further extend the scope of our results. Specifically, we consider the "standard model" of take-it-or-leave-it crisis bargaining that is also the focus of Slantchev & Tarar's (2011) discussion of mutual optimism and war. We begin by addressing the specific model presented in Slantchev & Tarar (2011), in which only one side has private information. 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 unjustified given their misleading 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 even interpret the Slantchev & Tarar (2011) results as an approximation of a mutual optimism model with "almost one-sided" incomplete information.

Having dealt with the standard bargaining model of war with one-sided incomplete information, we next examine this same model with two-sided incomplete information. We first show that in a symmetric two-type model, mutual optimism is never necessary for war. Specifically, the optimistic type of the side receiving the offer will always choose to fight with positive probability. In the final part of this section we examine a more general model with a continuum of types and show that mutual optimism is either not sufficient or not necessary for war (or both). In any model in this class, it must be that some mutually optimistic type pair will agree on a peaceful settlement, or some type pair without mutual optimism will go to war, or both. We conclude that our findings about mutual optimism in models with a fixed peaceful settlement extend to bargaining models, further calling into question the mutual optimism explanation for war.

Summing up these results, we find that in a very broad class of situations the presence or absence of mutual optimism is an irrelevant condition for determining whether war will or will not occur. In an important way, mutual optimism and war are just coincidental. Unilateral optimism, however, is a more important marker for war. While optimistic states may fight or may not in any realized state of the world, the possibility of unilateral optimism implies that the *ex ante* probability of war is always positive in any pure strategy equilibrium. Moreover, we show that at any state of the world with mutual optimism, even one in which war does not occur, we can change mutual optimism to unilateral optimism by giving one side full information and the result must be war. Here again, war is associated with one-sided optimism rather than mutual optimism.

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.

Second, our study of environments with uncertainty about the probability of winning leads us to focus on the "strategic inferences" that a rational decision-maker should make about the information of an opponent in an equilibrium theory of war. These inferences do not arise in models with privately valued costs because in such models the value of war to one side does not depend on the private information of the other side. In particular, in such models country 1 does not care which types of country 2 choose war; it only cares about how likely it is that war is chosen. With uncertainty about a commonly valued parameter like the probability of victory in war, however, the value of war depends on the private information of both sides. Therefore, country 1 now cares which types of country 2 choose war, which is something that is determined by the equilibrium strategy of country 2. In this way, country 1 must make a strategic inference about the information of country 2 in choosing its optimal decision. These inferences take a particularly powerful form in games with unilateral war, which are the focus of our study. Specifically, in such a strategic context, if your opponent is choosing to go to war, then your choice does not matter—the outcome is war no matter what. Put another way, your choice matters only if your opponent is choosing to not unilaterally start a war. Thus, in deciding on an optimal choice of action, a decision-maker should condition on the fact that their opponent is not fighting. This, in turn, means that in equilibrium a decision-maker acts in equilibrium as if she has different information than she possesses. In this way, our analysis is related to the winner's curse in auction theory (Thaler 1994) and the swing voter's curse in voting theory (Feddersen & Pesendorfer 1996). In international conflict, as we will show, strategic inferences can, among other things, lead pessimistic leaders to start wars and optimistic leaders to accept settlements.

As a final thought, perhaps an analogy would be useful to understand our criticism of the mutual optimism explanation for war in settings where either side can choose to fight. Suppose there are a number of offices, each of which is shared by two graduate students, and either student can turn on the light in a given office. Suppose also that we see there is an office occupied by two students with the light on. It may be tempting to conclude that mutual occupancy is an explanation for the light being on. But what if we also see an office in which the light is on but only one student is present? We must conclude that even though mutual occupancy is associated with the light being on, it is not the correct casual explanation. The cause of the light being on is unilateral occupancy, not mutual occupancy. In the same way, we argue that mutual optimism is not the correct casual explanation for war. One of our main results is that whenever war occurs and mutual optimism is present, there is another state of the world in which war occurs and mutual optimism is not present. Therefore we conclude that there will always be an office with the light on but without mutual occupancy. This casts serious doubt on the mutual optimism explanation of war. Moreover, war is not limited by the occurrence of optimism. It turns out that war can arise even if neither side is optimistic. Surprisingly, the logic of strategic inference leads us to the existence of situations in which the light is on but no one is home!

2 Two Examples

We begin by describing two simple examples that illustrate our main findings. The two examples differ only by what the two sides know about their relative power. The examples share the same game-theoretic structure, with the same actions in the same order. This structure is chosen to be as simple as possible to highlight some important incentives that exist in international conflict. However, as we show in Section 4, our main findings apply to a broad class of models that may more accurately reflect the complexities of crisis bargaining.

Our first example illustrates that mutual optimism is neither a necessary nor a sufficient condition for war. Our second example shows that, in fact, war can occur even in a model lacking any optimism at all.

2.1 Setup

In both of our examples, we have two countries, labeled 1 and 2, which are involved in an international crisis. This crisis can be resolved peacefully or by the use of force. The process by which one of these two outcomes is reached is described by a *game form.* Although our results apply to a wide variety of game forms, as described in Section 3, in order to keep the examples simple we use as simple a game form as possible. In particular, we suppose country 1 begins by choosing to fight or not fight, represented by the choice of actions F or N. If country 1 chooses to fight, then war results. If country 1 chooses to not fight, then country 2 chooses to fight or not fight. If country 2 chooses to fight, then war results, otherwise the peaceful settlement results. Thus, as illustrated in Figure 1, if *either* country chooses action F, then war results. If both countries choose action N, then the peaceful settlement results. This simple game is a game in which either side can choose to fight. One way to conceptualize this game is that there a status quo allocation that can only be changed by war. Either player can start the war and only if both players accept the status quo does peace prevail.[2]

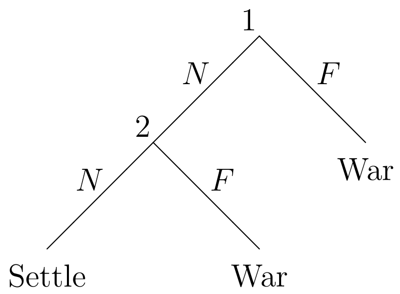

Figure 1: Game Form of Examples

We next describe the payoffs and information structure of our examples. There are two possible outcomes to the game we describe: war and peaceful settlement. For simplicity, we suppose that the peaceful settlement is fixed at a payoff of 1/2 for both countries. This can be viewed as the status quo of equal division of some resource of unit size, for example. In case of war, the outcome is either victory for country 1 and defeat for country 2 or vice versa. We normalize the value of victory to be 1 and the value of defeat to be 0. Regardless of which side wins, war is costly and each side must pay a cost $c_i > 0$ in the event of war.

[2]In fact, the choices of the two sides can be simultaneous or sequential—the following analysis does not depend on the timing of the choices.

Therefore, if we let p_i denote the probability that country i wins the war, the expected utility for war for country i is given by $p_i - c_i$.

In order to analyze mutual optimism and its connection to war, we suppose that each side has private information about its war-fighting ability and that the probability of victory for each side depends on the war-fighting ability of both sides. Formally, both of our examples have two-sided incomplete information, which we now define. We suppose each country can be one of three possible types, A, B, and C, and that these types are equally likely. We denote the type of country i by $t_i \in \{A, B, C\}$ and thus a type profile $t = (t_1, t_2)$ gives the types of both countries. As is standard, we suppose a country knows its own type but is uncertain about the type of its opponent. We suppose the probability of victory for country i depends on the realized type profile t, which we denote $p_i(t_1, t_2)$. As the probability of winning and therefore the value of war depends on the types of both players, both of our examples have "interdependent values."

Now that we have defined the information structure and payoffs for our game, we next define what it means for a country to be optimistic. Informally, we say that a country is optimistic if, based solely on its own private information, it thinks it will be better off fighting a war than receiving the peaceful settlement. Formally, we let $\hat{p}_i(t_i)$ denote the *naive* conditional probability that type t_i of country i will win a war. Thus, we have

$$\hat{p}_i(t_i) = \frac{1}{3} \sum_{t_j \in \{A,B,C\}} p_i(t_i, t_j),$$

the average probability of victory against all of the possible types of the opponent. We say type t_i of country i is optimistic if $\hat{p}_i(t_i) - c_i > 1/2$ and we say there is mutual optimism at type profile $t = (t_1, t_2)$ if both type t_1 of country 1 and type t_2 of country 2 are optimistic. It is important to emphasize that these definitions are naive in that they refer to a country's likelihood of victory without conditioning on which types of their opponent would actually choose to fight.

2.2 Example 1: Information and Equilibrium

The exact way in which the probability of victory varies with the type profile in our first example is given by the table in Figure 2. In this table, the type of country 1 corresponds to the rows and the type of country 2 corresponds to the columns.

The precise values for the probability of winning are chosen to make the presentation of

	A	B	C
A	(.5,.5)	(.3,.7)	(.9,.1)
B	(.7,.3)	(.5,.5)	(.5,.5)
C	(.1,.9)	(.5,.5)	(.5,.5)

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

	A	B	C	
A	(.5,.5)	(.3,.7)	(.9,.1)	Opt.
B	(.7,.3)	(.5,.5)	(.5,.5)	Opt.
C	(.1,.9)	(.5,.5)	(.5,.5)	Pess.
	Opt.	Opt.	Pess.	

Figure 3: Optimism and pessimism ($c_i < 1/15$)

the results in this example as clear as possible. But it is possible to provide some motivation for these values, as follows. Consider an emerging, unproven, new technology for warfare such as poison gas or the tank in World War I or unmanned aerial vehicles (UAVs) in the early 1990s. Now suppose that the two countries in our example have varying abilities in regards to this new technology and these abilities are private information. Specifically, a type A country has the ability to use this new technology in an offensive role, such as stock piles of poison gas and a reliable delivery system. A type B country has an effective defense against this technology, such as the distribution of gas masks, but no ability for offensive use. Finally, a type C country has neither of these abilities. This interpretation of the types of the countries motivates the values in Figure 2 in the following way. If the two countries are the same type, then neither has an advantage: if both are type A they both have an equal offensive advantage and if both are type B or C then the new technology is not used offensively. If one country is type B and the other is type C, then again neither has an advantage because the new technology is not used offensively. On the other hand, if a type A country fights a type C country, then the type A country has an overwhelming advantage and wins with probability .9. Finally, if a type A country faces a type B country in war, the defensive capability of the type B country eliminates the offensive abilities of type A country which makes victory more likely for the type B country. Of course, this description is only meant to make the values in our example plausible, it is not an attempt to explain war generally.

In our example, it is easy to calculate the *naive* conditional probabilities that country i

	A	B	C	
A	(.5,.5)	(.3,.7)	(.9,.1)	Opt.
B	(.7,.3)	(.5,.5)	(.5,.5)	Opt.
C	(.1,.9)	(.5,.5)	(.5,.5)	Pess.
	Opt.	Opt.	Pess.	

Figure 4: Type pairs at war and peace

will win a war:

$$\hat{p}_i(A) = \hat{p}_i(B) = \frac{17}{30} \quad \text{and} \quad \hat{p}_i(C) = \frac{11}{30}.$$

Thus, types A and B of country i are optimistic when $c_i \in (0, 1/15)$ and type C of country i is never optimistic. Intuitively, the requirement that c_i be relatively small to be optimistic should make sense; if war is extremely costly then war will never be a better choice than peace, regardless of a country's private information. We summarize the optimism or pessimism of each type in Figure 3.

We are now ready to identify the equilibria of this game. In fact, when $c_i \in (0, 1/15)$ for $i = 1, 2$, this game has a unique perfect Bayesian equilibrium. In this equilibrium, each country plays action F if its type is A and plays action N if its type is B or C. To see that this strategy profile is indeed an equilibrium, consider country 2. Because of the timing of the game, country 2's choice matters only when country 1 is choosing action N. Given country 1's strategy, this occurs precisely when $t_1 = B$ or $t_1 = C$. Therefore, for type A of country 2, conditional on its action mattering, its expected payoff for war is

$$E[p_2(t_1, A) \mid t_1 \in \{B, C\}] - c_2 = \frac{.3 + .9}{2} - c_2 = .6 - c_2$$

As $c_2 < .1$, choosing F is superior to choosing N. Similarly, for type B or C of country 2, its expected payoff for war conditional on its action mattering is $(.5 + .5)/2 - c_2 = .5 - c_2$. As $c_2 > 0$, these types of country 2 prefer to choose N rather than F. Turning now to the choice of country 1, note that country 1's choice matters only when country 2 is choosing action N. Given country 2's strategy, this occurs precisely when $t_2 = B$ or $t_2 = C$. Therefore, the analysis for country 1 is exactly symmetric to the analysis just described for country 2.[3] Thus the given strategy profile is an equilibrium.[4]

[3] This symmetry in the analysis is why the exact timing of the game in our example is inconsequential.

[4] The proof that this equilibrium is unique is somewhat involved and, therefore, is presented in a Reviewer's Appendix.

2.3 Example 1: Implications

The equilibrium in our first example illustrates several important aspects of the connection between mutual optimism and war. The equilibrium outcomes are given in Figure 4. In this figure, the type pairs for which war occurs are shaded and the type pairs for which the peaceful settlement occurs are unshaded.

There are two main observations to make about the equilibrium in this example. The first observation is that mutual optimism is not necessary for war. If it were necessary, then it would be the case that mutual optimism is present at every type profile for which there is war. But for both type profiles (A, C) and (C, A), war occurs but mutual optimism is not present, because type C of both countries is not optimistic. Thus this example demonstrates that war can occur without mutual optimism, and we can conclude that mutual optimism is not a necessary condition for war.

The second observation is that mutual optimism is not sufficient for war. If it were, then it would be the case that war occurs at every type profile at which mutual optimism is present. But for the type profile (B, B), there is mutual optimism but war does not occur. For both sides, type B has the naive expectation that it will do better in war than in a peaceful settlement, so mutual optimism is present at type profile (B, B). But in the unique perfect Bayesian equilibrium of the game, type B of both countries chooses not to fight. Thus war does not occur at this type profile. We see, then, that mutual optimism can occur without war, and therefore mutual optimism is not a sufficient condition for war.

In sum, mutual optimism is neither necessary nor sufficient for war in this example. In this way, this simple example illustrates our main objection to mutual optimism as a causal story for war. War can occur without mutual optimism and mutual optimism can occur without war.

2.4 Example 2: Information and Equilibrium

In our second example, we maintain the game form and general information framework described above. This example differs, though, in how the probability of victory varies with the type profile of the two sides. These probabilities are given by the table in Figure 5.

It is not difficult to motivate the probabilities in this figure. A type C country has some vulnerability that can be exploited by a type A or B country, while those two types of a country have no significant advantage or disadvantage against each other. Using these probabilities, we can calculate the *naive* conditional probabilities that country i will win a

	A	B	C	
A	(.5,.5)	(.5,.5)	(.65,.35)	Pess.
B	(.5,.5)	(.5,.5)	(.65,.35)	Pess.
C	(.35,.65)	(.35,.65)	(.5,.5)	Pess.
	Pess.	Pess.	Pess.	

Figure 5: Probabilities of winning and pessimism ($c_i > .05$)

war:

$$\hat{p}_i(A) = \hat{p}_i(B) = .55 \quad \text{and} \quad \hat{p}_i(C) = .4.$$

Thus, if $c_i > .05$, all types of country i are pessimistic. This fact is also displayed in Figure 5.

We are now ready to identify the equilibria of the game given in Figure 1 with this information structure. If $c_i > .05$ for $i = 1, 2$, then it is easy to see that this game has a perfect Bayesian equilibrium in which all types of both countries play action N. In this peaceful equilibrium, no type chooses to fight because its expected war payoff is equal to its naive conditional probability of victory minus its cost c_i and for every type this payoff is strictly less than .5. More interestingly, if $.05 < c_i < .15$ for $i = 1, 2$, then there also exists a perfect Bayesian equilibrium in which each country plays action F if its type is A or B and plays action N if its type is C. To see that this strategy profile is indeed an equilibrium, consider country 2. The choice of country 2 matters only if country 1 is choosing N, which only occurs if $t_1 = C$. Therefore, for type A of country 2, conditional on its action mattering, its expected payoff for war is $p_2(C, A) - c_2 = .65 - c_2$. As $c_2 < .15$, choosing F is superior to choosing N. The expected payoff for war for type B of country 2, conditional on its action mattering, is also $.65 - c_2$, so F is also optimal for for type B. Finally for type C of country 2, its expected payoff for war conditional on its action mattering is $p_2(C, C) - c_2 = .5 - c_2$. As $c_2 > 0$, this types of country 2 prefers to choose N rather than F. Turning now to the choice of country 1, note that country 1's choice matters only when country 2 is choosing action N. Given country 2's strategy, this occurs precisely when $t_2 = C$. Therefore, the analysis for country 1 is exactly symmetric to the analysis just described for country 2, which establishes that this strategy profile is an equilibrium.[5] It should also be noted that this is a strict perfect Bayesian equilibrium and therefore no type of either country is playing a weakly dominated action. Thus, this equilibrium is not ruled out by any standard refinement argument.

[5] For this range of c_i, there also exists a perfect Bayesian equilibrium in which some types of both countries mix. This equilibrium thus also involves a positive probability of war.

	A	B	C	
A	(.5,.5)	(.5,.5)	(.65,.35)	Pess.
B	(.5,.5)	(.5,.5)	(.65,.35)	Pess.
C	(.35,.65)	(.35,.65)	(.5,.5)	Pess.
	Pess.	Pess.	Pess.	

Figure 6: Type pairs at war and peace

2.5 Example 2: Implications

We summarize the outcomes of the equilibrium we have just described in Figure 6. In this figure, the type pairs for which war occurs are shaded and the type pairs for which the peaceful settlement occurs are unshaded.

We again make two main observations about this example. The first observation is that war can occur in the absence of optimism on either side. Although this game has a peaceful equilibrium, it also has an equilibrium in which war occurs, even though no type of either country is optimistic. This example thus strengthens our earlier argument that mutual optimism is not necessary for war.

The second observation is that the probability of war can be high, even in the absence of optimism. The ex ante probability of war in this equilibrium is 8/9. Comparing this information structure to that given in Example 1, we see that even though there is less optimism, there is a higher probability of war. This finding again calls into doubt the connection between mutual optimism and war.

3 General Model

The example given in the previous section generates several suggestive observations about the connection between mutual optimism and war. But are these observations limited to our specific example or are they more broadly applicable? To address this question in this section we develop a general model of war in order to provide general results.

Two countries face a potential conflict that can be settled either by force or by a negotiated settlement. We suppose that any negotiated settlement is efficient, but that war is inefficient. We also suppose that a war can be started by either side, unilaterally. To explore the role that private information plays in this choice, we assume that there is a set Ω of possible states of the world. Each possible state of the world, denoted ω, is a complete description of both countries' capabilities and prospects for war. As is standard, we suppose both countries share a *common prior* π on Ω and focus on how differences in information might lead to the choice of war.

In order to incorporate these states of the world into a conflict game, Fey & Ramsay (2007) present a general model of knowledge, which we summarize briefly here and connect to the perhaps more familiar framework of Bayesian games. In this model of knowledge, we associate information or knowledge with the ability to distinguish between various states ω in Ω. We assume Nature initially draws the true state of the world according to the common prior π. Nature then provides information to players in the form of a signal about the true state of the world. The (deterministic) signal function of player i is denoted $t_i(\omega)$. In this setting, the type space of player i, T_i, is the range of the function $t_i(\omega)$. That is, the set of *types* of player i is just the set of all possible signals for player i. As Ω is assumed to be finite, the set T_i is also finite. The inverse image of the signal function, $t_i^{-1}(t_i^k)$, gives the set of states that could give rise to type t_i^k. These inverse image sets are important in the following way. Let $P_i(\omega) = \{\omega' \mid t_i(\omega) = t_i(\omega')\}$. We call $P_i(\omega)$ a *possibility correspondence*. For each $\omega \in \Omega$, $P_i(\omega)$ is interpreted as the collection of states that individual i thinks are possible when the true state is ω. This is one example of an *event*, which are naturally defined as subsets of Ω. A possibility correspondence $P_i(\omega)$ for Ω is *partitional* if there is a partition of Ω such that for any $\omega \in \Omega$ the set $P_i(\omega)$ is the element of the partition that contains ω. As discussed in Fey & Ramsay (2007) (and proved by others), a fully rational player must have a partitional possibility correspondence.

We know turn to incorporating this model of knowledge into a general model of war. We define two functions, $p_1(\omega)$ and $p_2(\omega)$, that specify the probability that country 1 and 2 will win a war, given the true state of the world ω. Of course, $p_1(\omega) + p_2(\omega) = 1$ and $0 \leq p_i(\omega) \leq 1$ for all values $\omega \in \Omega$. Consider an arbitrary event E. If a country knows an event $E \subseteq \Omega$ has occurred, it can combine this information with the prior π via Bayes' Rule to form a posterior belief about the value of p_i as follows:

$$\mathrm{E}[p_i|E] = \frac{\sum_{\omega \in E} p_i(\omega)\pi(\omega)}{\sum_{\omega \in E} \pi(\omega)} \tag{1}$$

From this expression, it is easy to verify that if $\mathrm{E}[p_i|E'] \geq x$ and $\mathrm{E}[p_i|E''] \geq x$ for disjoint sets of states E' and E'', then $\mathrm{E}[p_i|E' \cup E''] \geq x$. This result is known as the Sure Thing Principle (Savage 1954).

We normalize the utility of countries to be 1 for victory in war and 0 for defeat, and we suppose there is a cost $c_i(\omega) > 0$ of fighting a war for country i. Thus, in the event of war at state ω, the expected utility of country i is $p_i(\omega) - c_i(\omega)$. Similarly, it is possible that the potential negotiated settlement will depend on the private information of the two sides. Therefore, we define two additional functions, $r_1(\omega)$ and $r_2(\omega)$, that specify the bargaining outcome when the true state of the world is ω. Since bargaining is efficient, we assume that $r_1(\omega) + r_2(\omega) = 1$ for all values $\omega \in \Omega$. Given a true state ω, a country can combine its knowledge of $P_i(\omega)$ with the prior π via Bayes' Rule (equation 1) to construct its individual belief about the probability it will win, $\hat{p}_i(\omega) = \mathrm{E}[p_i|P_i(\omega)]$, the cost of fighting $\hat{c}_i(\omega) = \mathrm{E}[c_i|P_i(\omega)]$, and its expected payoff from bargaining, $\hat{r}_i(\omega) = \mathrm{E}[r_i|P_i(\omega)]$. In this setting, we say that country i is *optimistic* at ω if $\hat{p}_i(\omega) - \hat{c}_i(\omega) > \hat{r}_i(\omega)$. If exactly one country is optimistic at ω, then we say *unilateral optimism* occurs at ω; if both sides are optimistic at ω, then we say *mutual optimism* occurs at ω.

We end this section by describing the class of games that we analyze. In order to be as general as possible and to cover as many different varieties of strategic interaction, we describe an abstract class of games. Let the set of actions for player i in a two-player strategic form game be given by the set A_i. The result of the choice of actions for the two sides will be either war or a peaceful settlement. We assume that war is a *unilateral act*, so that either side can start a war. Formally, war is a unilateral act if, for each i, there is an action $\bar{a}_i \in A_i$ such that whatever action is chosen by the opponent, the outcome is war. To avoid redundancy, we assume that the action $\bar{a}_i \in A_i$ is the unique action with this property and, to avoid triviality, we assume that there is some action profile that results in a peaceful settlement, as well.

Finally, we define a pure strategy $s_i \in S_i$ as a function $s_i : \Omega \to A_i$ with the restriction that

$$P_i(\omega) = P_i(\omega') \quad \Rightarrow \quad s_i(\omega) = s_i(\omega').$$

This condition states that if a country cannot distinguish state ω from state ω', then its action must be the same in both states. For a given strategy profile (s_1, s_2), if there is a positive probability that the war outcome results from the play of this strategy profile, we say that (s_1, s_2) is a strategy profile in which *war occurs*. Since we have specified a strategic form game with incomplete information, the appropriate solution concept is Bayesian-Nash equilibrium. Note that under the assumption that war is a unilateral act, a pure strategy Bayesian-Nash equilibrium always exists, namely the strategy profile in which every type of country 1 chooses action $\bar{a}_1$ and every type of country 2 chooses action $\bar{a}_2$.

4 General Results

In this section we present two general results that apply to the broad class of games defined in the previous section. We first show that the existence of unilateral optimism precludes peace. We then give a result that show that mutual optimism is never necessary for war. Finally, we discuss how these results extend to cases in which actors are not fully rational in their decision makers.

Throughout this section, let G denote an arbitrary strategic form game of incomplete information that satisfies our assumptions on the information structure, payoffs, and strategies given in the preceding section.

4.1 Unilateral Optimism

Our first result states that if at least one type of one country is optimistic, then there is a positive probability of war in equilibrium. That is, the possibility of unilateral optimism precludes peace. In addition, the converse of this statement is also true. If neither country has an optimistic type, then there exists a peaceful equilibrium.

Theorem 1 *Let G denote an arbitrary strategic form game of incomplete information in which war is a unilateral act. Then there is a positive probability of war in every pure strategy Bayesian-Nash equilibrium of G if and only if there is a state ω and a country i that is optimistic at ω.*

Proof: We begin by showing that if there is a state ω and a country i that is optimistic at ω, then in every pure strategy Bayesian-Nash equilibrium of G there is a positive probability of war. For a proof by contradiction, suppose that there is a game G with a state ω and a country i that is optimistic at ω and a pure strategy Bayesian-Nash equilibrium with zero probability of war. This means that type $t_i(\omega)$ of country i is not choosing action $\bar{a}_i$ in equilibrium and, moreover, this type's equilibrium payoff is $\hat{r}_i(\omega)$. If this type deviates to action $\bar{a}_i$, however, its payoff is $\hat{p}_i(\omega)-\hat{c}_i(\omega)$. Since country i is optimistic at ω, $\hat{p}_i(\omega)-\hat{c}_i(\omega) > \hat{r}_i(\omega)$ and therefore this deviation is profitable. This contradicts the existence of such an equilibrium.

For the reverse direction, we suppose that there is no state ω for which either side is optimistic and show the existence of a peaceful equilibrium. To do so, fix a strategy profile that gives the peaceful settlement at every state $\omega \in \Omega$. This is possible because, by assumption, there exists an action profile that results in the peaceful settlement. For this strategy profile, the expected payoff of type $t_i(\omega)$ of country i is $\hat{r}_i(\omega)$. Deviating to some other action will either result in a peaceful settlement with probability one, war with probability one, or both outcomes with some positive probability. Thus, the payoff to deviating of type $t_i(\omega)$ of country i is a convex combination of $\hat{p}_i(\omega) - \hat{c}_i(\omega)$ and $\hat{r}_i(\omega)$. But because no type is optimistic, we have $\hat{p}_i(\omega) - \hat{c}_i(\omega) \leq \hat{r}_i(\omega)$. Therefore this is not a profitable deviation. Thus, such a strategy profile is indeed a peaceful equilibrium. ∎

This theorem states that the existence of optimism on the part of a single country is enough to ensure that war occurs in equilibrium. The logic is simple: in a completely peaceful equilibrium, an optimistic type has an incentive to fight. Importantly, optimism on only one side is enough for this result; it does not require mutual optimism. This fact illustrates that it is not mutual optimism that drives the occurrence of war, instead one-sided optimism is enough. Of course, mutual optimism can be present when war occurs, but our result points out that it is not a requirement for war. Moreover, if both sides *always* lack optimism, then there exists a peaceful equilibrium.

It may be tempting to conclude from Theorem 1 that although the mutual optimism explanation is unsatisfactory, we are replacing it with an unsurprising result that countries fight when they think they are better off by fighting and they choose not to fight when they don't think this. However, the truth turns out to be significantly more subtle than this. First, as we demonstrated in Example 1 in Section 2, it is not the case that all optimistic types end up fighting. It is possible that a type who initially thinks fighting is better can rationally choose not to fight based on the strategic inference about what is actually true in the states where its choice matters. Second, Example 2 in Section 2 illustrates that it is not the case the optimism is needed for war to occur. To be clear, in this example there exists a peaceful equilibrium, as required by Theorem 1, but there also exists an equilibrium in which war occurs, even though no type of either country is optimistic. Again, the strategic inference that a side that is initially not optimistic makes can lead it to rationally choose to fight. This reinforces the point that mutual optimism is not always necessary for war by showing that even unilateral optimism is not always necessary for war.

4.2 Mutual Optimism

Our second result is a general statement about the relationship between mutual optimism and war. It proves that mutual optimism is not necessary in the entire class of games with unilateral war and not just in specific examples. That is, there are *no* examples of games with unilateral war in which war occurs *only* when mutual optimism holds.

Theorem 2 *Let G denote an arbitrary strategic form game of incomplete information in which war is a unilateral act. In every pure strategy Bayesian-Nash equilibrium of G in which war occurs, there is a state ω at which war occurs but mutual optimism does not hold.*

Proof: We begin by supposing that the strategy profile (s_1^*, s_2^*) is a Bayesian-Nash equilibrium in which war occurs. Denote the set of states for which the outcome of the game is war by W and denote the set of states for which the outcome is a peaceful settlement by T. As we are considering pure strategies, these two sets form a partition of Ω. Consider a state $\omega' \in T$. As each player can impose war by playing $\bar{a}_i$, and this deviation changes the payoff to player i only if war would not have occurred anyway, equilibrium requires

$$\begin{aligned} E[r_i(\omega) \mid P_i(\omega') \cap T] &\geq E[p_i(\omega) - c_i(\omega) \mid P_i(\omega') \cap T] \\ E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid P_i(\omega') \cap T] &\leq 0 \end{aligned} \tag{2}$$

for every $\omega' \in T$.

Now define the events

$$O_i = \{\omega \in \Omega \mid \hat{p}_i(\omega) - \hat{c}_i(\omega) > \hat{r}_i(\omega)\}$$

for $i = 1, 2$. To prove the theorem, suppose that the conclusion is false. That is, suppose that in every state that war occurs, mutual optimism also occurs. Formally, this requirement is that $W \subseteq O_1 \cap O_2$. Now, take an arbitrary $\omega \in W$. Because $\omega \in O_i$ for $i = 1, 2$, it follows that

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid P_i(\omega)] > 0, \quad i = 1, 2. \tag{3}$$

We claim that for an arbitrary $\omega \in W$,

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid P_i(\omega) \cap W] > 0, \quad i = 1, 2. \tag{4}$$

If $P_i(\omega) \cap T$ is empty, then $P_i(\omega) \cap W = P_i(\omega)$ and the claim follows from inequality (3). If $P_i(\omega) \cap T$ is nonempty, then there is some $\omega' \in P_i(\omega)$ such that $\omega' \in T$. Therefore, by inequality (2),

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid P_i(\omega) \cap T] \leq 0.$$

As W and T form a partition of Ω, this implies that inequality (4) must hold because otherwise the Sure Thing Principle would generate a contradiction with inequality (3). Thus, in either case, inequality (4) holds.

As the correspondence P_i is partitional, we can define a set of states D_i^* with $D_i^* \subseteq W$ such that the sets $\{P_i(\omega)\}_{\omega \in D_i^*}$ are disjoint and

$$\bigcup_{\hat{\omega} \in D_i^*} [P_i(\hat{\omega}) \cap W] = \bigcup_{\hat{\omega} \in W} [P_i(\hat{\omega}) \cap W] .$$

Since $D_i^* \subseteq W$, we have from inequality (4) that

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid P_i(\omega) \cap W] > 0$$

for every $\hat{\omega} \in D_i^*$. As this holds for each disjoint set $P_1(\hat{\omega})$, then by the Sure Thing Principle the same conditional expectation inequality holds over the union of these disjoint sets. Therefore,

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid \bigcup_{\hat{\omega} \in D_i^*} [P_i(\omega) \cap W]] > 0$$
$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid \bigcup_{\hat{\omega} \in W} [P_i(\omega) \cap W]] > 0.$$

As $\omega \in P_i(\omega)$ for every ω, it follows that

$$\bigcup_{\hat{\omega} \in W} [P_i(\omega) \cap W] = W.$$

We conclude that, for $i = 1, 2$,

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid W] > 0.$$

From this, it follows that

$$E[p_1(\omega) - c_1(\omega) - r_1(\omega) \mid W] + E[p_2(\omega) - c_2(\omega) - r_2(\omega) \mid W] > 0$$
$$E[p_1(\omega) + p_2(\omega) \mid W] - E[r_1(\omega) + r_2(\omega) \mid W] > E[c_1(\omega) + c_2(\omega) \mid W]$$

As $p_1(\omega)+p_2(\omega) = 1$ and $r_1(\omega)+r_2(\omega) = 1$ for all $\omega \in \Omega$, it follows that $E[p_1(\omega)+p_2(\omega) \mid$

$W] = 1$ and $E[r_1(\omega) + r_2(\omega) \mid W] = 1$. Thus we have

$$0 > E[c_1(\omega) + c_2(\omega) \mid W]$$

But this contradicts the fact that $c_i(\omega) > 0$ for all $\omega \in \Omega$, which proves the result. ∎

The theorem shows that there cannot be an equilibrium to a game in which countries have the ability to unilaterally start a war where mutual optimism is a necessary condition for costly conflict. This does not mean that mutual optimism and war cannot occur together in equilibrium, but rather that it is never necessary. Any game with such an equilibrium must have other realizations of the state of the world—in the particular equilibrium—where war occurs and there is no mutual optimism. Put more simply, Theorem 2 establishes that mutual optimism is never a necessary condition for war by showing that if war happens with mutual optimism, it must also occur without mutual optimism.

4.3 Bounded Rationality

In earlier work, Fey & Ramsay (2007) show that their arguments about the logical connection between mutual optimism and war continue to hold even when actors are not fully rational in their decision making. Thus, while Theorem 2 is true for a broad class of games in which the decision-makers rationally process information, here we pause to consider whether this result depends on strictly rational learning. We find that even if actors' information processing suffers from cognitive biases, the link between mutual optimism and war is still quite weak. In particular, even if both players ignore "bad news" or are inattentive, then mutual optimism is never necessary for war.

Following Fey & Ramsay (2007), we examine an information structure that captures several possible kinds of information processing errors found in the psychological international relations literature (Jervis, Lebow & Stein 1985, Jervis 1976). This information structure is defined as follows.

Definition 1 *Let P_i be a possibility correspondence for individual i. We say P_i is*

1. nondeluded *if, for all $\omega \in \Omega$, $\omega \in P_i(\omega)$, and*

2. nested *if for all $\omega, \omega' \in \Omega$, either $P_i(\omega) \cap P_i(\omega') = \emptyset$, or $P_i(\omega) \subseteq P_i(\omega')$, or $P_i(\omega') \subseteq P_i(\omega)$.*

An individual with this kind of possibility correspondence may "ignore" or "throw out" information that would be known to a fully rational Bayesian. This formalization is consistent with many forms of information processing bias, because it is agnostic to the reason information is ignored. Individuals could fail to learn in some states because acquiring information is costly, because they are inattentive, or because they would rather not think about the implications of the information in front of them.

We now establish that even with actors that do not process information in a fully rational way, mutual optimism is still never necessary for war.

Theorem 3 *Let G denote an arbitrary strategic form game of incomplete information in which war is a unilateral act, countries have a common prior, and P_i is nondeluded and nested for $i = 1, 2$. In every pure strategy Bayesian-Nash equilibrium of G in which war occurs, there is a state ω at which war occurs but mutual optimism does not hold.*

Proof: The proof of this theorem is very similar to the proof of Theorem 2. As in that proof, fix a pure strategy equilibrium (s_1^*, s_2^*) in which war occurs and let the set of states for which the outcome of the game is war be W and the set of states for which the outcome is a peaceful settlement be T. Using the exact same argument as in the proof of Theorem 2, we can establish that equation (4) continues to hold. That is, for an arbitrary $\omega \in W$,

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid P_i(\omega) \cap W] > 0, \quad i = 1, 2. \tag{5}$$

Because information partitions can be nested, a given state of the world could belong to multiple information partitions. So let $M_i(\omega)$ be the largest set (with respect to set inclusion) of the collection of sets $\{P_i(\omega') \mid \omega \in P_i(\omega')$. By nestedness, $M_i(\omega)$ is well-defined for all $\omega \in \Omega$. Moreover, by non-deluded and nestedness, for all $\omega, \omega' \in \Omega$, either $M_i(\omega) = M_i(\omega')$ or $M_i(\omega) \cap M_i(\omega') = \emptyset$. Therefore the $M_i(\omega)$ sets form a partition of Ω. Enumerate the sets that make up this partition for i as $M_i^1, M_i^2, \ldots, M_i^K$. For each set M_i^k such that $M_i^k \cap W \neq \emptyset$, let

$$\bar{P}_i^k = \bigcup_{\omega \in M_i^k \cap W} P_i(\omega) \cap W$$

By nestedness, $\bar{P}_i^k = P_i(\omega') \cap W$ for some $\omega' \in W$. Therefore, by equation (5), we have

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid \bar{P}_i^k] > 0, \quad i = 1, 2.$$

Moreover, by non-deluded, $\bar{P}_i^k = M_i^k \cap W$. As the M_i sets form a partition of Ω, the $\bar{P}_i^k$ sets just defined form a partition of W. Thus we can then write W as the union of disjoint sets $\bar{P}_i^k$, defined by some collection of states $\hat{D}^*$ all contained in W, i.e., $\hat{D}^* \subseteq W$. The result then follows as in Theorem 2. ∎

Theorem 3 show that for some plausible types of "boundedly rational" actors, mutual optimism cannot be necessary for war. In other words, if war and mutual optimism occur simultaneously at some state in an equilibrium in G, then there must be some another state where there is war and no mutual optimism. Therefore, the mutual optimism result in Theorem 2 is not fragile. Clearly, some departure from rational Bayesian learning is acceptable and consistent with our results. In particular, if decision-makers sometimes ignore unpleasant information or behave as if they have imperfect memory, then our result survives.

5 Mutual Optimism and Bargaining

Up to this point, we have focused on the class of models in which each side chooses between starting a war and accepting a peaceful settlement fixed at each state ω. But the requirement that the settlement may only vary by the underlying state of the world may be too strong—there is a robust literature that views international conflict as being a type of bargaining failure. That is, the peaceful settlement may be endogenous to the actions of the two sides and it could be that this changes the connection between mutual optimism and war. In this section we examine this question. We first consider some existing work that purports to connect mutual optimism and war in a bargaining context. Next, we consider two versions of a standard bargaining model with two-sided uncertainty. In the first version each side has two possible types ("strong" or "weak") and in the second version each side has an infinite number of possible types. We show that in the first version mutual optimism is not necessary for war and in the second version mutual optimism is either not necessary or not sufficient for war. Thus we conclude that, as in our earlier models, in the bargaining context mutual optimism is not a satisfactory explanation for war.

5.1 Existing work on mutual optimism and bargaining

In recent work, Slantchev & Tarar (2011) give an example of a take-it-or-leave-it bargaining game with one-sided incomplete information that they claim shows that mutual optimism can be both necessary and sufficient for war. 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 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.

Leaving this matter aside, there is an additional difficulty with Slantchev & Tarar's (2011) claim that mutual optimism is necessary and sufficient for war in their example. Understanding this difficulty 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. Instead, mutual optimism needs to be specified in terms of some given parameters of the model, independent of the occurrence of war. We do this in the next section.

But before we do so, we can say even more about the example of Slantchev & Tarar (2011). Forgetting for the moment the concerns just raised, let us take their claims about their example with one-sided incomplete information at face value. As a proper analysis of mutual optimism really requires two-sided incomplete information, we will modify the example 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 then 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 $p_w > p_s$.[6] 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).[7]

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 every such 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*

[6]This notation differs from Slantchev & Tarar (2011) in that we define p to be the probability that country 1, rather than country 2, wins a war.

[7]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.

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.

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 variant type has a higher war payoff, the D1 refinement requires that country 2 place probability one on such an offer coming from a variant 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.

5.2 Two-sided uncertainty with two types

In the previous section, we considered a bargaining model with two-sided uncertainty in which the two sides had very different levels of uncertainty. Country 1 had substantial uncertainty about the strength of country 2, but country 2 had only a small amount of

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

Figure 7: Probabilities of winning: (p_1, p_2)

uncertainty about country 1. In this section, we consider a bargaining model that has two-sided uncertainty and two possible types for each side, but unlike the previous model, has equal levels of uncertainty on both sides. We show that in this model, mutual optimism is not necessary for war. Thus, our main conclusion from the previous section continues to hold in a model with symmetric uncertainty.

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 7. 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.

How should mutual optimism be defined in this context? While we do not have a fixed settlement to compare against, we would like to maintain the intuition of our earlier definition. That is, we seek to define mutual optimism as a situation in which the unconditional beliefs about winning for the two sides are incompatible. Specifically, we say that a type profile (t_1, t_2) has mutual optimism if the naive conditional value of war for type t_1 of country 1 plus the naive conditional value of war for type t_2 of country 2 is greater than one. From Figure 7, we can see that the value for war for the W type is $1/2 - a/2 - c$ and the value of war for the S type is $1/2 + a/2 - c$. Thus mutual optimism is only possible 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 in all equilibria 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 in Section 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.[8] This makes our case that mutual optimism is not necessary for war even stronger.

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

5.3 Two-sided uncertainty with a continuum of types

Is the result in Proposition 2 limited by each side having only two types or does it hold in a more general model? To answer this question, we next consider a version of this model in which both counties have a continuum of possible types. We suppose each country's type t_i is independently drawn from a continuous distribution F with support $[\underline{t}, \bar{t}] = T$. As in the rest of the paper, these types determine the probability that country 1 prevails in a war, which we denote $p(t_1, t_2)$. We assume that p is continuous, strictly increasing in t_1 and strictly decreasing in t_2. As in the previous section, we assume that this war-fighting technology is symmetric with respect to the types of the two countries. Formally, we suppose that for all $a, b \in T$, $p(a, b) = 1 - p(b, a)$. The structure of the game remains the same. That is, country 1 makes an offer $(x, 1-x)$ which is accepted or rejected by country 2. If country 2 accepts, the payoffs are $(x, 1-x)$. If country 2 rejects, both sides receive their war payoff, which is $p(t_1, t_2) - c$ for country 1 and $1 - p(t_1, t_2) - c$ for country 2. To avoid complications involving boundary conditions, we assume that for all (t_1, t_2) pairs, $p(t_1, t_2) + c \leq 1$ and $p(t_1, t_2) - c \geq 0$. Note that we do not impose any other restrictions on the functional form for this war-fighting technology. We also allow the distribution of uncertainty, represented by F, to take on any (continuous) shape. Thus our results apply to a wide range of potential conflict scenarios.

We define mutual optimism in the same way as in the two-type model. That is, mutual optimism occurs at a type pair (t_1, t_2) if the expected (unconditional) value for war for type t_1 of country 1 and the expected (unconditional) value for war for type t_2 of country 2 sum to more than one. Formally, we say mutual optimism holds at the type pair (t_1, t_2) if

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

which, by symmetry, can be written as

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

The equilibrium concept we utilize is perfect Bayesian equilibrium. As is well known, this equilibrium concept allows complete latitude in specifying beliefs off the equilibrium path. Because of this, the set of perfect Bayesian equilibria to our game is very large. To deal with this, we focus on perfect Bayesian equilibria that satisfy the additional requirements of the D1 refinement (Cho & Kreps 1987). 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 our game. We let $x^*(t_1)$ be the equilibrium offer made by type t_1 of country 1.[9] 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

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

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 $(x, 1-x)$.[10] 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

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

[10] 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.

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.

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

Proof: We begin by showing that 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. So 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 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

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

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.

We now use this result to prove our statement about mutual optimism and war. Specifically, we will show that mutual optimism is either not necessary or not sufficent 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 simpler models carry over to a general version of the standard bargaining model with continuous types.

6 Conclusion

It has long been argued that mutual optimism is an important causal explanation for war between countries. The work of uncovering how this mechanism operates in a strategic environment has largely been eschewed in favor of the collection of historical anecdotes of high level political and military decision-makers espousing optimistic views of the upcoming war. But before evidence can be brought to bear on the usefulness of a theory, the validity of the underlying argument must be verified. Fey & Ramsay (2007) attempt such a validation only to find that a rigorous analysis of the mutual optimism argument shows no link between mutual optimism and war. In that article, the analysis assumes that both countries must choose war for it to occur. Here we have taken up the question of how mutual optimism fairs as a theory for explaining war if we allow war to be the result of a unilateral choice by a single country or the rejection of offers in a bargaining game.

Our summary finding is that the mutual optimism argument fares no better in an environment with unilateral war or explicit bargaining. As argued above, mutual optimism would be a useful theory of war if it were either a necessary or sufficient for war in equilibrium. We have presented a simple example in Section 2 that shows mutual optimism fails to be either one. We have also presented several general results that hold for every game with unilateral war. In these environments the possibility of unilateral optimism is sufficient for war to occur in equilibrium. We also show that mutual optimism is never necessary for war.

Moreover, we have established that these findings extend to bargaining games with two-sided incomplete information as well. In sum, we have demonstrated that mutual optimism is not necessary or not sufficient for war, or both. Together these results seriously undermine the case for mutual optimism as a rationalist explanation for war. In addition, these results hold even if we relax the assumption that decision-makers are perfect Bayesian learners.

Therefore, in an important way, mutual optimism and war are just coincidental. Unilateral optimism, however, is a more important marker for war. While optimistic states may or may not fight in any realized situation, the possibility of unilateral optimism alone implies that the *ex ante* probability of war is always positive in any pure strategy equilibrium. The analogy in the introduction is instructive. Just because we often see two students in their office with the light on does not mean that their mutual presence is necessary; either one could turn the light on by themselves. This is precisely what we show must happen: if there is a state of the world with mutual optimism and war, there must also be a state of the world with war but without mutual optimism.

The various approaches we take when modeling mutual optimism are united by some important common themes. Our models all focus on uncertainty about the balance of power. More specifically, we analyze conflict games in which the two sides have private information about the likelihood of prevailing in war. There are three important aspects of this focus. First, it is crucial that any model of mutual optimism must have two-sided incomplete information. In our view, there is no other way to study the informational condition of *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. Third, our study of environments with uncertainty about the probability of winning leads us to focus on the "strategic inferences" that a rational decision-maker should make about the information of an opponent in an equilibrium theory of war. Future work on such models will help us to understand in what way they are different from the standard models of war and how these differences affect our thinking about international conflict.

References

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.

Feddersen, Timothy J. & Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review* 86(3):408–24.

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.

Jervis, Robert. 1976. *Perception and Misperception in International Relations.* Princeton, NJ: Princeton University Press.

Jervis, Robert, Richard Ned Lebow & Janice Gross Stein. 1985. *Psychology and Deterrence.* Baltimore, MD: The John Hopkins University Press.

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

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

Savage, Leonard J. 1954. *The Foundations of Statistics.* New York: John Wiley.

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

Thaler, Richard H. 1994. *The Winner's Curse: Paradoxes and Anomalies of Economic Life.* Princeton University Press.

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

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

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.

Reviewer's Appendix

Here we show that there is a unique perfect Bayesian equilibrium to Example 1 in Section 2. The set of actions of both countries is $A_i = \{F, N\}$ and the set of types of each country is $T_i = \{A, B, C\}$. Thus a (mixed) strategy for country i gives the probability that each type t_i plays N, which we denote by $\sigma_i(t_i)$. As the game in the example is an extensive form game, we must also define the beliefs of country 2. Let $\mu_k = P[t_1 = k \mid a_1 = N]$ for $k \in \{A, B, C\}$ denote such a belief. Using these beliefs, it is sequentially rational for type t_2 of country 2 to choose F if

$$\begin{gathered} Eu_2(F \mid t_2, \mu) \geq Eu_2(N \mid t_2, \mu) \\ \mu_A p_2(A, t_2) + \mu_B p_2(B, t_2) + \mu_C p_2(C, t_2) - c_2 \geq 1/2 \\ \mu_A p_2(A, t_2) + \mu_B p_2(B, t_2) + \mu_C p_2(C, t_2) \geq 1/2 + c_2. \end{gathered}$$

Turning now to the choice of country 1, it is sequentially rational for type t_1 of country 1 to choose F if

$$\begin{gathered} Eu_1(F \mid t_1, \sigma_2) \geq Eu_1(N \mid t_1, \sigma_2) \\ (1/3)[p_1(t_1, A) + p_1(t_1, B) + p_1(t_1, C)] - c_2 \geq (1/3)[(1 - \sigma_2(A))(p_1(t_1, A) - c_1) + \sigma_2(A)(1/2) \\ +(1 - \sigma_2(B))(p_1(t_1, B) - c_1) + \sigma_2(B)(1/2) + (1 - \sigma_2(C))(p_1(t_1, C) - c_1) + \sigma_2(C)(1/2)] \\ \sigma_2(A)(p_1(t_1, A) - c_1 - 1/2) + \sigma_2(B)(p_1(t_1, B) - c_1 - 1/2) + \sigma_2(C)(p_1(t_1, C) - c_1 - 1/2) \geq 0. \end{gathered}$$

The fact that each term in this expression is weighted by the $\sigma_2(t_2)$ reflects the fact that country 1's choice of action only matters if country 2 is choosing N.

We must show that the only perfect Bayesian equilibrium to this game is one in which $\sigma_i(A) = 0$ and $\sigma_i(B) = \sigma_i(C) = 1$ for $i = 1, 2$. To begin, consider a type C of country 2. It is easy to see from the above condition that there is no belief μ that makes fighting sequentially rational. Therefore $\sigma_2^*(C) = 1$ in any perfect Bayesian equilibrium. Using this, we see that because $p_1(C, t_2) \leq .5$ for all types t_2, type C of country 1 will never play F. Thus, $\sigma_1^*(C) = 1$ in any perfect Bayesian equilibrium. Now we turn to type A of country 1. For this type, F is strictly preferred to N if

$$\sigma_2(A)(-c_1) + \sigma_2(B)(-.2 - c_1) + (.4 - c_1) > 0.$$

But note that

$$\sigma_2(A)(-c_1) + \sigma_2(B)(-.2 - c_1) + (.4 - c_1) \geq (-c_1) + (-.2 - c_1) + (.4 - c_1) = .2 - 3c_1 > 0,$$

where the last inequality follows from $c_1 < 1/15$. This implies that $\sigma_1^*(A) = 0$ in any perfect Bayesian equilibrium. A similar argument establishes that $\sigma_2^*(A) = 0$ in any perfect Bayesian equilibrium. It then follows easily that $\sigma_1^*(B) = \sigma_2^*(B) = 1$. Thus, the equilibrium in which only the A type of each country fights is the unique perfect Bayesian equilibrium to this game.