Mutual Optimism and Unilateral War*

Mark Fey†
Kristopher W. Ramsay‡

Danger!
Preliminary and Incomplete
Work in Progress

April 10, 2009

Abstract

In this paper we examine whether mutual optimism can be supported as a cause of war in models that permit the unilateral use of force. Fey & Ramsay (2007) demonstrate that war cannot occur in any game in which *both* sides must agree in order to fight. With the more realistic assumption that either side can unilaterally use force, however, war is possible in equilibrium. But we show that mutual optimism is still not a valid explanation for war, because it is neither always necessary or sufficient for war in equilibrium.

*We thank .....
†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 cause of war that has received much attention in the literature is associated with optimism on the eve of conflict. As is well known, when leaders face uncertainty and have special or private knowledge that informs their assessment regarding the likely outcome of war, inconsistent expectations might be a cause of conflict. If leaders expectations are inconsistent in that both antagonists think they are likely to win a costly war, then there may not exist any peaceful settlement that both prefer to war. In these circumstances, the cause of war might then be inconsistent expectations that arise from private information (Fearon 1995, Blainey 1988, Wagner 1994, Kim & Bueno de Mesquita 1995).

In this paper we analyze a conflict game between two countries where both sides have private information regarding the likely outcome of war. In particular, we analyze the strategic incentives decision-makers face when choosing between fighting a war or opting for a peaceful settlement. The strategic problem for decision-makers in these circumstances is complicated by the interdependence of the players' payoffs in war. That is, unlike conflict games in which uncertainty is about "private valued" elements of a decision-maker's uncertainty, when the uncertainty concerns the probability of winning a war a player must account for the fact that their opponent's private information (or type) is relevant for their own payoff to war directly. This in turn leads us to focus on the "strategic inferences" that a rational decision-maker should make about the play of an opponent.

In its most common form, uncertainty and private information about the probability of war is associated with discussion of optimism and mutual optimism. In Fey & Ramsay (2007), this mutual optimism argument is formalized and shown to be an inconsistent rationalist explanation for war. A critical assumption in Fey & Ramsay's (2007) argument was that wars only occurred with mutual agreement. That is, war could only occur when both sides agreed to fight. It was argued that for the concept of "war by mutual optimism" to have any substantive content one needed to assume war occurred only when *both* sides preferred it to peace. Otherwise, it is not clear what is "mutual" in mutual optimism story.

Clearly, the assumption that "war is a mutual act" in Fey and Ramsay (2007), is not the only assumption one could make regarding the strategic interaction of countries in the shadow of war. In fact, given the anarchic nature of the international system, it seems natural to assume that countries can always choose to break off negotiations and use force instead.

This paper examines how the mutual optimism argument fairs as a causal story under the assumption that any single country can start a war. Following Fearon (1991), we say that mutual optimism causes war if war occurs when mutual optimism is present and war does not occur when mutual optimism is absent. When war is a mutual act, the argument is simple. As we show in Fey and Ramsay (2007), war does not occur whether mutual optimism is present or absent, so mutual optimism cannot be a cause for war. When war is a unilateral act, the argument is more subtle. Our first set of results show that if war occurs in equilibrium, there must be states of the world in which war occurs but only one side thinks it is going to win, i.e., is optimistic. Thus the requirement that war does not occur when mutual optimism is absent is not satisfied. In other words, mutual optimism is not a necessary condition for war.

We also show by way of an example that mutual optimism is not a sufficient condition for war. That is, we show that war does not always occur when mutual optimism is present. In this way, mutual optimism fails to be an explanation of war between rational decision-makers. Building on this example, we next illustrate that contrary to our intuition that reducing uncertainty reduces the risk of war, it is possible that providing information to one side in a conflict can increase the likelihood of war. We expand on this idea by establishing a general result that shows that in any game in which war occurs, if we fix a state of the world in which there is mutual optimism and war occurs, then we can give information to one side in a way that there is no longer mutual optimism and yet war still occurs. In the final section, we analyze the prospect of war as an equilibrium phenomenon in this class of games by looking at the setting where players are boundedly rational in their information processing. Here we find, like the results in Fey and Ramsay (2007), the weak link between mutual optimism and war is not fragile. It extends to situations where players are boundedly rational information processors.

All of the results in our paper depend in a significant way on the strategic inferences that rational decision-makers should make in any game in which war is a unilateral act. 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 has additional information that is not present outside of equilibrium. 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).

2 Model and Assumption

In this section, we describe our model and introduce the appropriate notation. We will follow the notation in Fey & Ramsay (2007) closely. Interested readers should consult the original paper for more details.

Two countries face a potential conflict that can be settled either by force or by a negotiated settlement. For our initial results 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. For later use, we define the *join* of two partitions as the coarsest common refinement of the two partitions. In terms of knowledge, the join of the possibility correspondences of two players represents what players would know if their information were public instead of private.

We represent the information of players by assuming that each player receives a *signal* about the true state of the world. The signal function of player i is denoted $\tau_i(\omega)$. To parallel our previous development, we take the signal function to be deterministic. In this setting, the type space of player i, T_i, is the range of the function $\tau_i(\omega)$. That is, the set of types of player i is the set of all possible signals for player i. The inverse image of the signal function, $\tau_i^{-1}(t_i)$, gives the set of states that could give rise to type t_i. Using this inverse function, we can specify the probability that player i places on the states of world, given player i's signal. This probability is given by

$$p(\omega \mid t_i) = \begin{cases} p(\omega)/\sum_{\omega' \in \tau_i^{-1}(t_i)} p(\omega') & \text{if } \omega \in \tau_i^{-1}(t_i) \\ 0 & \text{if } \omega \notin \tau_i^{-1}(t_i) \end{cases}$$

We can make this setup richer by allowing for the possibility of *noisy signals* that provide partial information about the state of the world. In this case, each player has a signal distribution η_i that maps states of the world to probability distributions on signals. Thus $\eta_i(t_i \mid \omega)$ is the probability that player i will receive the t_i signal given the state of the world is ω. Using this, we can give a more general statement for the probability that player i places on the states of world, given player i's signal:

$$p(\omega \mid t_i) = \frac{\eta_i(t_i \mid \omega)p(\omega)}{\sum_{\omega' \in \Omega} \eta_i(t_i \mid \omega)p(\omega')}$$

Certain kinds of events have special importance. First, we say that an event E is *self-evident* event for a possibility correspondence P_i if and only if for all $\omega \in E$, $P_i(\omega) \subseteq E$. In other words, an event E is self-evident if, for any state in E, a player knows E has occurred. Second, a *public* event, unlike a private signal, is known to all players when it happens. Specifically, if E is a public event, then E is self-evident to all players.

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's 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 $p_i(\omega) - c_i(\omega)$.

It equally likely that the negotiated settlement will depend on the underlying state of the world. We now 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 in each state $r_1(\omega) + r_2(\omega) = 1$. It is then immediate that countries' beliefs regarding the outcome of the bargaining process will depend on their private information as well.

We represent the private information of country i by a possibility correspondence $P_i : \Omega \to 2^{\Omega}$, which we assume is partitional. Recall that $P_i(\omega)$ is the set of states that country i views as possible, given the true state ω. Given a true state ω, a country can combine its knowledge of $P_i(\omega)$ with the prior π and equation 1 to construct its posterior 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 *unilateral optimism* occurs at ω when for exactly one player $\hat{p}_i(\omega) - \hat{c}_i(\omega) > \hat{r}_i(\omega)$. We say *mutual optimism* occurs at ω when $\hat{p}_i(\omega) - \hat{c}_i(\omega) > \hat{r}_i(\omega)$ for both players.

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 some 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.

Finally, we define a (pure strategy) 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 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 country i chooses action $\bar{a}_i$

3 Results

In this section we give our main results. We show that mutual optimism is neither necessary nor sufficient for war.

3.1 Mutual Optimism is not Necessary

Let G denote any strategic form game of incomplete information that satisfies our assumptions on the information structure, payoffs, and strategies. Our first result states that there is no game in which mutual optimism is a necessary condition for war.

Theorem 1 *Suppose countries have a common prior, war is a unilateral act, and P_i is partitional 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: 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 any game, in a broad class 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.

	t_2^1	t_2^2	t_2^3
t_1^1	.5	.3	.9
t_1^2	.7	.5	.5
t_1^3	.1	.5	.5

Figure 1: Information Structure of Example

3.2 Mutual Optimism is not Sufficient

Given that mutual optimism is never necessary for war, we now turn to sufficiency. As the following example demonstrates, it is possible that mutual optimism can occur without war. To describe this example, we start with the information structure. Specifically, suppose there are nine states of the world, given by the cells in Figure 1. Assume that each country believes initially that each state of the world is equally likely. The entry in each cell is $p_1(\omega)$, the probability that country 1 will prevail in war at the given states of the world. Country 1 knows which row the true state is in, while country 2 knows which column the true state is in. Thus, country 1 has three types, corresponding to the rows of the figure and likewise country 2 has three types, corresponding to the columns of the figure.

Although in general the cost of war and the peaceful settlement can depend on the state of the world, for simplicity we suppose that the cost is fixed at $c_i \in (0, .1)$ and the peaceful settlement is fixed at 1/2 for both countries. Also for simplicity, we consider a simple game form in which each country chooses to fight or not fight, represented by $A_i = \{F, N\}$. If either country chooses action F, then war results. If both countries choose action N, then the peaceful settlement results.

It is easy to check that there are two pure strategy Bayesian-Nash equilibria to this game with the information structure given in Figure 1. One, less interesting, equilibrium is for all types of both countries to choose F. The more interesting equilibrium is for each country i to play action F if its type is t_i^1 and to play action N if its type is t_i^2 or t_i^3. To see that this strategy profile is indeed an equilibrium, consider country 1. Recall that because war is a unilateral act, country 1's choice matters only when country 2 is choosing action N. Given country 2's strategy, this occurs precisely when country 2 is type t_2^2 or t_2^3. Therefore, for type t_1^1 of country 1, conditional on its action mattering, its expected payoff for war is

$$E[p_1 \mid t_1 = t_1^1, t_2 \in \{t_2^2, t_2^3\}] - c_1 = \frac{.3 + .9}{2} - c_1 = .6 - c_1$$

As $c_1 < .1$, choosing F is superior to choosing N. Similarly, for type t_1^2 or t_1^3 of country 1, its expected payoff for war conditional on its action mattering is $(.5 + .5)/2 - c_1 = .5 - c_1$. As $c_1 > 0$, these types of country 1 prefer to choose N rather than F. As the game is symmetric, the same analysis applies to country 2.

As the final step in the example, consider where mutual optimism occurs in Figure 1. In particular, consider the state of the world in the center cell of the figure, corresponding to the type pair (t_1^2, t_2^2). For country i, at this state of the world,

$$\hat{p}_i - c_i = \frac{.7 + .5 + .5}{3} - c_i = 17/30 - c_i$$

For $c_i \in (0, 1/15)$, $\hat{p}_i - c_i > r_i$ for both $i = 1, 2$ and thus mutual optimism occurs at this state of the world. However, in the interesting equilibrium we just described, both countries choose N at this state of the world. Therefore, we have found an equilibrium in which mutual optimism occurs at some state of the world but war does not occur. Therefore mutual optimism is not a general sufficient condition for war.

3.3 Converting Mutual Optimism to Unilateral Optimism

Theorem 1 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. More technically, the theorem showed that with a fixed information structure, if war occurred at some state of the world with mutual optimism then there must be another state with war but without mutual optimism. In our next result, we show that we can achieve the same result by holding the state of the world fixed and varying the information of the countries. Specifically, we show that if there is a state of the world with war and mutual optimism, then we can give information to one side in a way that war still occurs at this state, but mutual optimism no longer holds. Thus, the next theorem complements Theorem 1 and further undermines the claim that mutual optimism is a satisfactory explanation of war.

In this theorem, we modify the information structure of the given game by giving one country all of the information available to *either* country. This is accomplished by setting one side's information to be the join of the information partitions of the two sides. This is equivalent to letting one side see the realization of *both* countries' types, and not just their own.

Theorem 2 *Suppose countries have a common prior, war is a unilateral act, and P_i is partitional for $i = 1, 2$. Let P^* be the join of the partitions P_1 and P_2.*

For any game G, if mutual optimism holds at a state ω, then there exists a player $i^* \in \{1,2\}$ such that letting $P_i = P^*$ and leaving everything else unchanged, in every pure strategy Bayesian-Nash equilibrium of G, war occurs at ω but mutual optimism does not hold at ω.

Proof: Let ω^* be a state in which war occurs and mutual optimism holds. Then $\omega^* \in O_1 \cap O_2$, where O_1 and O_2 are the events described in Theorem 1.

To prove the theorem, let P^* be the join of the partitions P_1 and P_2. That is, P^* is the coarsest common refinement of P_1 and P_2. We first show that there exists $i^* \in \{1,2\}$ such that if $P_i = P^*$, then mutual optimism does not hold at ω^*. To show this, suppose not. That is, suppose that for both $i = 1,2$ we have

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

This implies that

$$\begin{aligned} E[p_1(\omega) - c_1(\omega) - r_1(\omega) \mid P^*(\omega^*)] + E[p_2(\omega) - c_2(\omega) - r_2(\omega) \mid P^*(\omega^*)] > 0 \\ E[p_1(\omega) + p_2(\omega) - c_1(\omega) - c_2(\omega) - r_1(\omega) - r_2(\omega) \mid P^*(\omega^*)] > 0 \\ E[1 - c_1(\omega) - c_2(\omega) - 1 \mid P^*(\omega^*)] > 0 \\ E[-c_1(\omega) - c_2(\omega) \mid P^*(\omega^*)] > 0. \end{aligned}$$

This contradiction establishes that such an i^* exists. So, for the rest of the proof, fix the country i^* and let $P_{i^*} = P^*$ and refer to the other country as j^*. We leave j^*'s information partition unchanged.

To complete the proof, we must show that in every equilibrium of the game G with the information structure given by information partitions P_{i^*} and P_{j^*}, war occurs at state ω^*. So, for a proof by contradiction, suppose that there exists a pure strategy equilibrium s^* to this game with this information structure in which war does not occur at state ω^*. As in the proof of Theorem 1, for this assumed equilibrium, 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 assuming that war does not occur at state ω^* in this equilibrium, we have $\omega^* \in T$ and therefore $s_i^*(\omega^*) \neq \bar{a}_i$ for $i = i^*, j^*$. In addition, define O^* to be the set of states in which country i^* prefers war to peace. That is,

$$O^* = \{\omega \in \Omega \mid E[p_{i^*}(\omega) - c_{i^*}(\omega) - r_{i^*}(\omega) \mid P_{i^*}(\omega)] > 0\} \tag{5}$$

First, note that if $\omega \in O^*$, then $\omega \in W$. This follows because if $\omega \in O^*$ and $\omega \in T$, then i^* can deviate to choosing action $\bar{a}_{i^*}$ at this state and thereby impose war at this state, which country i^* prefers because $\omega \in O^*$.

To proceed, observe that because $\omega^* \in O_1 \cap O_2$, $\omega^* \in O_{j^*}$ and so

$$E[p_{j^*}(\omega^*) - c_{j^*}(\omega^*) - r_{j^*}(\omega^*) \mid P_{j^*}(\omega^*)] > 0 \tag{6}$$

As $\omega \in O^*$ implies $\omega \in W$, we can write

$$P_{j^*}(\omega^*) = (P_{j^*}(\omega^*) \cap O^*) \bigcup (P_{j^*}(\omega^*) \cap O^{*C} \cap W) \bigcup (P_{j^*}(\omega^*) \cap T),$$

where O^{*C} is the complement of O^*. We consider two cases. The first case is the set $P_{j^*}(\omega^*) \cap O^*$ As P_{i^*} is the join of P_i and P_j, we know P_{i^*} is finer than P_{j^*} and therefore by equation (5), we have $E[p_{i^*}(\omega) - c_{i^*}(\omega) - r_{i^*}(\omega) \mid P_{j^*}(\omega^*) \cap O^*] > 0$, which implies that

$$E[p_{j^*}(\omega) - c_{j^*}(\omega) - r_{j^*}(\omega) \mid P_{j^*}(\omega^*) \cap O^*] < 0. \tag{7}$$

The second case is the set $P_{j^*}(\omega^*) \cap O^{*C} \cap W$. Pick an arbitrary ω' in this set. As $\omega' \in O^{*C}$, we have $E[p_{i^*}(\omega) - c_{i^*}(\omega) - r_{i^*}(\omega) \mid P_{i^*}(\omega')] \leq 0$. We claim that this expression cannot be negative. If it were, then at ω' country i^* would strictly prefer peace to war. But as $s_{j^*}(\omega') \neq \bar{a}_{j^*}$, there is some action available to country j that would change the outcome at this state of the world from war to peace. As this is a profitable deviation, we have a contradiction, which establishes that

$$E[p_{j^*}(\omega) - c_{j^*}(\omega) - r_{j^*}(\omega) \mid P_{j^*}(\omega^*) \cap O^{*C} \cap W] < 0. \tag{8}$$

From these two cases, we are able to conclude that

$$E[p_{j^*}(\omega) - c_{j^*}(\omega) - r_{j^*}(\omega) \mid P_{j^*}(\omega^*) \cap T] > 0. \tag{9}$$

If this expression were not positive, then using the Sure Thing Principle to combine this fact with inequalities (7) and (8), we would derive a contradiction with inequality (6). Therefore inequality (9) holds. But this inequality implies that country j^* would be strictly better off in state ω^* with war rather than peace. Therefore, country j^* should deviate to action $\bar{a}_{j^*}$. This contradicts the assumption that war does not occur at state ω^*, which proves the theorem. ∎

t_1^1	.5	.3	.9
t_1^2	.7	.5	.5
t_1^3	.1	.5	.5

Figure 2: Information Structure of Example

This theorem shows that anytime mutual optimism and war both occur, we can change just the information of one player, leaving everything else unchanged, removing the mutual optimism but leaving war as the unique equilibrium outcome.

3.4 More Information Can Lead to More War

Given the emphasis in the literature on uncertainty as a cause of war, it is intuitive to think that reducing the uncertainty in a conflict situation should make war less likely. However, when we apply the idea of providing one side with more information to the example given in Section 3.2, we arrive at the surprising conclusion that reducing uncertainty can actually lead to a greater chance of war.

To see this, recall that we found an equilibrium to the game in Section 3.2 in which each country chose to fight only when its type was t_i^1. Thus, war occurs in this equilibrium in five of the nine possible states of the world. Now give country 2 complete knowledge of the state of world, leaving the information for country 1 as before. This information structure is given in Figure 2. Here country 1's types correspond to the rows of the figure while country 2 knows which cell has been chosen by Nature.

As before, we suppose that cost of war is fixed at $c_i \in (0, 1/15)$ and the peaceful settlement is fixed at 1/2 for both countries. We also consider the simple game form in which $A_i = \{F, N\}$ and war occurs if and only if either country chooses action F.

In this setting, the following is a pure strategy Bayesian-Nash equilibrium. Country 1 chooses action F if its type is t_1^1 or t_1^2 and chooses action N if its type is t_1^3. Country 2, which knows that state of the world, chooses action F if and only if $p_1(\omega) < .5$. In other words, country 2 chooses to fight in exactly two states of the world, top row middle column and bottom row left column. Clearly, the strategy given for country 2 is optimal. To show that the strategy for country 1 is optimal, we again calculate the expected utility of war, conditional on country 2 choosing N. For type t_1^1, this expected utility is $(.5 + .9)/2 - c_1 = .7 - c_1$. For type t_1^2, this expected utility is $(.7 + .5 + .5)/3 - c_1 = 17/30 - c_1$. Finally, for type t_1^3, this expected utility is $(.5 + .5)/2 - c_1 = .5 - c_1$. Given that $c_i \in (0, 1/15)$, this establishes that country 1's strategy is optimal.[1]

The important point about this example is that even though country 2 no longer has any uncertainty, the likelihood of war goes up. Specifically, war now occurs in seven of the nine possible states of the world. Indeed, not only does the ex ante probability of war increase, but in fact it is true that in every state in which there was war before giving country 2 more information there is still war, plus there are additional states which go from peace to war.

What is the intuition for this result? The answer lies in the importance of the strategic inferences that rational decision-makers must make in this environment, namely that one side's choice only matters in states of the world in which the other side is not going to war. As demonstrated in Section 3.2, these strategic inferences lead both sides to sometimes choose war and sometimes not. Now think about what happens when we give country 2 complete information. How will country 2 use this information? The answer is that country 2 will now be more particular about which states of the world it choose to fight in. Specifically, it will continue to choose to fight in those states where it is likely to win, but it will choose to no longer fight in those states where it is likely to lose. But consider how this changes the choices of country 1. Remember that country 1's decision only matters in states of the world where country 2 is choosing not to fight. There are now more of these states, and because these states are those in which country 2 is likely to lose, *these states are exactly the ones in which country 1 is likely to win.* Therefore, war becomes more attractive to country 1. Put another way, because country 2 is starting fewer wars (those it thinks it will win), the average probability of victory for wars that country 1 can start has increased, which makes overall war more likely.

3.5 Mutual Optimism, Bounded Rationality, and Unilateral War

While Theorem 1 is true for any of the class of games in which the decision-makers rationally process information, one may wonder if the results depend on strictly rational learning. In this section, we consider a class of games where, again, two countries are choosing whether to fight a war or resolve the dispute by some other means. Here, we show that even if players' 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 war is only coincidently related to mutual optimism.

[1]There are exactly two other pure strategy equilibria. One has the same outcomes as described here—the only difference is that country 2 chooses N at the top row middle column state of the world. The other is the "total war" equilibrium in which both sides always choose F.

When it comes to information processing, a rational Bayesian may be able to deduce much more information from a "signal" than the signal carries at face value. That is, the rational Bayesian, like Sherlock Homes, learns from the dog that does not bark. There are, however, many cases in which we think that decision-makers, particularly the leaders of countries, may not be processing information rationally. Consider the information processing errors found in the psychological international relations literature (Jervis, Lebow & Stein 1985, Jervis 1976). For example, a decision-maker who has many responsibilities may face a volume of information that induces flaws in their learning. In particular, such a decision-maker may not update their beliefs when the state of the world is not explicitly brought to their attention. This error may occur because of a flaw in human psychology or it could be an information shortcut that allows decision-makers to deal with a world far more complex than the two state example above.

Alternatively, due to what Jervis, Lebow & Stein (1985, p.4) call *motivated bias*, a player's knowledge may be partly a matter of choice. So given that some people have strong predispositions to believe certain things to be true, this may prevent them from recognizing new information inconsistent with their world view. That is, sometimes decision-makers may consciously, or subconsciously, choose to ignore unpleasant information.

Next we consider a game with players whose information processing is flawed in ways consistent with the learning processes described above. A common component of these cognitive biases is that the player's information processing allows them to learn from new information in some states of the world, but not in others. To capture this idea formally, we define a new restriction on the players' possibility correspondences, P_i. In particular, while we still assume P_i is nondeluded, we now allow players to "ignore" or "throw out" information at a given state of the world that would be known to a fully rational Bayesian. To allow such pathologies, we must allow for the possibility that some information sets are nested within, or subsets of, other information sets. To do this, we define a new weaker possibility correspondence to capture these important departures from rationality.[2]

Definition 1 *A player's possibility correspondence is* nested *if for all $\omega, \omega' \in \Omega$, either (1) $P_i(\omega) \cap P_i(\omega') = \emptyset$ or (2) $P_i(\omega) \subseteq P_i(\omega')$ or (3) $P_i(\omega') \subseteq P_i(\omega)$.*

Possibility correspondences that satisfy nondeluded and nestedness represent a generalization of rational learning. That is, a decision-maker with a nested possibility correspon-

[2]For more on decision-theoretic approaches to bounded rationality in models of knowledge see Geanakoplos (1989) or Rubinstein (1998) for game theoretic approaches. The concept of nestedness is taken from Geanakoplos's (1989).

dence may processes information in a rational way or she may ignore new information at a number of different states. Such a formalization is consistent with many forms of bias, because it is agnostic to the reason information is ignored. Players 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 verify the robustness of Theorem 1.

Theorem 3 *Suppose 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 1. 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 1, we can establish that equation (4) continues to hold. That is, for an arbitrary $\omega \in W$,

$$E[p_i(\omega) - c_i(\omega) - r_i(\omega) \mid P_i(\omega) \cap W] > 0, \quad i = 1, 2. \tag{10}$$

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 (10), 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 1. ∎

Theorem 3 show that for some plausible types of "boundedly rational" actors, mutual optimism cannot be necessary for war. Or, in other words, if war and mutual optimism occur simultaneously at some state in an equilibrium in G, then there must be some other state where there is war and no mutual optimism.

As mentioned above, considering information structures that relax the requirements of strict Bayesian rationality can help us understand just how general our coincidence result is. On the one hand, the analysis in this section shows that the mutual optimism result in Theorem 1 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.

4 Conclusion

While much of the traditional literature on uncertainty and war in international relations focuses on uncertainty regarding the distribution of power and success in war, with few exceptions (Fey & Ramsay 2007, Powell 2004, Slantchev 2003, Wagner 1994) the formal literature on the causes of war has focused on uncertainty about the costs of war. In this paper we have focused on the special incentives that arise when uncertainty regarding the likelihood of winning a war is modeled directly. A few conclusions can be drawn. First, whether war is a unilateral act or some contest both players must agree to join, mutual optimism is not a good rationalist explanation for war. In particular, we show that there are no Bayesian Nash equilibria to a conflict game where players must choose between war and a negotiated settlement and where war occurs only when there is mutual optimism. Moreover, Theorem 2 says the distribution of equilibrium outcomes from an equilibrium where war and mutual optimism are coincidental can be replicated exactly in another equilibrium where there is no mutual optimism at any state of the world. These results hold even if we relax the assumption that decision-makers are perfect Bayesian learners.

Our second set of results grow out of the examples presented in Sections 3.2 and 3.4. The first of these examples showed that mutual optimism is not always sufficient for war. It is possible that mutual optimism can occur in some circumstance without war. The second of these examples illustrated the surprising result that reducing uncertainty can increase the likelihood of war. Together these results seriously undermine the case for mutual optimism as a rationalist explanation for war.

References

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

Fearon, James D. 1991. "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43(2):169–195.

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.

Geanakoplos, John. 1989. "Game Theory without Partitions, and Applications to Speculation and Consensus." *Cowles Foundation Discussion Paper No. 914* .

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.

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

Rubinstein, Ariel. 1998. *Modeling Bounded Rationality*. Cambridge, MA: MIT University Press.

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

Slantchev, Branislav L. 2003. "The Power to Hurt: Costly Conflict with Completely Informed States." *American Political Science Review* 47(1).

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

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