Mutual Optimism and War

Mark Fey†
Kristopher W. Ramsay‡

First Draft: August 2003
Current Draft: January 2006

Abstract

Working with the definition of mutual optimism as \textit{war due to inconsistent beliefs},
we formalize the mutual optimism argument and show that in a strategic setting; if war
is a “public event” when it occurs, then it cannot occur because of mutual optimism.
That is, we show that for a general class of games where countries contemplate war,
there is no Bayesian-Nash equilibrium where decision-makers choose war because of
mutual optimism.

\footnote{Earlier versions of this paper were presented at the 100th Annual meeting of the American Political Science Association, Chicago, IL, the 37th North-American meeting of the Peace Science Society, Ann Arbor, MI, and the University of Rochester’s Peter D. Watson Center for Conflict and Cooperation. We thank Songying Fang, Tanisha Fazal, Erik Gartzke, Shigeo Hirano, Adam Meirowitz, John Patty, Pablo Pinto, Robert Powell, Anne Sartori, Curt Signorino, Randy Stone, Branislav Slantchev, and Robert Walker, as well as other seminar participants. Any remaining errors are our own. Kris Ramsay acknowledges financial support from NSF grant SES-0413381.}

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

\footnote{Assistant Professor, Department of Politics, 033 Corwin Hall, Princeton University, Princeton, NJ 08544. phone: 609.258.2960, email: kramsay@Princeton.edu.}
1 INTRODUCTION

Why do states fight costly wars when less costly negotiated settlements are possible? Shouldn’t there be some mutually agreeable alternative to war that can produce the same result without incurring the social loss? Couldn’t decision-makers agree to distribute the disputed territory or assets in a way consistent with their beliefs about the likely outcome of conflict, saving both sides significant death and destruction? Blainey (1988) intimates that the high hopes on the eve of war suggest a sad conclusion; wars only occur when both rivals believe they can achieve more through fighting than through peaceful means. How might this be so? Obviously, if two countries are involved in a war, if one side wins then the other loses. We might then conclude that at least one side, in particular the loser, would prefer some peaceful method of resolving the dispute if she were certain of the outcome. But war is an uncertain process. Given this uncertainty the leaders of the two countries must each form expectations about the results of a conflict to guide their decision-making (Niou, Order-shook and Rose 1989, Stein 1990, Wagner 1994, Kim and Bueno de Mesquita 1995, Bueno de Mesquita, Morrow and Zorick 1997). Clearly, these expectations will be shaped by any special knowledge or information the leaders might poses. If these expectations are inconsistent in that both antagonists think their side will be better off fighting a war, the argument goes, then neither side would be willing to participate in a peacefully negotiated settlement (Blainey 1988, Wittman 1979). Such mutual optimism could thus lead to a rational choice of war by both countries (Morrow 1985, Werner 1998, Wagner 2000, Wittman 2001). In this setting, the root cause of war is the inconsistent expectations that arise because of private information (Fearon 1995, p.390).1 We can see this informally in our Figure 1, as the two shares that represent each countries expected benefit to war, when added together, are greater than the total pie there is to divide.

1Note that the requirement is not that both sides think they are more likely to win than lose. That is, the countries’ respective estimates of the probability of success need not be greater than 1/2, they need only be inconsistent with the fact that the actual probability of winning for the two sides sum to 1.
A similar argument is made by scholars who study why wars end. Given a war, it is argued that countries continue to fight until their individual assessments of the likely outcome of combat converge. At that point, both sides can agree to a settlement that both prefer to continued war. As Wittman says, inefficient wars start and continue because the probability that state 1 believes it may win need not equal the one minus the probability that state 2 wins, “. . . as the probability estimates [of each side] are based on different sources of information (Wittman 1979, p. 755).”

In this paper, we reconsider the mutual optimism hypothesis by presenting a formal model integrating players’ knowledge and strategy in a single framework and analyzing the equilibria of a class of games that capture the key features of the mutual optimism argument. We find that the simple logic of “war by mutual optimism” is misleading. That is, if war is a “public event” when it occurs, then it cannot occur because of mutual optimism. We show this result is robust by generalizing it to the case where leaders are not perfectly rational information processors.

Our goal here is significantly less ambitious than presenting an all-inclusive theory of war and peace. Rather, we aim to formalize the mutual optimism argument, show that the widely cited intuition is mistaken, and provide a clear reason why the relationship between mutual optimism and war is not a simple as it first appears. As such, our analysis provides a starting point for a theory of mutual optimism and war.

The rest of this paper proceeds as follows. In the next section we present a simple game that highlights the nature of our main results. Section 3 describes the class of games for which our results apply and gives our main result, that in such games mutual optimism cannot lead to war. Then in section 4 we extend the result to players who process information imperfectly. We end with a discussion of the implications of our results for theories of war and peace.

---

2For more recent work that builds on the ideas presented by Wittman, see Wagner (2000), Wittman (2001), Smith and Stam (2004), and Powell (2004).
We begin with an example that will provide some intuition for the general result that follows. Consider a game between two players, Alice and Bob, who have a choice between participating and not participating in a contest whose outcome is determined by chance. To start, each player is given a (fair) die to roll and the die are compared. If Alice’s die generates a higher number than Bob’s, Alice wins. If the number on Bob’s die is the larger he wins, and Bob and Alice tie otherwise. Suppose that each player maximizes expected utility and assigns a utility value of +1 to winning, -1 to losing, and 0 to a tie.

Before deciding whether to participate in the contest or not, each player receives some information about the roll of the dice. In particular, Alice observes the result of her die and Bob observes the result of his die. After receiving this information, the two players simultaneously announce whether or not they agree to play the game of chance. If they both agree, payoffs are awarded as above and, in addition, each player pays a small cost $c$. If one or both players do not agree to play, then both receive a payoff of 0.

To begin the analysis of this game, suppose a player naively considers only the private information he or she receives and does not make any additional inferences. That is, suppose Alice, for instance, observes the results of the first die and makes the assumption that the second die is equally likely to be any of the six possible values. Then it is easy to show via a simple expected utility calculation that if Alice observes 4, 5 or 6 on the first die, she would have positive expected utility for the contest, and if she observes 1, 2 or 3, her expected utility would be negative. Such non-strategic Alice would choose to play if she observed the first die is greater than or equal to 4 and would choose not to play for values less than 4. Clearly, the same would hold for Bob.

If we stop our analysis here, the game would be played when both players have seen a roll of their die that causes them to think they are likely to win. However, these decision rules neglect an important aspect of the game, and are therefore incomplete. Specifically, a strategic player must evaluate the effects of their opponent’s information on their opponent’s decision, and incorporate those inferences into their own calculations. Consider Alice. If the
contest is played in equilibrium, then it must be the case that it is “rational” for Bob to play. Therefore, in judging whether or not it is best response for Alice to play, she should consider the implications of Bob’s strategy when he participates in the contest. Alice can then infer, by the above analysis, that even an unsophisticated Bob will not play if he observes that the second die is 3 or less. So, given Bob’s strategy, Alice knows that the second die is at least a 4. But in that case, if Alice has observed that the first die is a 4, there is no chance of winning, because conditional on Bob agreeing to play, the best she can do is tie. Of course, by the same logic, Bob will not play if he observes the second die is a 4, because if Alice agrees to play, he can infer that the first die is greater than or equal to 4, and at best he will tie. Based on this analysis, then, Alice would only choose to play if the first die is a 5 or a 6 and Bob will also reach the same conclusion.

Now, taking this line of reasoning one step forward, consider the case in which Alice, having made the above inference about Bob, observes the first die is a 5. Now knowing that Bob will only play if the second die is a 5 or 6, agreeing to play has negative expected utility! Likewise, Bob would choose not to play if he sees a 5. We are therefore left with the decision rules in which Alice only agrees to play if she sees a 6 and Bob only agrees if he sees a 6. But, as Alice and Bob can infer these decision rules, they know that if they both agree to play the result is sure to be a tie. However, because of the cost $c$, both players strictly prefer not playing to a tie, and therefore neither player will agree to play, regardless of how favorable their private information is. Thus, there cannot be a Bayesian-Nash equilibrium in which the game is played because of differing beliefs that result from private information.

Given the intuition highlighted by this example, we can think about mutual optimism among countries. Suppose instead of Alice and Bob, we have two countries, East and West. Also, suppose that instead of looking at the roll of a die, the countries receive private information about the “quality” of their troops, the reliability of their allies, or some other relevant factor that directly affects their likelihood of winning a war. For East, a strong report would be equivalent to Alice seeing a high number on her die. Similarly, for West, a positive report about their troop quality would be like Bob seeing a high number on his die.
In such a world, if East gets a good report (a high roll for Alice) and so does the West (a high roll for Bob), each side’s expected utility from fighting might exceed that of the status quo, or some efficient settlement. Yet, like Alice and Bob, when East and West reflect upon what it means to fight an opponent who is going to resist, they must come to the same conclusion as Alice and Bob, namely that fighting cannot be a profitable alternative to peace.

The sequence of inferences described above reflect the fact that rational players in each example understand that it cannot be the case that both players expect to win in equilibrium.³

³This result is similar in spirit to ones found in the economics literature on efficient exchange (Aumann 1976, Milgrom and Stokey 1982, Tirole 1982, Sebenius and Geanakoplos 1983, Rubinstein and Wolinsky 1990).

³Niou, Ordershook and Rose (1989) and Fearon (1995) briefly discuss this result, but focus on Aumann’s assumption about the common knowledge of the players’ posterior beliefs. They fail to discuss the relationship between what is known in equilibrium and what beliefs

3 Mutual Optimism, Rationality, and War

A typical story for how war might result from mutual optimism is as follows. Suppose the leaders of two countries have information about their military forces and tactics that their opponent does not. Moreover, suppose that this information influences each leader’s assessment of their country’s likelihood of success in combat. If both leaders then believe that their side possesses the “stronger” force, both sides may think they will prevail militarily and thus both leaders may choose to fight rather than pursue a peaceful settlement. In such an environment, the leaders’ mutual sense of optimism could create a situation where there are no ex ante bargains both sides prefer to war, even though war is known to be ex post inefficient. The paradox is then that although rational leaders know that both sides cannot benefit from war at the same time, they still start wars that they would have preferred to avoid.⁴ In this section, we construct a game-theoretic model that formalizes the idea of

⁴This result is similar in spirit to ones found in the economics literature on efficient exchange (Aumann 1976, Milgrom and Stokey 1982, Tirole 1982, Sebenius and Geanakoplos 1983, Rubinstein and Wolinsky 1990).

⁴Niou, Ordershook and Rose (1989) and Fearon (1995) briefly discuss this result, but focus on Aumann’s assumption about the common knowledge of the players’ posterior beliefs. They fail to discuss the relationship between what is known in equilibrium and what beliefs
mutual optimism as a cause of war.

3.1 Model and Assumption

Suppose two countries are facing a potential conflict. The dispute can be settled by war or resolved without armed conflict. To represent war as a mutual decision we assume that war only occurs if both countries decide to stand firm. We also assume that war is an inefficient method of settling disputes. In this construction, we focus on the private information that the countries may have about their ability to prevail in the event of war and their expectations regarding the bargaining process. Initially, countries are uncertain about particular facets of the crisis situation, such as the balance of forces, technological differences, military strategy, latent resources of each side, support from allies, etc. The set of all such possible situations is denoted by $\Omega$. An element of $\Omega$ is therefore a complete description of one possible situation and thus we refer to $\omega \in \Omega$ as a (possible) state of the world. For simplicity, we assume $\Omega$ is a finite set. We may then ask, what do the two sides believe before they “know” anything about the true state of the world? If we assume that differences in peoples beliefs about the state of the world are the product of private information, such as their educational background, any inputs they may receive from advisors, life experience, etc., then it is logical to suppose all players share a “common prior.”\(^5\) Let $\pi = \Delta \Omega$ be this shared prior.

The “information” in our model can be naturally classified into two categories. First, there is information that players agree will effect the two sides’ respective likelihoods of success in the same way. For example, geography can favor one side, such as the advantage that being an island gave the British in the face of German attacks during World War II; it would indeed be strange to claim that the information pertaining to the “islandness” of Great Britain was private. Such commonly known information is described as “common knowledge.” The second type of information is private and only known to one side. This a player could hold. Patty and Weber (2001) discuss a similar problem in the context of the democratic peace.

\(^5\)For a paper that does not use a common prior, see Smith and Stam (2004).
information could describe such things as troop quality, military strategy, and the plans of other countries in the international system. For example, the capabilities and tactics of German units on the eve of World War II would be an obvious source of private information for Germany. This private information would then generate differences in the two sides’ respective assessments of the pre-war likelihoods of success.

To formalize this information structure we will need a model of knowledge. Generally, knowledge within a game is characterized by the ability of a player to distinguish between elemental states \( \omega \) in \( \Omega \), as a player may possibly distinguish between decision nodes in an extensive form game. We are also interested in *events*, which are naturally defined as subsets of \( \Omega \). For example, if \( \Omega = \{1, 2, 3, 4, 5\} \) then \( E = \{1, 3\} \) is an event. So here, events are related to information sets and describe a set of states consistent with the history of the game.

To formalize what players know and when they know it, we use a *possibility correspondence* \( P_i(\omega) \), that maps every state \( \omega \) to a non-empty set of states \( P_i(\omega) \subseteq \Omega \). 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 \( \omega \). Equivalently, a player’s knowledge can be formalized by a *knowledge correspondence* \( K_i(E) = \{ \omega : P_i(\omega) \subseteq E \} \), where \( K_i(E) \) is the set of states where player \( i \) knows an event \( E \) has occurred, for sure. So \( i \) knows \( E \) at every state in \( K_i(E) \).

As can be seen by the relationship between \( K_i \) and \( P_i \), a player’s knowledge can be discussed in terms of \( P_i \) or \( K_i \). For example, let \( \Omega = \{1, 2, 3, 4, 5\} \) and let player \( i \)’s knowledge be represented by \( P_i(\omega) \) taking on the following values:

\[
P_i(1) = P_i(3) = \{1, 3\} \quad \text{and} \quad P_i(2) = P_i(4) = P_i(5) = \{2, 4, 5\}.
\]

\( P_i \) implies that if the state is 1, the player thinks that the true state is either 1 or 3. Similarly, if the state is 4, then she thinks that the true state is 2, 4, or 5. Now define the events \( F = \{1, 2\} \) and \( F' = \{1, 2, 3\} \). By the definition of \( K_i \), player \( i \) “knows” an arbitrary

\*Note that our model of knowledge is simply a generalization of the typical information set structure of extensive form games of incomplete information.*
event $E$ at $\omega$ if $P_i(\omega) \subseteq E$. In this example, if $\omega = 1$ then $i$ clearly does not know $F$ because $P_i(1) \not\subseteq F$ ($P_i(1) = \{1, 3\}$). However, she can deduce the conditional probability of $F$, given $P_i$ and a prior $\pi_i$, using Bayes’ Rule. This conditional probability is her posterior belief that $F$ has occurred given her private information, $P_i(1)$. On the other hand, at $\omega = 1$ player $i$ does know $F'$ since $P_i(1) = \{1, 3\} \subseteq \{1, 2, 3\}$.

Notice that in the preceding example, the two sets $\{1, 3\}$ and $\{2, 4, 5\}$ form a partition of $\Omega$. It is typically the case that the structure of a player’s knowledge is represented by a collection of disjoint and exhaustive subsets of $\Omega$, called a partition.

**Definition 1.** A possibility correspondence $P_i(\omega)$ for $\Omega$ is partitional if there is a partition of $\Omega$ such that for any $\omega \in \Omega$ the set $P_i(\omega)$ is the element of the partition that contains $\omega$.

If $P_i$ is a partition and if $\omega$ and $\omega'$ are two states in $\Omega$, then when $\omega$ and $\omega'$ are in the same element of the partition, the decision-maker cannot tell the difference between them. However, if $\omega$ and $\omega'$ are not in the same element of the partition, the decision-maker can tell the two states apart.

At this point we place restrictions on $P_i$ to represent the types of properties, or axioms, we desire in the processing of information by players.

**Definition 2.** Let $P_i$ be a possibility correspondence for individual $i$. We say

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

2. a player $i$ knows that she knows [KTYK] if, for every $\omega' \in P_i(\omega)$, $P_i(\omega') \subseteq P_i(\omega)$,

3. a player $i$ knows that she doesn’t know if, for every $\omega \in \Omega$ and every $\omega' \in P_i(\omega)$, $P_i(\omega') \supseteq P_i(\omega)$.

These three properties are used to formalize the idea of rationality in knowledge. The first condition requires that a rational person always considers the true state of the world to be possible. The second condition requires that if any state an individual thinks is possible at the current state of the world were the true state, she would know at least her current
knowledge, $P_i(\omega)$. That is, $P_i(\omega)$ cannot occur without the individual knowing that she knows it has occurred. Formally, this implies that if $K_i(E)$ is the event that $i$ knows $E$, then $K_i(K_i(E)) \subseteq K_i(E)$. The final condition requires that players also know what they don’t know, i.e., $\neg K_i(E) \subseteq K_i(\neg K_i(E))$. These three conditions ensure that $P_i(\omega)$ is consistent and that a player’s possibility correspondence represents all that is knowable at each state.

To see an example of how rationality is related to the conditions of definition 2, let us consider a possibility correspondence that does not satisfy all three conditions. It is then easy to see why such a correspondence does not represent all that a rational player could know at a given state of the world, $\omega$. Suppose that there are five states of the world, $\Omega = \{1, 2, 3, 4, 5\}$. Since a rational model should never have players place zero probability on the true state of the world, let us consider a correspondence $P_i(\omega)$ that satisfies nondeluded, but not know that you know and know that you don’t know. In particular, let

$$P_i(\omega) = \{\omega - 1, \omega, \omega + 1\} \text{ if } \omega = \{2, 3, 4\},$$

$$P_i(\omega) = \{1, 2\} \text{ if } \omega = 1,$$

and

$$P_i(\omega) = \{4, 5\} \text{ if } \omega = 5.$$

Now suppose that $\omega = 2$, what could a rational player conclude? From $P_i(2)$ the player knows that the true state of the world is 1, 2, or 3. But player $i$ also knows that the state is not 1, because if the state were 1 she would know that the state was not 3. So $i$ can deduce that the true state is either 2 or 3. Moreover, if $\omega = 3$, $i$ would know that the other possible states would be 2, 3, 4, but since she knows $\omega \neq 4$, $\omega$ cannot be 3. Therefore, $i$ can deduce that the true state is 2, and a rational player knows more than what is described by $P_i(\omega)$.

In fact, we can justify the use of partitional possibility correspondences in a model with rational actors because it is easy, if somewhat tedious, to show that $P_i$ satisfies nondeluded, know what you know, and know what you don’t know if and only if it is partitional. That is, a partitional possibility correspondence is the only internally consistent representation of
a player’s knowledge and, therefore, represents all that can be known by a rational decision-maker at a given state of the world.\footnote{Notice that information sets in extensive form games are always partitional.}

Now to use this model of knowledge in equilibrium analysis, we also need to be able to say something about what players know about what others know, and what those others know that they know, etc. This is accomplished by considering events that are common knowledge. The concept of common knowledge was first explicitly described by Lewis (1969), and was later formalized by Aumann (1976) in terms of the *meet* of the partitions of players at a state \( \omega \). Informally, an event \( F \) is *common knowledge* between players \( i \) and \( j \), in state \( \omega \), if and only if \( \omega \) is a member of every set in an infinite series such that, player \( i \) knows \( F \), player \( j \) know \( F \), player \( i \) knows player \( j \) knows \( F \), player \( j \) knows player \( i \) knows player \( j \) knows \( F \), \ldots

Like common knowledge events, another important class of events is the *self-evident* event.

**Definition 3.** An event \( E \) is self-evident 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. Returning to our previous example where \( \Omega = \{1, 2, 3, 4, 5\} \) and \( P_i(\omega) \) induces a partition on \( \Omega \) of \( \{\{1, 3\}, \{2, 4, 5\}\} \), the event \( \{1, 3\} \) is self-evident, but \( \{1, 2, 3\} \) is not. The following useful fact is immediate. If \( P_i \) is nondeluded, then for a self-evident event \( E \),

\[
E = \bigcup_{\omega \in E} P_i(\omega).
\]

Self-evident events are useful because of the following result, which states that an event being self-evident to all players is equivalent to it being common knowledge.

**Lemma 1.** Suppose \( P_i \) is nondeluded for all \( i \). An event \( F \) is common knowledge at a state \( \omega \) if and only if there is an \( \omega \) and a self-evident event \( E \) such that \( \omega \in E \subseteq F \) for all \( P_i \).
We are often also interested in “public” events. A public event, unlike a private signal, is known to all players when it happens. Formally, we define a public event as follows:

**Definition 4.** An event $E$ is a public event if and only if, for all $i$, $K_i(E) = E$.

Note that a public event is self-evident to all players; this equivalence is given in the following lemma.

**Lemma 2.** Suppose $E$ is a public event, then for all $i$, $E$ is self-evident.

*Proof.* By definition of a public event $K_i(E) = \{\omega \in \Omega | P_i(\omega) \subseteq E\} = E$ for all $i$. Therefore, for all $\omega \in E$, $P_i(\omega) \subseteq K_i(E) = E$ and $E$ is self-evident. \qed

Now that we have specified a model of knowledge, we can talk about knowledge at a given state of the world and the decision to go to war.

Recalling that the state of the world is directly relevant to the question of which country will win a 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 $\omega$. 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 $\pi$ via Bayes’s Rule to form a posterior belief about the value of $p_i$ as follows:

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

(1)

From this expression, it is easy to verify that if $E[p_i|E'] \geq x$ and $E[p_i|E''] \geq x$ for disjoint sets of states $E'$ and $E''$, then $E[p_i|E' \cup E''] \geq x$.

It equally likely that the negotiated settlement will depend on the underlying state of the world. 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 $\omega$. 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.

---

8For a discussion of public events see Milgrom (1981).
We represent the private information of country \( i \) by a possibility correspondence \( P_i : \Omega \rightarrow 2^\Omega \), 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 \( \omega \). Given a true state \( \omega \), a country can combine its knowledge of \( P_i(\omega) \) with the prior \( \pi \) and equation 1 to construct its posterior belief about the probability it will win, \( \hat{p}_i(\omega) = \text{E} \left[ p_i | P_i(\omega) \right] \), and its expectations from bargaining, \( \hat{r}_i(\omega) = \text{E} \left[ r_i | P_i(\omega) \right] \). It is important to note that without additional assumptions or structure, it is certainly possible that \( \hat{p}_1(\omega) \neq 1 - \hat{p}_2(\omega) \) or \( \hat{r}_1(\omega) \neq 1 - \hat{r}_2(\omega) \) for some state \( (\omega) \). In this setting, mutual optimism occurs when \( \hat{p}_1(\omega) + \hat{p}_2(\omega) > \hat{r}_1(\omega) + \hat{r}_2(\omega) \).

In the rest of this section, we place the preceding informational assumptions in the context of a game and show that, in equilibrium, countries cannot have mutual optimism and therefore, war cannot occur as a result. Because the information structure is quite general, and can capture many aspects of the interaction, the decision problem can be represented by a simple finite two player game in strategic form. Denote the set of actions for country \( i \) by \( A_i \), with elements \( \{a^1_i, \ldots, a^k_i\} \). Depending on the choice of actions by both countries, the outcome of the game is either war or settlement. The expected payoff to war depends on the probability that a country will win, the utility of victory and defeat, and the inefficiencies present in fighting. We normalize the utility of countries to be 1 for victory in war and 0 for defeat, and we suppose there is a cost \( c > 0 \) for fighting a war. Thus the expected utility for country \( i \) of going to war is simply \( \hat{p}_i(\omega) - c \). On the other hand, the negotiation process provides an expected utility \( \hat{r}_1(\omega) \) for country 1 and \( \hat{r}_2(\omega) \) for country 2. In fact, the revelation principle tells us that any equilibrium of any bargaining game can be mimicked by a mapping that takes states of the world into outcomes (Myerson 1979). This is because a game is simply defined to be a mapping from players’ action profiles into outcomes, a strategy is a mapping of player types into actions, and so without loss of generality we can define a new mapping, \( r(\omega) \), that is a composition of the extensive form and the strategy. This new mapping takes states of the world into outcomes in the exact same way as an equilibrium to the game.

To focus on the mutual optimism explanation of war, we assume that both countries
must choose to “stand firm” for war to occur. Formally, this condition requires that when one country chooses stand firm the other country can stop a war by inducing the bargaining subgame instead. That is, war is an act of mutual consent and, by construction, also mutual optimism. Alternatively, we could assume that war is not a mutual act. Such models of unilateral war have received considerable attention in the literature and the results are well known (Filson and Werner 2002, Powell 2004, Slantchev 2003). These models generally show countries may face incentives to fight in hopes of a better settlement or that countries can use costly war to screen out unresolved types of challengers. But, if we assume that any single state can cause a war, the concept of war by mutual optimism loses meaning. In particular, it is hard to understand what is mutual about mutual optimism if only one side’s expectations enter into the decision to fight.

We now define strategies for each country. We reflect the fact that countries can condition their choice of action on their private information by defining a strategy \(s_i \in S_i\) by a function \(s_i : \Omega \rightarrow A_i\) with the restriction that

\[
P_i(\omega) = P_i(\omega') \Rightarrow s_i(\omega) = s_i(\omega').
\]

This condition states that if a country cannot distinguish state \(\omega\) from state \(\omega'\), then its action must be the same in both states.

Lastly, we discuss the event war. It follows from Lemma 2 that if the outcome of the game is a publicly observable war, then war is common knowledge whenever it occurs. Since strategies associate states with outcomes, we can now offer a rigorous definition of the statement that war is a public event. For a strategy profile \((s_1, s_2)\), let \(F\) be the set of states for which the outcome of the game is war. It follows that if war is publicly observable (for \((s_1, s_2)\)), then the event \(F\) is a public event. If the event \(F\) is nonempty for a strategy profile \((s_1, s_2)\), we say that \((s_1, s_2)\) is a strategy profile in which war occurs.
3.2 Results

Letting $G$ denote any strategic form game of incomplete information that satisfies the preceding assumptions on information structure, payoffs, and strategies. We are now prepared to state our main result.

**Theorem 1.** Suppose countries have a common prior, war is a public event, and $P_i$ is partitional for $i = 1, 2$. Then there is no Bayesian-Nash equilibrium of $G$ in which war occurs.

**Proof.** Suppose not. That is, suppose that the strategy profile $(s_1^*, s_2^*)$ is a Bayesian-Nash equilibrium in which war occurs. At state $\omega$, war has an expected payoff to country $i$ of $\hat{p}_i(\omega) - c$. By choosing action $\tilde{a}_i$, though, country $i$ can ensure itself a payoff of $\hat{r}_i(\omega)$. Define the following two events:

$$W_1 = \{\omega \in \Omega \mid \hat{p}_1(\omega) \geq \hat{r}_1(\omega) + c\}$$

$$W_2 = \{\omega \in \Omega \mid \hat{p}_2(\omega) \geq \hat{r}_2(\omega) + c\}.$$  

Here $W_i$ is the event that country $i$ will prefer war given that it knows the event $P_i(\omega)$ has occurred. At states outside $W_i$, country $i$ will prefer to deviate to $\tilde{a}_i$. Thus, if $(s_1^*, s_2^*)$ is a Bayesian-Nash equilibrium, the set of states for which the outcome of the game is war is $W = W_1 \cap W_2$. Moreover, since war occurs under this strategy profile, $W$ is not empty; there exists an $\omega^* \in W$.

Since we assume war is a public event, the event $W$ is a public event. By Lemma 2, $W$ is self-evident to 1 and 2, and therefore

$$W = \bigcup_{\forall \omega \in W} P_1(\omega) = \bigcup_{\forall \omega \in W} P_2(\omega). \quad (2)$$

That is, as $W$ is self-evident for nondeluded $P_i$, $W$ is the union of $P_i(\omega)$ for all $\omega$ in $W$, and this is true for each player $i$. As the correspondence $P_i$ is partitional, we can further write $W$ as the union of disjoint sets $P_1(\omega)$, defined by some collection of states $D^*$ with $D^* \subseteq W$. Since $D^* \subseteq W$, we have $E[p_1|P_1(\omega)] \geq E[r_1|P_1(\omega)] + c$ for every $\omega \in D^*$. That is,
if each disjoint set $P_1(\omega)$ has conditional expectation $E[p_1|P_1(\omega)]$ of at least $E[r_1|P_1(\omega)] + c$, then the conditional expectation over the union of these disjoint sets (i.e., $E[p_1|W]$) is also at least $E[r_1|W] + c$. By a symmetric argument for player 2, $E[p_2|W] \geq E[r_2|W] + c$. Therefore $E[p_1|W] + E[p_2|W] \geq E[r_1|W] + E[r_2|W] + 2c$. But as $p_1(\omega) + p_2(\omega) = 1$ and $r_1(\omega) + r_2(\omega) = 1$ for all $\omega \in \Omega$, it follows from Bayes’s Rule that $E[p_1|W] + E[p_2|W] = 1$, $E[r_1|W] + E[r_2|W] = 1$, and $1 \geq 1 + 2c$. This is a contradiction, which proves the result.

This theorem shows that there cannot be an equilibrium in which both sides think they are better off fighting, and as a result, go to war. The intuition underlying this theorem is as follows. Each country knows that the other is optimizing in equilibrium, knows when a war occurs and can deduce from the set of states, the prior probability, and the associated costs and benefits of each action for their opponent. Each player therefore knows that only in states where she is “likely to lose,” is the opponent willing to fight. Knowing this, each player should condition her decision on this fact. As a result, the conjecture that the players’ strategies form an equilibrium where war is a public event would unravel just like in the dice game between Alice and Bob discussed above. So the public knowledge that a war has occurred is inconsistent with inconsistent beliefs, even if leaders have private information about the likely outcome of conflict.

We conclude this section with some additional remarks regarding the theorem. First, our result does not require that bargaining be costless, but rather that fighting a war is more costly. This is because in any world where bargaining is costly, but less costly than war, we can always normalize the settlement outcome and think of the cost of war as the relative loss from fighting. Second, we note that Wittman’s (1979) classic mutual optimism argument is subsumed by a special case of Theorem 1. In particular, when $r_1(\omega) = r$ and $r_2(\omega) = 1 - r$ for all $\omega$, we see that for any $r \in [0, 1]$ there cannot be war by mutual optimism. That is, it would be false to conclude that mutual optimism can create a situation where countries

\footnote{Note the $(r_1(\omega), r_2(\omega))$ negotiation subgame need not be reached in any equilibrium of the game for the result to hold.}
can find no agreement that would dissuade at least one country from wanting to fight in equilibrium. In fact, Theorem 1 implies any efficient division will do.

Our result, however, is significantly stronger than this. Theorem 1 shows that there are no “optimistic” beliefs that allow a public war to be preferred to a peaceful settlement in equilibrium even when the probability of success in war and the value of a negotiated settlement are arbitrary functions that may depend on the state of the world. Moreover, as \( r_1(\omega) \) \( r_2(\omega) \) can represent any equilibrium of any bargaining game one could imagine, our results do not depend on the bargaining procedure that leads to a settlement. Third, one may wonder how our assumption that the game is in strategic form influences the generality of our theorem. First note that when it comes to the possibility correspondence we require only that they be partitional and that war is a public event. As a result player’s information may differ in a number of ways. One way they may differ is that one leader may know whether the other has chosen to stand firm or negotiate when making their decision. That is, our result applies equally to decisions made simultaneously and sequentially. Finally, our assumption that war is a public event is used in the proof to show that the event \( W \) is a public event. For \( \omega^* \in W \), by Lemma 1, this event is common knowledge at \( \omega^* \). So another way to state the conclusion of the theorem is that if war is common knowledge when it occurs, it cannot occur because of mutual optimism.

4 Mutual Optimism, War, and Bounded Rationality

While Theorem 1 is true for any strategically equivalent game in which the players rationally process information, one may wonder if the results depend on strictly rational learning. Can mutual optimism result in war if otherwise rational agents suffer from pathological misperception? 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, war still may not

\(^{10}\)For a more detailed discussion of this point see section 6.
be possible in an equilibrium of our game. In particular, even if both players ignore “bad news” or are inattentive, then war cannot occur because of mutual optimism.

4.1 Processing errors and bounded rationality

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. For example, consider a world where there are two possible states, \( \{a, b\} \). Suppose that when the true state is \( a \) it is brought to the player’s attention that the state is in fact \( a \). However, when the true state is \( b \) nothing is brought to the player’s attention. In this situation a rational Bayesian can always deduce the true state of the world. When the state is \( a \), the player is informed of that fact and knows it is \( a \). When the state is \( b \), the player knows that if the state were \( a \) she would have been told, but since she was not, the state must be \( b \). The rational Bayesian, like Sherlock Holmes, 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 and 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. So while they may learn that the state is \( a \) when there is an explicit signal to indicate that is so, they may not deduce that the state must be \( b \) in the absence of a signal that the state is \( a \). 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 and 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.

Now we consider a game with players whose information processing is flawed in ways consistent with the learning processes described above. Note that a common theme to 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 $P_i$ satisfies is nondeluded, and KTYK, we now allow players also to be blissfully ignorant.

**Definition 5.** A player $i$ is blissfully ignorant if, for all $\omega$ and $\omega' \in \Omega$, either $P_i(\omega') \cap P_i(\omega) = \phi$, or $P_i(\omega') = P_i(\omega)$, or else for some $\hat{\omega} \in \{\omega', \omega\}$ $P_i(\hat{\omega}) = \Omega$.

Blissful ignorance is therefore a generalization of rational learning. That is, such a decision-maker may processes information in a rational way or she may ignore new information. 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 information is costly, because they are inattentive, or because they would rather not think about the implications of the information in front of them.$^{11}$

### 4.2 Results

Together, these conditions are weaker than the three conditions for a rational partition, yet are still sufficient to exclude optimistic war. In general, we now have:

**Theorem 2.** Suppose countries have a common prior, war is a public event, and $P_i$ is nondeluded, KTYK, and blissfully ignorant for $i = 1, 2$. Then there is no Bayesian-Nash equilibrium of $G$ in which war occurs.

**Proof.** The proof of this theorem is similar to the proof of Theorem 1. Suppose that the strategy profile $(s_1^*, s_2^*)$ is a Bayesian-Nash equilibrium in which war occurs. Define the

---

$^{11}$For other approaches to bounded rationality in models of knowledge see Geanakoplos (1989).
following two events:

\[ W_1 = \{ \omega \in \Omega | \bar{p}_1(\omega) \geq \hat{r}_1(\omega) + c \} \]

\[ W_2 = \{ \omega \in \Omega | \bar{p}_2(\omega) \geq \hat{r}_2(\omega) + c \} \]

As before, if \((s_1^*, s_2^*)\) is a Bayesian-Nash equilibrium, the set of states for which the outcome of the game is war is \(W = W_1 \cap W_2\), with some \(\omega^* \in W\).

Also as before, war is a public event and so the event \(W\) is a public event. By Lemma 2, \(W\) is self-evident to 1 and 2, and therefore

\[ W = \bigcup_{\forall \omega \in W} P_1(\omega) = \bigcup_{\forall \omega \in W} P_2(\omega) \quad (3) \]

due to \(P_i\) being non-deluded.

Since each \(P_i\) satisfies non-delusion, KTYK, and blissful ignorance, there are only two kinds of events that can be self-evident between 1 and 2. On the one hand, \(W\) could consist of a set of states that are the union of partitional subsets of \(\Omega\) for each \(P_i\). In this case, the arguments in the proof of Theorem 1 apply directly and a contradiction results.

Alternatively, there could exist \(\hat{\omega} \in W\) such that \(P_i(\hat{\omega}) = \Omega\). It follows then that \(W = \Omega\) and from the definition of \(W\), \(E[p_1|\Omega] \geq E[r_1|\Omega] + c\) and \(E[p_2|\Omega] \geq E[r_2|\Omega] + c\). Once again, this yields a contradiction, as in the proof of Theorem 1.

So, even for less than rational actors, mutual optimism cannot be the reason two decision-makers go to war.

5 Discussion

Formalizing the mutual optimism hypothesis we see that if war is a public event when it happens, it cannot occur because of mutual optimism. Moreover, when we consider some intuitive forms of “bounded rationality” in the way players learn, the result is robust. This result is somewhat surprising, given the existing work on the subject. In this section we
discuss some specific aspects of our model as well as its implications and limitations for the study of international conflict.

Why doesn’t mutual optimism lead to war in our setting, given that others have argued that it can? One reason is that previous arguments linking mutual optimism to war are based on partial equilibrium analysis and did not fully account for the strategic nature of the interaction between decision-makers (Wagner 2000, Wittman 1979, Wittman 2001). For example, consider the canonical model in which each leader’s private information can induce both sides of a conflict to be optimistic about their prospects in a war, even though everyone knows war is *ex post* inefficient. From the observation that private information can generate such inconsistent beliefs, it is argued that if there is sufficient mutual optimism then leaders can choose to fight wars that, *ex post*, they would have preferred to avoid. Our results show that this conclusion is incorrect when war is common knowledge when, when it is known that the other leader is rational and when each leader has the ability to pursue negotiations instead. Specifically, Theorem 1 demonstrates that if war is a public event, then in equilibrium each player knows that the true state of the world must be one that could lead their opponent to fight. Yet in exactly those states of the world the at least one leader’s preference is for a settlement. Therefore, in any given state of the world, at least one player can reason that if the other is willing to fight, and this is equilibrium behavior, they can deviate to a strategy that induces negotiations and makes themselves better off.

This summary of the logic of our argument also highlights the importance of our assumption that both countries must commit to stand firm in order for war to occur. Indeed this assumption is required in order to establish our result. The fact that either side can choose negotiations rather than war makes it possible for one country to infer, in the case of war, the nature of the other country’s private information, namely that fighting would be better for that country than settling. Moreover, as previously discussed, we are in fact forced to make this assumption in order to formalize the explanation of war as a result of

---

12 Fey and Ramsay (2004) show that in certain cases, if a single country can start a war, it is impossible to write down any game form that guarantees peaceful outcomes.
mutual optimism - without this assumption we are left with an explanation of war as a result of unilateral optimism.

While our results are more general than many in the study war, we may still be interested in exactly which extensive forms satisfy the conditions presented in our two theorems. One way to think about the class of games covered by our results is to notice that in any given extensive form, we can represent a player’s information set as a partition function. The “possible nodes” are members of $P_1(\omega)$ and the other nodes are not. Noticing that information sets are always going to be partitional at any given stage of the game, the information structure we describe allows for any strategic history you might consider, as long as the theorem’s conditions are met. In particular, the differences in $P_1(\omega)$ and $P_2(\omega)$ allow for the possibility that one leaders knows that the other has already chosen stand firm or negotiate. Therefore, a mutual optimism game can have sequential or simultaneous moves.

On the other hand, we do require that the actions available to the players at these information sets must give them at least a choice between standing firm and negotiating, but there may be any number of other choices. It must also be the case that, conditional on one country standing firm, the other can offer negotiations and play a bargaining subgame that is described by $r_1(\omega)$, $r_2(\omega)$. The extensive form of the bargaining game subgame is inconsequential. Any game of this type cannot produce war by mutual optimism.

Note, however, that the strategic problem here is different from the risk-reward tradeoff characterized by most sequential bargaining games. One way to think about the relationship between these different models is that if it is common knowledge that countries are going to fight, and these countries have a “red phone,” then at least one side will make a call and a proposal that will be accepted and avoid the war. That is, our result applies to a situation where countries can discuss war before making it, but after a proposal has been made and rejected. Our result does not apply to situations where a firm offer is made that, if rejected, leads to certain war.

At this point, it should be emphasized that our model demonstrates that, in a fairly general setting, a game-theoretic formalization of the mutual optimism hypothesis does not
permit war as an equilibrium outcome. Our results do not, however, have anything to say about other explanations of war. For example, if the settlement option described in our setup suffers from a commitment problem, then war would be possible (Fearon 1998). Also, two key assumptions of the model are that the information partitions are themselves common knowledge, and that the players share a common prior over the states of the world. The first assumption is easily justified by the fact that war is clearly a public event, and thus its occurrence is commonly known. While we have noted above why we believe the common priors assumption is reasonable, and have discussed the assumption in more detail elsewhere (Fey and Ramsay 2005), this assumption is not without its critics (Gul 1998). When we assume common priors, we are assuming that the beliefs of rational players in a game are generated by updating, via Bayes’s rule, a single prior over a commonly known set of states. Obviously, such a assumption puts limits on what a rational player can believe. However, if we conceptualize players as having different theories as to how the world works resulting from private information, i.e., the way they were taught in graduate school, their cultural and social background, and their “life experiences,” it is relatively easy to show that even with common priors the same event can lead rational agents to draw different conclusions. In fact, the only time when players’ beliefs must agree is in the case where their beliefs are about a common knowledge event, when actions are common knowledge, or when the players posteriors are common knowledge (Geanakoplos 1992).

In any case, the common priors assumption is critical in the proof of our results; clearly war can occur due to mutual optimism if countries have different priors. Then again, with non-common priors it is easy to show that anything can be supported as equilibrium behavior, regardless of what we choose to model.

In sum, the results of our model are general, but not universal. Our results are not universal in the sense that we only address the use of mutual optimism as an explanation of war and have little to say about other mechanisms by which war occurs. However, our results are general in that the assumptions underlying the model are not heroic; the construction is consistent with, and even subsumes, arguments found in the informal and decision theoretic
literature on mutual optimism and war; and our results hold for a number of different extensive form games and less than rational players.

6 Conclusion

We have shown two things by formalizing the argument that countries fight wars due to the mutual optimism. First, if war is a public event among rational actors, war cannot occur because of mutual optimism. Second, by relaxing conditions in our model of knowledge the result is shown to be robust to imperfect, or “boundedly rational,” learning. Both of these claims are based on our demonstration that, for a general class of games where states contemplate war, there are no Baysian-Nash equilibria with war as an outcome.

These results are by no means the final word on the subject. The structure of the model points to a number of fruitful avenues for additional investigation. Within the framework of the current model, one could further weaken the rationality conditions of the knowledge model and characterize just how irrational decision-makers need be for mutual optimism to lead to war. Similarly, one could investigate the weakest possible conditions on decision-makers’ priors that still preclude war by mutual optimism, as Morris (1994) does with respect to inefficient trade.

Alternatively, one could go beyond the common value framework of this model and study the effects of incomplete information when players have unknown private values for war. If a tight characterization of the efficiency conditions is possible within such a framework, that model would speak to the constraints that rationality places on the effectiveness of treaties, settlement norms, and institutional design.


Wittman, Donald. 2001. “War or Peace.” Mimeo University of California, Santa Cruz.
Appendix

Proof of Lemma 1:

Proof. The proof of this lemma follows Rubinstein (1998) Proposition 3.5. Suppose $P_i$ is nondeluded for all $i$. Let $K_i$ be the knowledge function induced by $P_i$ and $E$, and recall that $K_i(E) = \{ \omega : P_i(\omega) \subseteq E \}$. Suppose there is a self-evident event $E \subseteq F$ for all $i$.

By definition of $E$ and nondeluded, $E \subseteq K_i(E)$, $K_i(E) \supseteq E$ and, therefore, $K_i(E) = E$ for all $i$. This implies $K_iK_j\ldots K_i(E) = K_iK_j\ldots K_j(E) = E$. Of course, $E \subseteq F$ implies $K_i(E) \subseteq K_i(F)$ for all $i$ and from this it follows that $E = K_iK_j\ldots K_i(E) \subseteq K_iK_j\ldots K_i(F)$ and $K_iK_j\ldots K_j(E) \subseteq K_iK_j\ldots K_j(F) = E$. Since $\omega \in E$, $\omega$ is a member of all the above sets and $F$ is common knowledge at $\omega$.

For necessity, suppose $F$ is common knowledge at $\omega$ and take $E$ to be the intersection of all the series of sets $K_iK_j\ldots K_i(F)$ and $K_iK_j\ldots K_j(F)$. Because $F$ is common knowledge at $\omega$, $\omega \in F$. By the nondeluded property, $E \subseteq F$, and since $\omega$ belongs to all the sets $K_iK_j\ldots K_i(F)$, $P_i(\omega) \subseteq E$ for all $\omega \in E$. \qed

27