Mutual Optimism as a Rationalist Explanation for War

Slantchev, Branislav L. and Tarar, Ahmer S.

University of California San Diego, Texas AM University

4 July 2010

MUTUAL OPTIMISM AS A RATIONALIST EXPLANATION OF WAR

Branislav L. Slantchev*
Department of Political Science, University of California – San Diego

Ahmer Tarar
Department of Political Science, Texas A&M University

June 4, 2010

Abstract

Blainey (1988) argued that crises are more likely to end in war when two nations disagree about their relative power. Fey and Ramsay (2007) claim that this widely used "mutual optimism" explanation is theoretically incoherent. Their criticism neglects the need to specify a behavioral causal mechanism that links beliefs to the outbreak of war. We show how the rationalist game-theoretic work on the causes of war provides such mechanisms: the risk-return trade-off and costly signaling, and demonstrate that these models are immune to Fey and Ramsay's critiques. We also show that the class of models Fey and Ramsay propose make the substantively unwarranted assumption that an actor can unilaterally impose peace on an opponent who strictly prefers war. Their finding that war does not occur in equilibrium has nothing to do with mutual optimism. We conclude that the mutual optimism explanation can be grounded on firm rationalist foundations.

*We have benefitted from illuminating discussions with Bob Powell, Kris Ramsay, and Mark Fey. We thank Christina Schneider and Megumi Naoi for helpful comments on previous drafts. This research was supported by the National Science Foundation (grants SES-0518222 and SES-0850435 to Slantchev and grant SES-0518945 to Tarar).

Blainey (1988, 114) offers what has become one of the most celebrated explanations of why war occurs: "war is usually the outcome of a diplomatic crisis which cannot be solved because both sides have conflicting estimates of their bargaining power." In this view, war happens when at least two states embroiled in a dispute cannot construct a mutually acceptable deal that would satisfy their conflicting demands, and when one of them—often in exasperation—resorts to arms in an attempt to impose its will by force. The crucial part of the explanation requires one to specify the reason, or reasons, for that inability to agree on a peaceful deal.

Blainey's argument is that the fundamental cause is the two sides' conflicting estimates about how much they can secure by force. Because a state can always go to war to enforce its demands, what it expects to gain by fighting constitutes the minimum it would require in any peaceful negotiation. These expectations are based on subjective estimates of one's probability of winning the war (which depends, among other things, on the distribution of power, and the quality and morale of one's armed forces), the costs required to do so, the duration of fighting, the behavior of allies and antagonistic states alike, the ability of one's economy and population to sustain the war effort, and so on. If both sides expect to gain a lot by fighting—perhaps because both expect to win with near certainty at an acceptably low cost—then there may exist no negotiated deal that can satisfy their minimum demands. War becomes the inevitable outcome.

This argument is now generally known as the *mutual optimism* explanation of war, and is among the most widely accepted explanations of why war occurs.[1] In a recent article in the *American Journal of Political Science*, Fey and Ramsay (2007, henceforth Fey & Ramsay) argue that this explanation is mistaken on logical grounds, and that war cannot occur between two actors

[1] Wittman (1979); Lebow (1981, 242–43); Levy (1983, 82–86); Blainey (1988); Van Evera (1999, 16); Johnson (2004); Stoessinger (2005, 211).

because of mutual optimism about the likelihood of victory. Furthermore, they argue that war occurs in many formal models under incomplete information not because of mutual optimism but because actors are 'locked in' by the extensive form of the game (i.e., war happens by assumption because analysts have not given actors enough flexibility to avoid it).

Motivated by Fey & Ramsay's critique of the mutual optimism argument, we have three goals in this article. The first is to provide a theoretical account of the mutual optimism mechanism that is more comprehensive and cohesive than what we have right now. One of the problems with Fey & Ramsay's analysis is that it is based on the informal version of the mutual optimism argument, which largely ignores the actual causal mechanisms through which mutual optimism might cause war. We show that the game-theoretic work on rationalist causes of war provides coherent accounts of that mechanism. Although these analyses generally call it "war due to incomplete information" rather than "war due to mutual optimism," they can naturally capture the mutual optimism explanation. Thus, our main contribution in this article is to point out rationalist mechanisms that implement the causal behavioral link between mutually optimistic beliefs and war. In our view, the game-theoretic work has advanced considerably beyond the underspecified informal treatments that Fey & Ramsay rely on but that are silent as to how precisely mutual optimism actually causes fighting to begin.

Fey & Ramsay make a number of arguments for why standard models of crisis bargaining cannot capture the mutual optimism idea, which is why they introduce a new class of models that they claim to be better suited for the purpose. They then show that war cannot occur in this class of models, which in turn leads them to conclude that the mutual optimism argument is incoherent. Our second goal in this article is to demonstrate that, contrary to their arguments, the standard models are immune to all of the ills that allegedly make them unsuitable for examination of the mutual optimism explanation. This obviates the need for Fey & Ramsay's class of models and also helps us elucidate the causal mechanisms that can implement the explanation in a rationalist framework.

Finally, we show that the class of models introduced by Fey & Ramsay cannot be used to study crisis bargaining. The fundamental feature of models in this class is that an actor can unilaterally impose peace terms on an opponent who strictly prefers war to those terms. We demonstrate how Fey & Ramsay's finding that war cannot occur in their class of models arises precisely from this substantively unappealing and distorting structural feature, and as such has nothing to do with mutual optimism. Thus, their results cannot be used to evaluate the coherence of the mutual optimism explanation even if one chooses not to adopt our rationalist specification of its mechanism.

We conclude that the standard game-theoretic models of crisis bargaining provide a behavioral causal connection between mutual optimism and war, and as such offer several coherent mechanisms that can implement the original insight. We do not claim that these rationalist mechanisms are the only ways in which optimistic beliefs might cause war but they are certainly sufficient to prove that the mutual optimism explanation can be grounded on firm theoretical foundations.

1 The Mutual Optimism Mechanism: A Synthesis

How does mutual optimism (MO) lead to war? Perhaps surprisingly, we do not have a clear and widely accepted definition of the MO mechanism even though references to it abound (see footnote 1). To begin with, one must distinguish between its rationalist and non-rationalist specifications. Most of the non-formal literature going back to Blainey (1988) is actually in the non-rationalist vein, as Fearon (1995, 391–93) points out, and it is essentially silent about the process through which mutual optimism leads to war; all it says is that optimism results in inflated expectations about fighting relative to peace, which makes actors unwilling to settle on terms their opponents are willing to offer. These inflated expectations might be due to psychological biases, nationalist fervor, bounded rationality, evolutionary adaptation, or private information about capabilities or resolve, among others. Whatever the source of optimism, the mutual unwillingness to offer better terms somehow causes negotiations to break down in war.

Observe that mutual optimism is only a starting point: it shows why the actors might be unwilling to offer each other war-avoiding peace terms. It does not tell us *how* this leads to war. This might not be a problem for a non-rationalist account: if an actor believes himself militarily superior to his opponent and thinks she is incapable of seeing the evident truth of that estimate, nothing she does would alter his conviction. He will even be quite willing to fight in order to teach her a lesson (presumably the outcome of the war would cause a revision of beliefs). If she is likewise optimistic, his willingness to fight will similarly signify nothing: she will be just as eager to fight to teach him a lesson.

Whatever the merits of such an account, one cannot maintain a fully rationalist explanation based on MO without specifying precisely *how* MO causes war in a world in which actors behave rationally and this fact is common knowledge (Fearon, 1995). The formal literature that has emerged over the past two decades has provided us with coherent accounts whose fundamental insight is that *MO causes actors to engage in behavior that ends in war even in environments where settlements exist that would make both better off and where they would be able to locate such settlements if they had better information.* It is not that MO simply makes actors *prefer* war to peace. Instead, when optimism makes actors unwilling to agree to the terms their optimistic opponents are willing to offer, it is the attempt to *overcome* this problem that sometimes results in war. In other words, MO leads to war not as a direct consequence of preferences but as a result of strategic behavior by the optimistic actors.

What are these war-causing behaviors? Costly signaling (actions that deliberately risk war to reveal the credibility of one's war expectations) and the risk-return trade-off (action that runs a higher risk of war in order to obtain better terms of peace) are two specific forms that the mechanism can take. We now present a very simple model of crisis bargaining to illustrate precisely how MO can lead to war.

1.1 The Standard Model

We begin by describing the standard setup for the inefficiency puzzle of war (Fearon, 1995). Two risk-neutral states, S (henceforth "she") and D (henceforth "he"), dispute the distribution of an infinitely divisible good whose size is normalized to 1. They can either agree to divide it peacefully or fight over it. War is a winner-take-all costly lottery: D wins with probability $p \in (0, 1)$, S wins with probability $1 - p$, and both suffer costs, $c_S, c_D > 0$. The expected payoffs from war are $p - c_D$ for D and $1 - p - c_S$ for S. Since they sum up to less than the total size of the benefit, there always exists a range of settlements, $B = [p - c_D, p + c_S]$, that both sides prefer to war. A fully rationalist account must explain why the two sides fail to reach an agreement when the existence of this range is common knowledge.

Consider now a simple variant of the ubiquitous ultimatum crisis bargaining model (henceforth, "standard model"). There is some status quo division of the good $(d, 1 - d)$, where $d \in [0, 1]$ is D's share and $1 - d$ is S's share. A state is "satisfied" with the status quo if its payoff from living peacefully at that distribution is at least as high as its expected payoff from war; otherwise it is "dissatisfied." If $d \in B$, both players are satisfied; if $d < p - c_D$, then D is dissatisfied but S is satisfied, and if $d > p + c_S$, then D is satisfied but S is dissatisfied. If both are satisfied, there is no crisis, there will be no revision of the status quo, and no war. These assumptions imply that at most one player can be dissatisfied (Powell, 1999). Suppose, without loss of generality, that D is dissatisfied. Then there is a crisis in which war cannot be avoided unless the status quo is revised in D's favor.

The sequence of moves is as follows: S makes a take-it-or-leave-it (TILI) offer $(y, 1 - y)$, where $y \in [0, 1]$ is D's proposed share and $1 - y$ is S's proposed share. If D accepts this offer, the status quo is revised accordingly and the game ends peacefully with players getting the payoffs from this new distribution. If D rejects this offer, the game ends in war.

With complete information, this model has a unique subgame-perfect equilibrium (SPE), in which D accepts any $y \geq p - c_D$, and S offers exactly $y^* = p - c_D$. The game ends peacefully with the status quo revised in D's favor. The fundamental point is that war will not occur under complete information; *states manage to avoid war because they agree to revise the status quo on mutually acceptable terms.*

To examine the mutual optimism explanation, consider an incomplete-information version of the model. To keep things as simple as possible for a crisp illustration of the results, we consider the two-type case with one-sided asymmetric information about military capabilities.[2] Assume that D can be strong, so his probability of winning the war is $p_h \in (0, 1)$, or weak, so his probability of winning is $p_w \in (0, 1)$ such that $p_w < p_h$. D knows his own type but S is uncertain: she believes that he is strong with probability q and weak with probability $1-q$. When she makes her

[2]Fearon (1995) considers uncertainty about the costs of fighting rather than relative military capabilities. We chose the latter because it is closer in spirit to the mutual optimism idea as originally proposed by Blainey (1988).

offer, S is unsure whether rejection would lead to a war with a strong opponent or a weak one. The solution concept is perfect-Bayesian equilibrium (PBE). The following proposition describes the well-known result. The proof is straightforward and is omitted.

PROPOSITION 1. *In all PBE, D accepts any $y \geq y_h = p_h - c_D$ if strong, any $y \geq y_w = p_w - c_D$ if weak, and rejects any other offer. The offer S makes depends on the critical belief threshold, $k = \frac{p_h - p_w}{p_h - p_w + c_D + c_S} \in (0, 1)$, as follows:*

(i) if $q > k$, then S offers y_h, which D always accepts;

(ii) if $q \leq k$, then S offers y_w, which D accepts if weak but rejects if strong.[3]

War occurs when $q \leq k$ and D happens to be strong. □

This, of course, is the risk-return trade-off result, which is among the most widely accepted rationalist explanations for war (Powell, 1999). The only way to avoid war in this setting is to offer D at least his expected payoff from war. S can always do this by offering peace terms that the opponent will accept if he is strong: they are so generous that he would certainly accept them if he happens to be weak. Even though S can always secure peace, doing so is not always optimal because such generous peace terms represent unnecessary concessions if D is weak. On the other hand, failing to offer them carries a risk of war if D is strong. S can resolve this dilemma by balancing the risk of having her offer rejected against the gain of obtaining better settlement terms if it is accepted. Because war is costly, this trade-off is only optimal if the risk is not too high; that is, if she believes that there is a good chance her opponent is weak.

In this and related models, the finding that war can occur under incomplete information but not under complete information is generally referred to as "war due to incomplete information"

[3]Technically, if $q = k$, S is indifferent between making the large and small offers, and could choose either one or mix. This is a knife-edge condition and is uninteresting.

rather than "war due to mutual optimism." However, we now offer a natural definition of optimism in this standard model, and show that war occurs if, and only if, there is mutual optimism. Thus, a very standard model establishes the coherence of the MO explanation for war. We then rebut every one of Fey & Ramsay's arguments that these models are incapable of capturing that explanation.

1.2 Conceptualizing Optimism in the Standard Model

Let us define optimism in the simple ultimatum crisis bargaining game that we have been considering. Consider the uninformed state, S, first. She can be said to be optimistic about her military prospects when she is sufficiently confident that she faces a weak opponent (when $q < k$). On the other hand, she can be said to be pessimistic when she is sufficiently confident that she faces a strong one. This is so because in our specification of the standard model, a war against a strong opponent ends in victory with a lower probability than a war against a weak opponent. If this were a model of two-sided incomplete information, D's optimism would be defined analogously. Because we assumed that he knows the actual military balance, the definition here boils down to the actual state of that balance. He can be said to be optimistic when he is the strong type (because his probability of winning is high), and pessimistic when he is the weak type (because that probability is low).

	D pessimistic (weak, p_w)	D optimistic (strong, p_h)
S pessimistic ($q > k$)	peace (S offers y_h, D accepts)	peace (S offers y_h, D accepts)
S optimistic ($q \leq k$)	peace (S offers y_w, D accepts)	**war** (S offers y_w, D rejects)

Table 1: Optimism and War in the Standard Model.

Table 1 shows the conditions for optimism under which war occurs in the standard model (this is just a tabular form of Proposition 1). When S is pessimistic, she make the generous offer y_h that D accepts regardless of type. Similarly, when D is pessimistic, he accepts any offer that S makes in equilibrium (because the worst terms she ever offers match his expected war payoff). Hence, with unilateral optimism war cannot occur: *mutual optimism is a necessary condition for war in the standard model*. If, on the other hand, both players are optimistic, S makes the low-ball offer y_w and D rejects it (because it is worse than his expected war payoff). Hence, with mutual optimism war always occurs: *mutual optimism is a sufficient condition for war in the standard model*. To summarize, in the standard model, *war occurs if, and only if, there is mutual optimism*. This provides a mechanism through which mutual optimism can lead to war and establishes the coherence of that explanation.

2 Why the Standard Model Is Appropriate

Fey & Ramsay make a number of arguments for why existing models of crisis bargaining (including the standard model that we have been analyzing, but more complicated variants as well) cannot appropriately examine the mutual optimism explanation for war. We can parcel them into five claims: (i) war is a unilateral act in the standard model and does not require the agreement of both players to occur; (ii) because one player can start a war without the consent of the other, MO is irrelevant to its occurrence (it is enough that only one side is optimistic); (iii) the risk-return trade-off is an alternative explanation for war, which should be considered separately from MO; (iv) when war does occur, players are not optimistic in the instant before war begins; and (v) players would want to avoid war on the “eve of war” but are prevented from doing so by arbitrary restrictions on the game-tree. These arguments are rather persuasive at first glance, and hence we rebut each one in some detail. In doing so, we also elucidate the rationalist mechanism of the mutual optimism explanation.

2.1 WAR *IS* A MUTUAL ACT

Fey & Ramsay argue that if war is a unilateral act—meaning that an actor can impose that outcome on the opponent—then "the concept of war by mutual optimism loses meaning" (745). We agree: for MO to provide an explanation for war, it should be the case that war occurs when optimistic actors make choices that only *collectively* lead to bargaining failure. For instance, if one side has a very high expected payoff from war but the other was not allowed to make any concessions, war would occur but it would not be caused by MO. The problem with Fey & Ramsay's argument is that, contrary to their claim, war is *not* unilateral in the standard model.

How do Fey & Ramsay propose to capture the idea of war as a mutual act? They require that each side has an action ("negotiate") that guarantees the peace outcome regardless of the action of the other player.[4] It might appear that in the standard model D unilaterally causes war because it is his rejection that starts it. Since S has no way to "stand firm," her consent appears unnecessary for war to occur.

However, observe that in the standard model D certainly has a strategy that guarantees peace regardless of what S does: accept any offer. Moreover, S *also* has a strategy that guarantees peace, at least for anything that D might rationally do: offer some $y > y_h$. It is true that this is slightly weaker than Fey & Ramsay's requirement because it does not guarantee peace for *any*

[4]Fey & Ramsay sometimes refer to this assumption as the requirement that both sides have to choose to "stand firm" for war to occur. As they note on pp. 739 and 745, this boils down to the same thing.

strategy of D (e.g., a strategy of rejecting all offers). However, it does guarantee peace for any strategy that D would rationally play: accept offers that are greater than his expected war payoff. These are the only strategies that D would ever play in equilibrium, and since Fey & Ramsay maintain the rationality assumption, our definition is very close to theirs.

Note further that in the endogenous peace-terms setting of the standard model (and unlike the Fey & Ramsay setup where these terms are specified exogenously), the satisfied state will never *choose* war. As Proposition 1 shows, it is always better to make an offer that would satisfy at least the weak opponent because doing so buys some positive probability of peace at terms that are strictly better than war. However, because that state also has the option to make the large peace-guaranteeing offer, the choice to make the limited one that carries a risk of war can naturally be interpreted as that state "standing firm."

This means that war is a mutual act in the standard model: it can occur *only* if S *chooses* to make a limited offer (which she certainly knows carries a positive risk of being rejected), and if D *chooses* to reject it. That is, war occurs only when both actors choose to forsake the strategies that guarantee a peaceful outcome. It takes two to make war here.

Finally, as we explain in some detail in the last part of the article where we consider Fey & Ramsay's model, the definition of mutuality they require is exceedingly demanding for it implies an ability to impose peace terms on an actor whose expected payoff from war is higher.

2.2 Mutual Optimism *Is* Relevant for the Occurrence of War

Fey & Ramsay further argue that "if the correct model of war is one in which any single country can start a war, the presence of mutual optimism is irrelevant and, therefore, not a coherent rationalist explanation of war" (751). However, as we demonstrate in Proposition 1 and Table 1, war occurs in the standard model *if, and only if, mutual optimism is present*, and hence the presence of mutual optimism is not just relevant, it in fact entirely determines the occurrence or non-occurrence of war.[5]

2.3 The Risk-Return Trade-Off Is Not an *Alternative* to Mutual Optimism

We have identified the risk-return trade-off as one means through which mutual optimism can result in war. Fey & Ramsay state that there might be other "viable rationalist explanations" for war, such as "incomplete information mechanisms that are not associated with mutual optimism" (752). The only other such mechanism they cite is the risk-return trade-off, and they are very explicit that they consider it an alternative to mutual optimism (750). (In fact, they reject the risk-return trade-off itself as dependent on arbitrary restrictions in the extensive form, a claim we deal with in the next section.)

It is not entirely clear to us what position Fey & Ramsay wish to take. As far as we have been able to discern, their argument is basically about private information and how it can directly cause war (e.g., when they write that "in this setting, the root cause of war is the inconsistent expectations that arise because of private information" and that they are "working with the definition of mutual optimism as war due to inconsistent beliefs," 738).

But this is puzzling. *Mutual optimism* is just a set of conditions that describe beliefs: both sides having high expectations about war (which might be due to private information about capabilities

[5]This also contradicts Fey & Ramsay's claim that their results "apply to any game where peace prevails in the absence of mutual optimism" (740). In the standard model, peace does indeed prevail in the absence of mutual optimism and yet their 'no-war' theorem does not hold.

but which also might be due to a host of other factors, as we noted above). The mutual optimism *explanation* must show how these expectations cause war; it has to specify the reasons these expectations cannot be reconciled without fighting. In particular, since peace requires that both sides agree to its terms, the mechanism must explain why actors persist in their unwillingness to offer terms that the opponent demands in order to preserve the peace. Our puzzlement stems from Fey & Ramsay's attempt to critique the *explanation* while simultaneously omitting the mechanism through which it operates.

The non-rationalist literature has skirted this requirement by arguing that actors would not reduce their optimism even in the face of abundant evidence that contradicts it simply because they are not rational. This venue is not open to us, but neither it is to Fey & Ramsay who are also interested in a rationalist specification of the MO mechanism.

In contrast to Fey & Ramsay, we have provided such a mechanism: the risk-return trade-off. We have argued that this is not an *alternative* to the MO explanation, it is one mechanism through which MO can cause war. High expectations about war (because she believes D is likely weak) cause S to forsake the strategy that guarantees peace and to make a limited offer, which she is fully aware carries a risk of war, to D. High expectations about war cause D to reject this offer even though he is fully aware that doing so will result in war. Thus, when MO is present, the actors engage in specific behaviors and their interaction ends in war.

There are other means through which MO can cause war. For instance, when actors believe their opponent has unreasonably high expectations about war, they might attempt to lower them, which usually entails taking actions that run a risk of war (Schelling, 1966). As is well-known, when both sides are very optimistic, they can end up taking actions that commit them to war (Fearon, 1994). Thus, *costly signaling,* or the attempt to overcome the problem created by MO, is another way through which MO can lead to war. Another example is in contexts where military preparations are very costly: an optimistic actor might choose to under-prepare in the belief that his opponent is weak, and his force levels might prove inadequate to compel an optimistic strong opponent (Slantchev, 2005).

In their desire to purge their model of any such 'alternatives', Fey & Ramsay have ended up voiding the MO mechanism. It is no surprise that they find that mutual optimism cannot cause war; after all, they have ruled out the very mechanisms through which mutual optimism is theorized to do so.

2.4 Players *Can* Be Optimistic "On the Eve of War"

Fey & Ramsay argue that when war occurs under incomplete information in the standard model, it is not war due to MO because the uninformed actor is no longer optimistic "on the eve of war." This can be seen very easily in the model that we have been analyzing. Suppose that $q \leq k$ and D is strong, so that mutual optimism is present. In equilibrium, S offers y_w, which D rejects. Now, in the instant after that rejection but before war begins, S is no longer optimistic. Since D's strategy is to accept y_w if weak, rejection clearly reveals that he is strong. Thus, upon observing that rejection, S will immediately update her beliefs and conclude that D is strong. And we know that in an environment where war is costly, she is better off satisfying the strong type instead of fighting: $1 - y_h = 1 - p_h + c_D > 1 - p_h - c_S$. In other words, S would prefer to offer y_h and avoid war but is prevented from doing so by the extensive form of the game. As Fey & Ramsay put it, there is no longer mutual optimism "on the eve of war," and hence it is not really war due to mutual optimism.

This line of reasoning has two components—the claim that following rejection of her offer, S

would want revise its terms to *ensure peace*, and the claim that she is *artificially constrained by the game-tree*. We deal with the second claim in the next section, and in this section show that although the first claim does indeed obtain in the two-type model, it fails in models with more than two types (which preserve the MO risk-return trade-off results intact).

Consider a variant of the standard model where D can be either weak, p_w, moderately strong, p_m, or very strong, p_h, with $p_h > p_m > p_w$. S is unsure which type she is facing, but believes that her opponent is strong with probability $q_h \in (0, 1)$, moderate with probability $q_m \in (0, 1)$, and weak with probability $1-q_m-q_h \in (0, 1)$. The following proposition establishes that mutual optimism will cause war through the risk-return trade-off mechanism in this model as well.

PROPOSITION 2. *In all PBE, D accepts any $y \geq y_h = p_h - c_D$ if very strong, any $y \geq y_m = p_m - c_D$ if moderately strong, any $y \geq y_w = p_w - c_D$ if weak, and rejects any other offer. The offer S makes depends on the critical belief thresholds*

$$k_1 = 1 - q_m\left(1 + \frac{C}{p_m - p_w}\right), \qquad k_2 = \frac{p_h - p_m}{p_h - p_m + C}, \qquad \text{and}$$

$$k_3 = \frac{p_h - p_w - q_m(p_m - p_w + C)}{p_h - p_w + C},$$

where $C = c_D + c_S$, as follows:

(i) if $q_h > \max\{k_2, k_3\}$, then S offers y_h, which D always accepts;

(ii) if $q_h < \min\{k_1, k_3\}$, then S offers y_w, which D accepts only if weak;

(iii) if $k_1 < q_h < k_2$, then S offers y_m, which D accepts only if weak or moderately strong.

War occurs if and only if S is sufficiently optimistic and D sufficiently strong. □

Proof. The strategy for D follows from subgame perfection, and implies that if p_h accepts some offer y, then so will p_m and p_w, and if p_m accepts some offer, then so will p_w. It follows that S will choose among three possible offers: y_h, which all three types accept; y_m, which only the weak and moderate types accept; and y_w, which only the weak type accepts. S will always prefer to offer at least y_w rather than some unacceptable offer $y' < y_w$ that would certainly lead to war: $U_S(y_w) - U_S(y') = (1 - q_h - q_m)C > 0$. Therefore, we only need to consider her preference among the three offers that provide for a chance of peace. As before, we ignore knife-edge conditions. Algebra shows that S prefers y_w to y_m when $q_h < k_1$, prefers y_m to y_h when $q_h < k_2$, and prefers y_w to y_h when $q_h < k_3$. Conditions (i), (ii), and (iii) follow immediately. ■

Suppose now that $q_h < \min\{k_1, k_3\}$, so that S is very optimistic and offers D the worst possible terms that he might ever accept, y_w. Since D would only accept this if he is weak, rejection signals that he is either very strong or moderately strong. The only way for S to ensure peace now would be to offer y_h. It is easy to see that there exist conditions under which S *will not want to ensure peace* with such a generous offer (if she could make another take-it-or-leave-it offer) but would instead make another limited offer, y_m this time, which the strong opponent would still reject, causing war. As we know from Proposition 1, the condition for this preference is that S's belief that D is strong (which is the posterior $q_h/(q_m+q_h)$ by Bayes rule) is sufficiently small. Since we already supposed that q_h is small relative to q_m, this condition is easy to satisfy.[6] In other words, S has retained sufficient optimism even "on the eve of war" and would not make

[6]As a numerical example, suppose that $c_D = c_S = 0.1$, $p_h = 3/4$, $p_m = 1/2$, $p_w = 1/4$, and $q_m = 1/3$. Then, $k_1 = 0.4$ and $k_3 = 0.5$, and so S will offer y_w if $q_h < 0.4$. If this offer is rejected, then if S could make another take-it-or-leave-it offer, she would offer y_m rather than y_h if $q_h/(q_m + q_h) < 5/9$, which simplifies to $q_h < 5/12$ (> 0.4). Thus, if $q_h < 0.4$, then S initially offers y_w, and if this offer is rejected and she could make another take-it-or-leave-it offer, she would still be optimistic enough to offer y_m rather than y_h.

the offer that would guarantee peace. Fey & Ramsay's claim is an artifact of the two-type case (it also fails if we assume any countable number of types greater than two or a continuum).

One possible response to this would be to agree that S will retain optimism after her initial demand but to argue that she would not be optimistic if she made a second offer and *that* got rejected as well. (It might take many more rejections depending on the degree of uncertainty and amount of initial optimism.) In other words, there will always be a "final" offer whose rejection would reduce S's optimism to the point that she would prefer to make the offer that would guarantee peace. But of course, this just begs the question of how precisely this optimism gets reduced, a question that Fey & Ramsay completely avoid but that is crucial to the MO explanation for war. This leads us to a fundamental point that also addresses perhaps Fey & Ramsay's most sweeping claim about the inability of existing models to examine the mutual optimism argument.

2.5 War Is *Not* an Artifact of Arbitrary Restrictions on the Game-Tree

The final criticism that Fey & Ramsay level at the standard model's ability to capture the mutual optimism argument is that in it, war occurs under incomplete information only because the extensive-form of the game does not permit the actors to avoid it once they realize that they do not wish to fight and that they would rather reach a peace settlement. Their argument is clearest in the two-type case that we have analyzed: after her offer is rejected, S knows that she is facing the strong D, and would strictly prefer to make a war-avoiding last-ditch proposal that would satisfy him. However, the game-form does not allow her to make that choice because D's rejection automatically causes war.[7] As Fey & Ramsay put it,

> Reflecting on this example, we can give an intuitive statement of our main result in the following way. If it is common knowledge that countries are going to fight, and these countries have a "hot line" available, then at least one side will always want to 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 not only before making it, but after a proposal has been made and rejected. In situations where a firm offer is made that, if rejected, leads to certain war, equilibria with war can exist, but not because of mutual optimism; one side would prefer to settle but is locked into a war by the extensive form (751).

If this argument is correct, it has serious consequences for almost *all* game-theoretic models of crisis bargaining under incomplete information, not just the standard ultimatum model, because these models all have "final" decision-nodes. That is, decision-nodes where the choice to "fight"

[7] It seems that this is also why Fey & Ramsay reject the risk-return trade-off, where "there is no way for $[S]$ to react to the private information of $[D]$" (750), as a viable explanation for war. They cite Leventoğlu and Tarar (2008) as allegedly having shown that the risk-return result "is driven to a large extent by (somewhat arbitrary) assumptions regarding the extensive form of the bargaining process" (fn. 17). However, what the latter actually show is that private information does not necessarily cause risk-return behavior, as suggested by Powell's (1999) finding of a unique equilibrium with that feature. They show that uniqueness depends on one actor having all the bargaining power, and that if that is not the case, multiple equilibria become possible, and some of them do not exhibit the risk-return trade-off. This implies that mutual optimism may not always "activate" this particular war-causing mechanism, but does not imply that the mechanism is unreasonable or unrealistic.

starts a war irrespective of how this choice would update the other actor's beliefs; that is, regardless of whether she would want to make an "eve of war" offer. Because these models disallow such presumably war-avoiding offers, they impose war by analyst fiat, not as a consequence of optimal behavior. As a result, they do not really explain war at all.[8]

What features do Fey & Ramsay require of a legitimate explanation of war under incomplete information? They write that "in practice, if one country chooses to stand firm, the other country can stop a war by inducing the bargaining procedure instead" (745). This implies that any actor who would rather bargain than fight should be able to avoid war even after his opponent has chosen to fight. As they put it, "at any given moment before war begins, a state could continue negotiations with the hopes of avoiding a fight" (739). From a substantive standpoint, it is very doubtful that 'in practice' an actor can always avoid war when the opponent has decided to fight; and it is perhaps even more doubtful that he can do so 'at any given moment before war begins'. We leave these problems aside to focus on the theoretical implications that Fey & Ramsay's claim, if true, would have for crisis bargaining behavior.

To examine their "hotline" argument, we now modify the standard model to implement Fey & Ramsay's requirements by allowing S to "induce the bargaining procedure" and "continue negotiations" after D has chosen to fight. Consider the following infinite-horizon game: in each period S makes an offer, which D can either accept or reject. If he accepts, the game ends on the peace terms accepted. If he rejects, S can either let the rejection stand or continue negotiations. If she lets the rejection stand, the game ends in war; otherwise, the game continues to the next

[8]As we show in the next part of the article, Fey & Ramsay's own model is vulnerable to a (much stronger) analogue of this argument because it does not allow D to react to S's unilateral imposition of peace terms that he finds worse than war.

period, where S makes a new offer. The status quo distribution of benefits is $(d, 1-d)$, where d is D's share. The payoffs are as follows. If an agreement is reached on some division $(y, 1-y)$ in period t $(t = 0, 1, 2, \ldots)$, then D's payoff is $\sum_{i=0}^{t-1} \delta^i d + \sum_{i=t}^{\infty} \delta^i y$, and S's payoff is $\sum_{i=0}^{t-1} \delta^i (1-d) + \sum_{i=t}^{\infty} \delta^i (1-y)$, where $\delta \in (0,1)$ is the common discount factor. (If the players never reach an agreement, the status quo distribution remains in place forever.) If players go to war in some period t, then D's payoff is $\sum_{i=0}^{t-1} \delta^i d + \sum_{i=t}^{\infty}(p - c_D)$, and S's payoff is $\sum_{i=0}^{t-1} \delta^i (1-d) + \sum_{i=t}^{\infty}(1 - p - c_S)$. As before, assume that S is satisfied but D is not. This modification is sufficient to implement Fey & Ramsay's 'hotline' notion that an actor that prefers negotiations to war should have the opportunity to avoid fighting by continuing negotiations even on the "eve of war." The following proposition fully characterizes the set of equilibrium outcomes.

PROPOSITION 3. *Every subgame-perfect equilibrium of the modified infinite-horizon game is peaceful and the status quo is never revised.* □

Proof. The only way the game can ever end in war is for S to let D's rejection stand. In any arbitrary period, she can ensure the status quo payoff by any strategy that always continues bargaining. Because she is satisfied, this payoff is strictly higher than her expected war payoff. Hence, S will never let rejection stand. To see that the status quo will never get revised, note that any revision requires that S propose it and D accept it. Because D is already dissatisfied, he will never agree to any redistribution that would leave him worse than the status quo (i.e., any $y < d$), so S cannot obtain better terms. Similarly, S would never propose terms worse than the status quo (i.e., any $y > d$) if these have any chance of being accepted (because she can always keep the status quo payoff). ■

In equilibrium, the interaction can end in one of two ways: either in some period S offers the status quo terms and D accepts them, or D always rejects her offers but she never lets the rejection stand (so "negotiations" continue forever). It is straightforward to construct SPE with immediate acceptance of the status quo or interminable negotiations.[9] The structure and degree of incomplete information about D—indeed, whether it is even present—are completely irrelevant. Regardless of how her beliefs might have changed during the interaction, S will never agree to a revision of the status quo in D's favor. *In every equilibrium, the dissatisfied actor is worse off than his expected payoff from war.*

This result is obviously substantively silly, but it illustrates the problem with Fey & Ramsay's hotline argument. They essentially claim that when war occurs under incomplete information in existing models, (i) at least one actor would want to make a peace-ensuring offer on the "eve of war" but (ii) is prevented from doing so by the structure of the game-tree. We have shown that these two assertions are incompatible: if we fix (ii) along the lines they require in a model of crisis bargaining, then (i) no longer follows. If the satisfied state can always make a new offer to forestall war, then she never has an incentive to make an acceptable offer because she strictly prefers the status quo and can ensure it by repeatedly making unacceptable offers whenever the dissatisfied state opts for war. In the hotline world, peace can be had for free: no concessions are necessary to induce the opponent not to fight because he cannot fight without her consent.

Note that in the standard ultimatum model, war is also avoided under complete information, but for sensible reasons: S makes an offer that D (as well as S) finds preferable to war. Having "final" decision-nodes is a necessary condition for a model to give sensible results, and is not a means for allowing war to artificially occur under incomplete information. Such decision nodes

[9] See the online appendix for SPE that have these features.

are also warranted on substantive grounds: would anyone doubt that if D's alternatives were interminable negotiations and an unpalatable status quo that he would cut the knot by attacking? It is also highly doubtful that war can be avoided up to the last instant before it begins, even with an offer of concessions. Crises may very well have (endogenous) "deadlines", and it is the presence of these (usually modeled as final decision nodes) that exerts the coercive pressure on the participants.

A model without "final" decision-nodes eliminates by assumption another mechanism through which mutual optimism can lead to war: credible information transmission. In the simple ultimatum model with two types of D, D's rejection of S's small offer is *informative*: D only rejects if strong. Moreover, S *always* makes at least an offer that is acceptable to the weak type of D, and hence rejection of that offer must be (at least partially, if there are more than two types) informative. However, in the hotline world, S always proposes an offer that is unacceptable to even the weak type of dissatisfied state because she prefers the status quo to satisfying even the weak type. Hence, rejection of the offer is not informative, and S's prior belief is never revised. Information transmission cannot occur in the hotline world because there is no risk of war: the satisfied state can perpetually (and unilaterally) ensure the status quo by always just making a new offer. A core insight of the literature on credible signaling in international relations is that for a signal to be informative, it generally must create a real risk of war. In models that explicitly incorporate signaling mechanisms such as audience costs or military mobilization, this risk is created endogenously, and in models without explicit signaling mechanisms, such as the simple ultimatum game, this is captured through the existence of "final" decision-nodes which leads to rejection creating a risk of war and hence allowing for information-transmission. In the hotline world, it is not surprising that mutual optimism cannot cause war because this world eliminates the very mechanism through which mutual optimism is overcome when initially present: credible signaling. In the hotline world, the satisfied state can unilaterally impose peace on terms that the dissatisfied state finds worse than war, and hence signaling becomes irrelevant.

3 Why War Does Not Occur in the Fey-Ramsay Model

So far, we have presented a "modern synthesis" of the mutual optimism explanation for war based on the rationalist work on the causes of war. In particular, we have gone beyond the simple informal argument that mutual optimism creates a situation where incompatible beliefs about the likely outcome of a war create a situation where no mutually acceptable agreement is obvious, and hence war inevitably occurs. Instead, we have shown how existing work provides behavioral mechanisms that link optimistic beliefs and the outbreak of war. In these accounts, incompatible beliefs cause actors to *engage in behavior* (often to try to *overcome* the very same incompatible beliefs) that causes war to occur with positive probability. The two most prominent such mechanisms are the risk-return trade-off and credible signaling. These are not alternative (to mutual optimism) explanations for how incomplete information can lead to war; they are the *very means* through which mutual optimism can lead to war. That this has not been recognized until now is probably because these works do not explicitly seek to validate the mutual optimism explanation, but we have shown here how they in fact do so.

We do, however, recognize that one need not accept our rationalization of the MO argument. The remaining question then would be whether this also entails the acceptance of Fey & Ramsay's class of models with the corresponding rejection of the MO explanation. In this final section, we briefly show that Fey & Ramsay's own model does not (and cannot) invalidate the mutual optimism explanation. This is because, like the infinite-horizon hotline model that we presented above, their model allows one actor to unilaterally impose a negotiated settlement that the other side finds worse than war. As in the hotline model, it is not surprising to find that war cannot occur in equilibrium in such a setting, but the reason has nothing to do with mutual optimism.

3.1 The Basic Fey & Ramsay Model

Fey & Ramsay's approach to invalidating the mutual optimism explanation for war is to analyze a general class of models that supposedly more accurately captures the mutual optimism argument than do standard crisis bargaining models, and show that in this class of models there exists no (Bayesian Nash) equilibrium in which war occurs (we shall henceforth refer to this, their main result in Theorem 1, as the "no-war" result.) While their approach permits the analysis of an entire class of models, it is quite abstract and they never offer an example of an actual model that belongs to this class. We construct just such a model and use it to show that the "no-war" result follows from the unilateral-peace assumption that they make, rather than from anything having to do with mutual optimism.

In Fey & Ramsay's model, two states, which we label S and D for comparability with our earlier model, are embroiled in a crisis and simultaneously choose from a set of actions. Given Fey & Ramsay's assumption that the negotiation payoffs are unique in each state of the world, we can reduce the set of actions to two: "stand firm" (F) and "negotiate" (N). Since neither war nor negotiation payoffs can depend on how that outcome is reached, nothing is added by considering more complicated action sets. To capture the notion of war as a mutual act, Fey & Ramsay assume that it only occurs if both actors choose to stand firm, $\langle F, F\rangle$. Otherwise, the outcome is a negotiated settlement where payoffs are identical for the strategy profiles $\langle F, N\rangle$, $\langle N, F\rangle$, and $\langle N, N\rangle$.

Whereas payoffs are not allowed to depend on the crisis behavior of the actors, they can depend on the true state of the world denoted by ω, which can be one from a countably finite set $\Omega = \{\omega_1, \omega_2, \ldots, \omega_K\}$. Actors may obtain private information about the true state of the world before they make their choices. The war payoff in some state ω is specified as in the common costly-lottery winner-take-all model we have been using, $W_i(\omega) = p_i(\omega) - c_i(\omega)$, where $p_i(\omega) \in (0, 1)$ is actor i's probability of winning if the true state of the world is ω, and $c_i(\omega) > 0$ denotes his war costs in that state. The usual assumptions apply: ties are not allowed, $p_S(\omega) + p_D(\omega) = 1$, and war is inefficient: $W_S(\omega) + W_D(\omega) < 1$ for all $\omega \in \Omega$. Negotiated settlements, denoted by $r_i(\omega)$, on the other hand, are efficient: $r_S(\omega) + r_D(\omega) = 1$. (This is just for convenience: the results hold as long as negotiations are less costly than war.) The negotiation payoffs are specified exogenously for each state of the world and cannot vary with the strategies used during the crisis.

		D: N	D: F
S	N	$r_S(\omega), r_D(\omega)$	$r_S(\omega), r_D(\omega)$
	F	$r_S(\omega), r_D(\omega)$	$W_S(\omega), W_D(\omega)$

Figure 1: Fey & Ramsay's basic model.

Figure 1 shows Fey & Ramsay's basic model. Private information about ω obeys standard rationality postulates (e.g., a player's information partitions are such that he cannot exclude the true state of the world from the set of states he believes possible; see Osborne and Rubinstein (1994, Ch. 5) for formal definitions). Fey & Ramsay's "no-war" theorem establishes that $\langle F, F\rangle$ can never be a Bayesian-Nash equilibrium. Because this result applies to all games in this class, they conclude that "our result that war cannot occur in equilibrium implies that mutual optimism is not a valid rationalist explanation for war… Our result shows that one prominent explanation, war by mutual optimism, is not a coherent and internally consistent theory of war within the rationalist framework" (752).

How does this inference work? Let us grant for a moment that this model is ideally suited to examine the MO explanation, as they claim.[10] Fey & Ramsay insist, correctly, that if MO is to make any sense as an explanation, *war should not occur if actors have complete information, only when they have private information.* Fey & Ramsay reason that if war does not occur in equilibrium in the latter case under *any* information partitions satisfying the standard game-theoretic postulates, then MO is not a valid rationalist explanation for war. They further conclude that in models in which war occurs only under incomplete information, it does so only because of artificial restrictions in the game-tree. We have already dealt with the last claim, and so here we just examine the reasonableness of Fey & Ramsay's own model.

3.2 War Does Not Occur with Complete Information... Even When It *Should*

Consider Fey & Ramsay's model with complete information, i.e., there is only one state of the world, ω_1. Suppose now that in this state of the world $p_i(\omega_1) = 1/2$ and $c_i(\omega_1) = 1/4$, so that $W_i(\omega_1) = 1/4$. Bargaining is efficient and $r_i(\omega_1) = 1/2$, so each actor strictly prefers the exogenous negotiated settlement to war. The model is shown in Figure 2(a).

It is readily verified that only $\langle N, N\rangle$, $\langle N, F\rangle$, and $\langle F, N\rangle$ are Nash equilibria. In particular, $\langle F, F\rangle$ is not an equilibrium, and so the "no-war" result holds. The Fey & Ramsay assumptions

[10]For example, on p.739, they write that they are analyzing "a class of games that capture the key features of the mutual optimism argument.", and on p.750 that they are "formalizing the mutual optimism hypothesis and using assumptions designed to test this hypothesis" and "Our game-theoretic setting is thus designed to create conditions in which there is a clear link between mutual optimism and war."

		D	
		N	F
S	N	$1/2$, $1/2$	$1/2$, $1/2$
	F	$1/2$, $1/2$	$1/4$, $1/4$

(a) State of the world is ω_1.

		D	
		N	F
S	N	$4/5$, $1/5$	$4/5$, $1/5$
	F	$4/5$, $1/5$	$1/4$, $1/4$

(b) State of the world is ω_2

Figure 2: Parameterized Specifications of Fey & Ramsay's Model.

do not appear problematic.

Consider now an alternative state of the world, ω_2, where everything is the same as in ω_1 except that $r_S(\omega_2) = 4/5$ and $r_D(\omega_2) = 1/5$. This model is shown in Figure 2(b). Here, S still prefers the negotiated settlement to war but now D strictly prefers war. It is easily verified that only $\langle N, N \rangle$ and $\langle N, F \rangle$ are Nash equilibria. Again, $\langle F, F \rangle$ is not an equilibrium, and the "no-war" result holds. However, we now have a problem because given the payoffs, *war should occur* in this scenario!

To see this, observe that in both Nash equilibria D "accepts" a peace settlement that gives him a payoff that is strictly worse than his payoff from war. In other words, D would rather fight than accept such a meager settlement but cannot do so. Why not? Because if he chooses F, S's best response is to choose N and avoid the war. Fey & Ramsay innocuously refer to this ability as "inducing the bargaining procedure" but what it really means is an ability to unilaterally impose peace on an opponent who strictly prefers war. Therefore, Fey & Ramsay's model *artificially precludes war through structural assumptions that have nothing to do with information or mutual optimism*.

Compare this with the no-war result that obtains in the standard ultimatum model with complete information. In that model, peace obtains for sensible reasons: S makes an offer that both sides find at least as good as war. Crucially, peace obtains precisely *because* of the lack of mutual optimism: S knows exactly how much she needs to offer D, and chooses to make such an offer. In Fey & Ramsay's model, on the other hand, peace obtains simply because S unilaterally imposes a peace that D finds worse than war, and does not even have to offer concessions.

Fey & Ramsay rightly insist that in a model that purports to capture the mutual optimism argument, war must be a mutual act in that it only occurs if *both* actors "stand firm," i.e., forgo strategies that would guarantee peace. Contrary to their claim, we have argued that standard models of crisis bargaining in fact capture this. Standard endogenous-offer models *also* capture the requirement of any sensible model of crisis bargaining that any negotiated settlement obtained in equilibrium must make *both* actors, given their beliefs, at least as well off as war. Fey & Ramsay's model does not capture this second crucial requirement, and hence it is not surprising that war cannot occur in equilibrium, but such a finding tells us nothing about the mutual optimism explanation for war.[11]

[11]In the online appendix, we use a simple sequential-move game of incomplete information that captures all of Fey & Ramsay's assumptions, and present a very accessible proof (that does not rely on the technology of models of knowledge) of their "no-war" result. We then show that in this model, the two Fey & Ramsay assumptions of (i) unilateral peace (each player can unilaterally avoid war by choosing to negotiate), and (ii) behavior-independent peace payoffs (the terms of the negotiated settlement cannot be influenced by the behavior of the players), are, given the standard assumptions, necessary and sufficient for the "no-war" result to hold, and that relaxing either of them leads to war occurring when there is sufficient optimism. We also argue that these assumptions are at odds with our modern notions of crisis bargaining, and are hence unjustified even if one's goal is only to examine the mutual optimism argument rather than crisis bargaining in general.

4 Conclusion

The mutual optimism explanation is perhaps one of the most venerated explanations for how war can occur. Until now, however, it has been presented primarily in informal terms. The informal literature does not explicitly describe mechanisms for exactly *how* inconsistent beliefs about the likely outcome of war will actually cause war to occur, perhaps because it seems intuitively obvious. The rationalist research agenda, on the other hand, is much more demanding of a purported explanation because it requires one to show how initial conditions (mutual optimism) result in rational (equilibrium) behavior that involves starting a war. The particular vein of this research agenda that we have examined is even more demanding for it also assumes that for these actors war is the costliest dispute-resolution method and carries no inherent benefits. Even then, we have identified two rationalist mechanisms through which mutual optimism can lead to war: the risk-return tradeoff and costly signaling.

The need for such a synthesis was brought to our attention by Fey & Ramsay's provocative article, in which they write:

> 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. (741)

This simple story does not consider the type of behavior that mutual optimism causes actors to engage in and how that behavior might influence the terms of settlements they pursue. In contrast, the rationalist literature examines how peace-seeking (because war is costly) but dissatisfied states may try to overcome mutual optimism so that a mutually beneficial settlement can be reached. In this view (and in contrast to the informal work on MO), actors do not simply "choose to fight rather than pursue a peaceful settlement." Instead, they try to reach a peaceful settlement but when mutual optimism is initially present, this attempt may entail actions that generate a real risk of war. Such actions are often the only way that mutual optimism can be overcome. Hence, although they have not been previously characterized this way, standard models of crisis bargaining in fact provide coherent mechanisms for how mutual optimism can lead to war, and thus provide a coherent elucidation and justification of this venerated explanation for war.

In closing, we contrast the crisis bargaining setting with that of efficient economic exchange (the analysis of the latter is where results analogous to Fey & Ramsay's "no-war" theorem were first obtained). The differences between them are so radical that the approach used to analyze one setting cannot readily be imported to analyze the other. As Fey & Ramsay note (740 and the cites therein), a central result of the efficient exchange literature is a "no-trade theorem" that establishes that "it cannot be common knowledge between two individuals that both will gain from a trade of a risky asset." They also write that the "fundamental reason that mutual optimism cannot lead to war is that if both sides are willing to fight, each should infer that they have either underestimated the strength of the opponent or overestimated their own strength. In either case, these inferences lead to a peaceful settlement of the dispute" (738). This is precisely why a trade of a risky asset cannot occur in equilibrium but what does it tell us about whether war can occur due to mutual optimism?

The decision to enter into a trade of a risky asset (the equivalent of war in our setting) is voluntary and non-coercive: if some player does not like what her private information tells her about the likely consequences of the trade, she can simply decline to trade, which would keep the status quo intact. Even if the other player's private information tells him that the trade is much more preferable than the status quo, he cannot force her to enter into the trade, or force her to compensate him for *allowing* her to decline the trade. In the trade setting, it is indeed quite sensible to allow either actor to unilaterally impose "no-trade" (the equivalent of peace in our setting) on the other, no matter how unpalatable the other actor finds the absence of trade (i.e., the status quo). But this is certainly not acceptable in any legitimate model of crisis bargaining.

While it is possible for an actor, whose private information tells her that war is undesirable, to avoid war "unilaterally" in the crisis bargaining setting as well, she can only do so by making an offer that the other side finds at least as good as war. That is, she cannot unilaterally *impose* peace on just any terms. Thus, we would argue that a crucial requirement of any legitimate model of crisis bargaining is that any peace agreement that is reached in equilibrium must make each actor at least as well off as going to war: peace must be mutually acceptable.

Another fundamental difference is in the role of private information and signaling. In the trade setting, if I obtain private information that the trade is very likely to benefit me (i.e., that I am strong), I want to convince the other side that I am weak. This would induce her to trade and allow me to benefit from my strength. If I am weak, on the other hand, I have no incentive whatsoever to trade or pretend to want to. In this setting, the strong has incentives to sandbag and the weak has no incentives to bluff. The only motivation to signal anything comes when one is strong, otherwise signaling is irrelevant.

This is not so in crisis bargaining where war is costly and negotiated settlements can avoid it. Here, I want to reach a peaceful settlement even when I am strong. However, I also want terms that are sufficiently favorable given my expectations about fighting. Since my benefit comes from the other side conceding to my terms rather than fighting, my incentive is to reveal that I am strong, not to trick the opponent to blunder into a war. In contrast with the trading scenario, I benefit from sincerely signaling my strength. Unfortunately, and also in contrast with the trading scenario, I also have an incentive to bluff when I am weak because doing so might convince the other side to give me a better deal. Thus, whereas the core problem in the trade setting is that the strong cannot conceal their strength, the core problem in the crisis bargaining setting is that the strong cannot reveal it.

The implications of this difference become clearer when we consider the informational content of behavior. Suppose that agreement to trade or start a war can be signified by a handshake, so it is a mutual act. In the trade setting, when I extend my hand to shake the other side's hand, my doing so is a clear indication of my willingness to trade. This is because I would only trade if I believe that I am in a strong position. I would not extend my hand if I believe I am in a weak position because there is no benefit in doing so and there might be a serious cost if she grabs my hand and shakes it before I can withdraw. Thus, extending my hand is unambiguously related to my belief in my strength, and she can use that to update her beliefs accordingly. Since she has the exact same incentives that I do, if she extends her hand after observing my behavior, I can also infer that she believes she is in a strong position, and will thus update my beliefs accordingly. In this way, the gradual extending of hands leads to *frictionless learning* about each other's beliefs, and in the end will prevent the trade from actually occurring. The stronger player will not be able to benefit from his strength. He would dearly love to be able to make his behavior less revealing but since the trade can only occur upon a handshake, he cannot do so because he must still extend his hand.

As the quotation above shows, Fey & Ramsay's reasoning about why war cannot occur relies precisely on this sort of reasoning which, as we have just seen, requires frictionless learning. The problem is that in a crisis bargaining setting, extending my hand is no longer a credible signal of my willingness to fight. When the other side sees me extend my hand, she would not know whether I am strong and sincere or whether I am weak and bluffing. I might extend my hand hoping that she would agree to give me good terms rather than actually shake it and start a war. Learning can only occur if this behavior can be made informative, and the only way to do that is to provide some sort of disincentive for the weak type to engage in it. In this simple setting, the required disincentive will come from a higher risk of a handshake: the other player must start extending her hand, hoping that the risk will cause the weak type to give up before the handshake occurs. Now we are playing a "game of chicken" rather than sincerely revealing our willingness to trade, and the results are well-known. Frictionless learning of the type envisioned by Fey & Ramsay cannot occur, and the friction that does permit some learning is provided by the risk of the very outcome both players would rather avoid.

The strategic gap between the trade and crisis bargaining settings is so staggering that there is no reason to expect that results obtained in one setting will readily transfer to the other. Aumann's (1976) "impossibility of agreeing-to-disagree" result is deservedly famous but evidently has no implications for the mutual optimism explanation of war.

REFERENCES

Aumann, Robert J. 1976. "Agreeing to Disagree." *Annals of Statistics* 4(6): 1236–9.

Blainey, Geoffrey. 1988. *The Causes of War*. 3rd ed. New York: The Free Press.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(September): 577–92.

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

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

Johnson, Dominic D. P. 2004. *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Cambridge: Harvard University Press.

Lebow, Richard Ned. 1981. *Between Peace and War: The Nature of International Crisis*. Baltimore: The Johns Hopkins University Press.

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

Levy, Jack S. 1983. "Misperception and the Causes of War: Theoretical Linkages and Analytical Problems." *World Politics* 36(October): 76–99.

Osborne, Martin J., and Ariel Rubinstein. 1994. *A Course in Game Theory*. Cambridge: The M.I.T. Press.

Powell, Robert. 1999. *In the Shadow of Power*. Princeton: Princeton University Press.

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven: Yale University Press.

Slantchev, Branislav L. 2005. "Military Coercion in Interstate Crises." *American Political Science Review* 99(November): 533–547.

Stoessinger, John G. 2005. *Why Nations Go To War*. 9th ed. Belmont: Wadsworth.

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

Wittman, Donald. 1979. "How a War Ends: A Rational Model Approach." *Journal of Conflict Resolution* 23(December): 743–63.

Online Appendix for
“Mutual Optimism as a Rationalist Explanation of War”

Branislav L. Slantchev
Department of Political Science, University of California – San Diego

Ahmer Tarar
Department of Political Science, Texas A&M University

June 4, 2010

Abstract.

In this appendix to the paper published in the *American Journal of Political Science,* we elaborate on two points. First, we exhibit two SPE for the infinite-horizon model that satisfies the Fey & Ramsay requirement of unilateral war-avoidance: in one, agreement is never reached, and in the other, agreement is reached immediately on the status quo distribution. Second, we demonstrate that the prevalence of peace in the Fey & Ramsay framework is due to their structural assumptions of unilateral peace and payoff-irrelevant behavior.

---

1 The Strange SPE in the Infinite-Horizon Model

In the article we showed that when we implement Fey & Ramsay's assumptions in a dynamic bargaining model with infinite horizon, there exists no SPE in which the status quo is revised or war occurs even though one of the players strictly prefers war to his status quo share. Now we exhibit two SPE that have this property: in the first one, agreement is never reached, and in the second one agreement is immediately reached without revision of the status quo. In either case war does not occur.

The model is described in the main text. In the proofs of the following two propositions, we use the "one-stage-deviation principle," henceforth OSDP, for infinite horizon games with discounting of future payoffs (Fudenberg and Tirole, 1991, 108-10). This principle states that, to verify that a profile of strategies comprises a SPE, one just has to verify that, given the other players' strategies, no player can improve her payoff at any history at which it is her turn to move by deviating from her equilibrium strategy at that history and then reverting to her equilibrium strategy afterwards.

PROPOSITION 1. *The following is a SPE in which an agreement is never reached, with the status quo forever remaining in place:*

(a) S always proposes some $y < d$. At any history at which D has rejected S's offer, S continues negotiations (rather than letting the rejection stand).

(b) D always accept any $y \geq d$. □

Proof. Does S's offer satisfy the OSDP? By offering $y < d$, and then continuing with her equilibrium strategy in the future, agreement is never reached and hence S's payoff is $\frac{1-d}{1-\delta}$. If S adopts a one-shot deviation by making a *different* unacceptable offer and then reverting to her equilibrium strategy, her payoff is the same, and hence such a one-shot deviation is not profitable. If she makes a one-shot deviation by making an *acceptable* offer in the current period, then the best (for herself) acceptable offer that she can make, given D's acceptance rule, is $y = d$, in which case her payoff is the same. Hence, her proposal satisfies the OSDP.

Now consider a history at which D has rejected S's offer. Does S's equilibrium strategy of continuing negotiations rather than allowing the rejection to stand satisfy the OSDP? By continuing negotiations, agreement is never reached and S's payoff is $\frac{1-d}{1-\delta}$. If she adopts a one-shot deviation by letting the rejection stand, her payoff is $\frac{1-p-c_S}{1-\delta}$, which is strictly lower. Hence, continuing negotiations satisfies the OSDP.

Finally, we need to verify that D's acceptance rule satisfies the OSDP. If D rejects S's offer and then uses his equilibrium strategy in the future, agreement is never reached and hence D's payoff is $\frac{d}{1-\delta}$. Hence, D's acceptance rule must be to accept any $\frac{y}{1-\delta} \geq \frac{d}{1-\delta}$, or $y \geq d$. Hence, D's acceptance rule satisfies the OSDP. ■

The following proposition characterizes a SPE in which agreement is reached on the status quo in every period, and hence in the first period. It is also easy to construct SPE in which agreement is first reached on the status quo in any later period (2nd, 3rd, 4th, etc.).

PROPOSITION 2. *The following is a SPE:*

1. *S always proposes $y = d$. At any history at which D has rejected S's offer, S continues negotiations (rather than letting the rejection stand).*

2. *D always accept any $y \geq d$.* □

Proof. Does S's offer satisfy the OSDP? By offering $y = d$, agreement is reached in the current period and hence S's payoff is $\frac{1-d}{1-\delta}$. If S adopts a one-shot deviation by making a *different* acceptable offer (i.e., some $y > d$), S's payoff is lower, and hence such a one-shot deviation is unprofitable. If she makes a one-shot deviation by making an *unacceptable* offer (i.e., some $y < d$) in the current period and then reverts to her equilibrium strategy, agreement is reached on $y = d$ in the next period, and hence her payoff is still $\frac{1-d}{1-\delta}$. Hence, her proposal satisfies the OSDP.

Now consider a history at which D has rejected S's offer. Does S's equilibrium strategy of continuing negotiations rather than allowing the rejection to stand satisfy the OSDP? By continuing negotiations, agreement is reached on $y = d$ in the next period and hence S's payoff is $\frac{1-d}{1-\delta}$. If she adopts a one-shot deviation by letting the rejection stand, her payoff is $\frac{1-p-c_S}{1-\delta}$, which is strictly lower. Hence, continuing negotiations satisfies the OSDP.

Finally, we need to verify that D's acceptance rule satisfies the OSDP. If D rejects S's offer and then uses his equilibrium strategy in the future, agreement is reached in the next period on $y = d$ and hence D's payoff is $\frac{d}{1-\delta}$. Hence, D's acceptance rule must be to accept any $\frac{y}{1-\delta} \geq \frac{d}{1-\delta}$, or $y \geq d$. Hence, D's acceptance rule satisfies the OSDP. ∎

2 The Structural Causes of Peace in the Fey & Ramsay Model

In the section titled "Why War Does Not Occur in the Fey & Ramsay Model" of our article, we showed that the Fey & Ramsay framework precludes war even when the terms of peace leave one of the players worse off than going to war; that is, it precludes war even in situations where it *should* occur under complete information. Briefly, because of the *unilateral peace* assumption (each player can unilaterally avoid war by choosing to negotiate), war is easily avoided when one actor imposes the peace negotiations on the other. Combining this with their *behavior-independent peace payoffs* assumption (that the terms of the negotiated settlement cannot be influenced by the crisis behavior of the actors) creates situations in which a satisfied actor can avoid both war and concessions by imposing terms that the dissatisfied opponent finds worse than war. The complete information analysis reveals the potential problems that the twin structural assumptions of unilateral peace and payoff-irrelevant behavior might pose. We now show that they actually provide the core of Fey & Ramsay's main "no-war" result under incomplete information as well.

Consider a sequential-move variant of Fey & Ramsay's basic model. As they note on p. 746, their result "applies equally to decisions made simultaneously and sequentially," and such a model makes it easier to illustrate our points. Assume that there are only two states of the world, ω_1 and ω_2, and there is incomplete information about the realization. D knows the true state of the world, but S does not—she believes that it is ω_1 with probability $1 - q$, and ω_2 with probability q. S can negotiate or stand firm. If she negotiates, the game ends with the peace settlement. If she stands firm, D can negotiate or stand firm in turn. If he negotiates, the game ends with the peace settlement. If he stands firm, the game ends in war. Payoffs depend on the state of the world but not on crisis behavior. Figure 1 shows the extensive-form of this game.

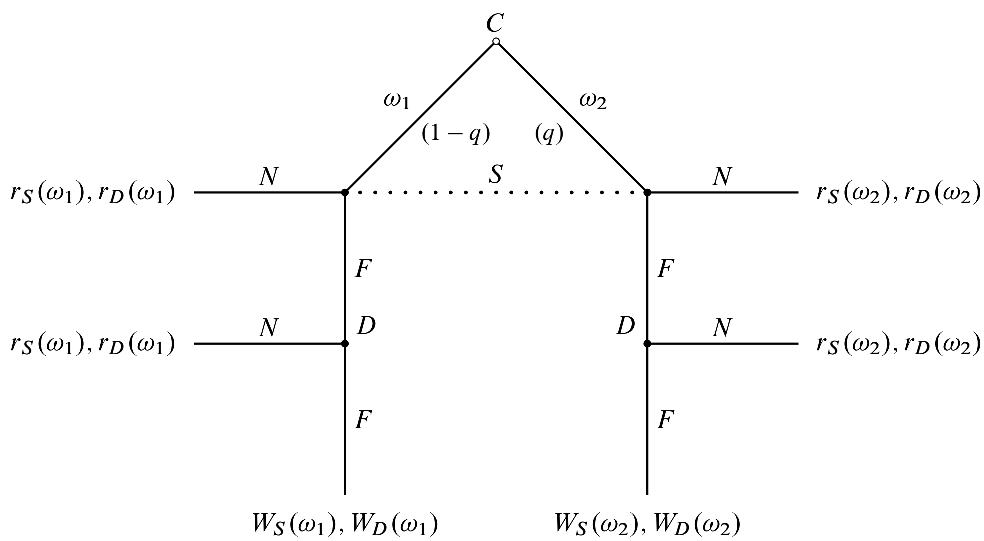

Figure 1: A game that conforms to all Fey & Ramsay assumptions.

This model satisfies all assumptions of Fey & Ramsay's class of models. In particular, (i) war is inefficient, (ii) the negotiation outcomes are efficient, (iii) war only occurs if both sides stand firm, (iv) each player has a strategy that guarantees a negotiated settlement regardless of the other side's strategy, (v) the negotiation payoffs are unique in each state of the world, (vi) there is a common prior, and (vii) each actor's possibility correspondence is partitional. Consequently, the "no-war" theorem holds for this particular model as well, as is established in the following result.

PROPOSITION 3. *The game in Figure 1 has no perfect Bayesian equilibrium (PBE) in which war occurs with positive probability.* □

Proof. Let $\sigma_S(F)$ denote the probability with which S stands firm. Let $\sigma_D(F|\omega_i)$ denote the probability with which D stands firm if the state of the world is ω_i. War can only occur in strategy profiles where $\sigma_S(F) > 0$ and either $\sigma_D(F|\omega_1) > 0$ or $\sigma_D(F|\omega_2) > 0$ (or both). There are four possible cases to consider depending on D's equilibrium strategy.

Suppose first that $\sigma_D(F|\omega_1) > 0$ and $\sigma_D(F|\omega_2) > 0$ in equilibrium. Because his strategy is optimal and he can always induce the negotiated outcome, the fact that he prefers to stand firm implies that $W_D(\omega_1) \geq r_D(\omega_1)$ and $W_D(\omega_2) \geq r_D(\omega_2)$; that is, he is dissatisfied in both states of the world. This implies that S must be satisfied:

$$W_D(\omega_1) \geq r_D(\omega_1) = 1 - r_S(\omega_1) \Rightarrow r_S(\omega_1) \geq 1 - W_D(\omega_1) > W_S(\omega_1)$$
$$W_D(\omega_2) \geq r_D(\omega_2) = 1 - r_S(\omega_2) \Rightarrow r_S(\omega_2) \geq 1 - W_D(\omega_2) > W_S(\omega_2),$$

where the last inequalities in both cases follow from $W_S(\omega_i) + W_D(\omega_i) < 1$. Because S is satisfied in both states of the world, she will negotiate with certainty, $\sigma_S(F) = 0$, and so the equilibrium probability of war is zero.

Suppose now that $\sigma_D(F|\omega_1) = 0$ and $\sigma_D(F|\omega_2) = 0$ in equilibrium. Then the outcome is negotiations regardless of S's strategy, and the equilibrium probability of war is zero.

Suppose now that $\sigma_D(F|\omega_1) = 0$ and $\sigma_D(F|\omega_2) > 0$ in equilibrium. As we now show, in this case S's unique best response is to negotiate with certainty. Consider S's expected payoff given D's strategy. If she chooses to negotiate, she would obtain the state-dependent negotiation payoff, so her expected utility will be:

$$U_S(N) = (1-q)r_S(\omega_1) + qr_S(\omega_2).$$

If, on the other hand, she chooses to stand firm, then she would still obtain the negotiation payoff in ω_1, but she might end up fighting if the true state of the world is ω_2. The probability of war in ω_2 is the probability that D also stands firm there. Hence, S's expected utility from standing firm is:

$$U_S(F) = (1-q)r_S(\omega_1) + q\left[\sigma_D(F|\omega_2)W_S(\omega_2) + (1-\sigma_D(F|\omega_2))r_S(\omega_2)\right].$$

For war to occur with positive probability in this equilibrium, it must be the case that $\sigma_S(F) > 0$, which implies that $U_S(F) \geq U_S(N)$ must hold. Observe now that the $(1-q)r_S(\omega_1)$ term is the same in both expressions. That is, if the true state of the world happens to be ω_1, then S's payoff will be exactly the same wether or not she chooses to stand firm. These terms cancel, which allows us to rewrite the necessary condition for this equilibrium, $U_S(F) \geq U_S(N)$, as follows:

$$\sigma_D(F|\omega_2)W_S(\omega_2) + (1-\sigma_D(F|\omega_2))r_S(\omega_2) \geq r_S(\omega_2). \tag{1}$$

In other words, the optimality of S's action depends entirely on D's behavior in state ω_2. (This allows us to dispense with the probability with which that state of the world occurs, as the q term also cancels out.)

Because D stands firm with positive probability in ω_2, it follows that $W_D(\omega_2) \geq r_D(\omega_2)$. As shown above, this further implies that $W_S(\omega_2) < r_S(\omega_2)$. But this means that:

$$\sigma_D(F|\omega_2)W_S(\omega_2) + (1-\sigma_D(F|\omega_2))r_S(\omega_2) < r_S(\omega_2),$$

which contradicts the necessary condition we derived in (1). Therefore, in state ω_2, S's unique best response is to negotiate with certainty. Because she is indifferent between standing firm and negotiating in state ω_1, this implies that if state ω_2 is expected to occur with positive probability, then S's unique optimal strategy is to negotiate with certainty, so $\sigma_S(F) = 0$. The equilibrium probability of war is zero.

Finally, suppose that $\sigma_D(F|\omega_1) > 0$ and $\sigma_D(F|\omega_2) = 0$ in equilibrium. The proof of this case is symmetric, *mutatis mutandis,* to the one we just made. The equilibrium probability of war is zero. This exhausts all possible strategy profiles in pure as well as mixed strategies. ■

Intuitively, the proof is straightforward and proceeds along the lines of the one case we discuss here. Assume that D strictly prefers the negotiated settlement to war in ω_1 but not in ω_2. That is, $W_D(\omega_1) < r_D(\omega_1)$ and $W_D(\omega_2) > r_D(\omega_2)$, so D is satisfied in ω_1 and dissatisfied in ω_2. With these parameter values, D's equilibrium strategy must be to negotiate in ω_1 and stand firm in ω_2.

Consider now the optimal strategy for the uninformed player. Because S does not know the true state of the world, from her perspective D is potentially dissatisfied so her expected payoff from standing firm is $U_S(F) = (1-q)r_S(\omega_1) + qW_S(\omega_2)$. Her expected payoff from negotiations is $U_S(N) = (1-q)r_S(\omega_1) + qr_S(\omega_2)$. She would strictly prefer to negotiate if, and only if, $U_S(N) > U_S(F)$, which simplifies to $r_S(\omega_2) > W_S(\omega_2)$.

Note that this condition does not depend on her beliefs: it is completely irrelevant whether she is optimistic or not. Her choice is *entirely* determined by her preferences in the state of the world where D is known to be dissatisfied. As we know, our standard assumption of war being costlier than peace allows only one actor to be dissatisfied in any given state of the world. If D is dissatisfied in ω_2, it must be that S is satisfied in that state of the world. This means that $r_S(\omega_2) > W_S(\omega_2)$, which in turn implies that $U_S(N) > U_S(F)$. Therefore, she will choose N, ending the game immediately with the negotiated settlement.

Because S's decision only depends on her preferences in ω_2 but not ω_1, her belief about the true state of the world—and hence any possible optimism—is irrelevant. Unlike the standard model, S does not care about D's private information. To see why this happens, consider first ω_1, where D is known to be satisfied. If this is the true state of the world, then standing firm would cause him to negotiate, and S's payoff would be $r_S(\omega_1)$. If she simply chose to negotiate immediately, her payoff would again be $r_S(\omega_1)$, courtesy of the *behavior-independent payoff assumption*. This holds for any possible state of the world where D is satisfied: if the true state of the world happens to be among them, peace would occur whether S chooses to stand firm or negotiate, and her payoff would be the same either way.

This now means that from her perspective, the only difference between her expected payoff from standing firm and negotiating would arise from the states of the world, like ω_2, where D is dissatisfied—if she stands firm and the true state is among them, he would stand firm too and war would occur. To see why her beliefs about the likelihood of different states in which D is dissatisfied are also irrelevant, observe that in any such state she has to be satisfied:

$$\begin{aligned} W_D(\omega_2) \geq r_D(\omega_2) &\Rightarrow W_D(\omega_2) \geq 1 - r_S(\omega_2) \qquad &&\text{(because } r_D(\omega_2) = 1 - r_S(\omega_2)\text{)} \\ &\Rightarrow r_S(\omega_2) > W_S(\omega_2) &&\text{(because } 1 - W_D(\omega_2) > W_S(\omega_2)\text{).} \end{aligned}$$

Is it easy to see that this holds for *any* state in which D is dissatisfied, so in any such state S's payoff from war is strictly worse than her payoff from peace. Therefore, for any positive probability that D is dissatisfied, her expected payoff from standing firm is *strictly lower* than her expected payoff from negotiating. Consequently, in any equilibrium she strictly prefers to negotiate, and can obtain a settlement, courtesy of the *unilateral peace assumption*.

Both assumptions are necessary and sufficient for the result. Sufficiency follows immediately from the argument above. To see their necessity, suppose players could not impose peace unilaterally but their peace payoffs were still independent of behavior. War would then occur in any state of the world in which D is dissatisfied: S cannot alter the exogenous terms to make him satisfied and cannot simply impose the terms to avoid fighting. Thus, the "no-war" result cannot obtain without the *unilateral peace assumption*.

To see that this result also requires payoff-irrelevant behavior, suppose players could impose peace unilaterally but their peace payoffs depended on how that peace was obtained. We now show that an *arbitrarily small* difference in these payoffs in just *one* state of the world causes the "no-war" result to break down.

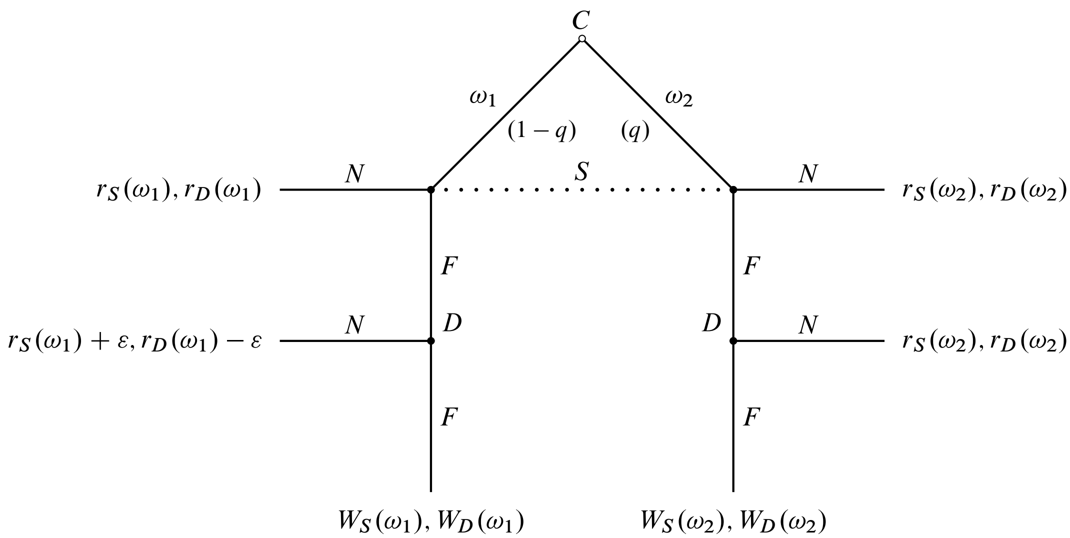

Figure 2: A game that conforms to all Fey & Ramsay assumptions except behavior-independent peace payoffs.

Consider the sequential-move game in Figure 2. It is the same as the game in Figure 1 except that in ω_1, S gains a slight advantage, $\varepsilon > 0$, if she stands firm and D induces the negotiations compared to the settlement in which she negotiates first. We have in mind an advantage so small that D would still strictly prefer to negotiate in ω_1: $0 < \varepsilon < r_D(\omega_1) - W_D(\omega_1)$. Suppose, as before, that D strictly prefers war in ω_2 which, of course, implies that S strictly prefers peace there. Compared to the game in Figure 1, we now allow negotiation payoffs to differ in just one state of the world and the difference is so small that both players are still satisfied there. War would not occur with complete information: $\langle F, N\rangle$ is the unique Nash equilibrium under ω_1, and $\langle N, F\rangle$ and $\langle N, N\rangle$ are Nash equilibria under ω_2. With incomplete information, however, the result changes dramatically, as the following proposition shows.

PROPOSITION 4. *If* $q < k = \frac{\epsilon}{r_S(\omega_2) - W_S(\omega_2) + \epsilon} \in (0, 1)$*, then the game in Figure 2 has a unique perfect Bayesian equilibrium: S stands firm, and D stands firm if dissatisfied (i.e., in ω_2) and negotiates if satisfied (i.e., in ω_1). War occurs with probability* $q > 0$. □

Proof. Given the parameter assumptions, in any PBE D must negotiate in ω_1 and stand firm in ω_2. Therefore, $U_S(F) = (1-q)[r_S(\omega_1) + \epsilon] + qW_S(\omega_2)$, and $U_S(N) = (1 -$

$q)r_S(\omega_1) + qr_S(\omega_2)$. She would strictly prefer to stand firm if, and only if, $U_S(F) > U_S(N)$, which simplifies to $q < k$. ■

It is not just that war occurs with positive probability in some equilibrium, it does so in the *only* PBE that exists when players are mutually optimistic. The necessity of mutual optimism follows because war occurs only when both players stand firm: S does so if her belief that D is satisfied is sufficiently high, and D does so only if he is actually dissatisfied (so knows that his war payoff is higher than the available peace terms). If the true state of the world happens to be ω_2, war would occur even though each player could impose peace unilaterally.

Why does such a minor relaxation of the behavior-independent peace assumption cause a breakdown of the "no-war" result? After all, the negotiation payoffs differ only in a state of the world where D is satisfied and would never stand firm; how can it then lead to war? It can can lead to war because now S is no longer indifferent *how* the peace terms are obtained. She strictly prefers D to induce the negotiations, which he would do whenever he is satisfied. If S is very optimistic, she believes that the risk from standing firm is so small that the minor gain from forcing D to negotiate trumps the probability that he would stand firm as well. It is precisely because the gain accrues in a state of the world where war does not occur that she is willing to run that risk. S's optimism leads her to discount the probability that war will occur when she stands firm. Of course, when war does occur, S has reasons to regret her decision (recall that she is satisfied in ω_2). However, this does not alter the fact that her decision is optimal *ex ante* given her optimistic beliefs. The reader would naturally wonder whether it would be possible to avoid war if she could offer D better terms. This, however, would take us out of Fey & Ramsay's world back to the standard endogenous-offer models.

We conclude that the absence of war in Fey & Ramsay's model has nothing to do with private information, but is a direct consequence of their two structural assumptions. We believe these assumptions to be indefensible from a substantive perspective. However, even if we were to grant them, Fey & Ramsay's model would remain silent about the validity and coherence of the MO explanation. Since nothing in their result actually depends on beliefs, the "no-war" theorem cannot be used to evaluate an explanation for which beliefs are crucial.

3 Conclusion

When we first learned of Fey & Ramsay's result, we were quite startled. Their mathematical logic is impeccable, and the implications seemed to reach far beyond being just a challenge to the MO explanation. We were at a loss when trying to defend the occurrence of war in our existing models against a general result under such apparently reasonable assumptions. Our discussions with Mark and Kris have been enormously beneficial, and we came to realize that we disagree with them about what mutual optimism is. In doing so, we were forced to define very precisely our understanding of the rationalist MO explanation. As we have shown in this article, the dispute is not a matter of semantics, it is a fundamental disagreement about the essence of crisis bargaining. Our position can be summarized succinctly: *the modern rationalist version of the MO explanation specifies the mechanism that leads* *from mutual optimism to fighting: war occurs as a result of strategic behavior intended to overcome the inability to reach an agreement caused by this optimism.* In our construction of this argument, we identified one feature that we believe *any* reasonable model of crisis behavior should possess, and another that is desirable of most.

First, in any model of crisis bargaining, *peace should be a mutual act; the avoidance of war cannot be under the individual control of any actor*. The standard models all assume that peace is only under collective control. While it is true that war is also consensual, it is only so in the very limited sense that both sides must agree to fight for it to occur. The standard model implicitly assumes mutuality of war when it assumes that unconditional surrender to the attacker's demands—what it would take to get him to halt his attack—is worse than fighting for the target. The unilateral peace assumption that Fey & Ramsay make is substantively implausible: if we lived in such a world, any state could preserve a favorable status quo by refusing to fight revisionist opponents dissatisfied with its terms. For instance, the U.S. could have won the second phase of the war in Iraq by the simple expediency of declining to fight the insurgents. (Hussein could have won the first phase by not fighting the invasion… which he tried to no avail.) Thus, any model used to study crisis behavior must specify every peace outcome in such a way that the expected payoffs for both players are no worse than their expected war payoffs when that outcome occurs. This is just a variant of the venerable assumption of anarchy in international relations theory. In the endogenous-offer model this arises naturally from strategic behavior that ensures that at the time of agreement both sides find the terms acceptable.

Second, in most model of crisis bargaining, *crisis behavior should be payoff-relevant*. The standard endogenous-offers model assumes that behavior can affect the terms of the settlement (through the offers), and the signaling models assume that it can affect the peace payoffs (through audience costs, mobilization costs, or through which actor capitulates first) and even the war payoffs (through military preparations). The only exception we can think of that would still be reasonable are "burning bridges" models where an actor can irreversibly commit to fighting when attacked. We have our scruples about those as well, but at any rate Fey & Ramsay's model is not among them.

Fey & Ramsay seem to think that crisis behavior should affect expected payoffs only through changes in beliefs. They write that the "fundamental reason that mutual optimism cannot lead to war is that if both sides are willing to fight, each should infer that they have either underestimated the strength of the opponent or overestimated their own strength. In either case, these inferences lead to a peaceful settlement of the dispute" (738). As we have seen, *that* is not the reason they obtain their "no-war" result, but the claim reveals the fundamental flaw in their approach to crisis bargaining in general. The problem is that these inferences about willingness to fight cannot come for free as they do in their model. There is a price tag attached to any learning that can happen in a crisis, and it usually involves a higher risk of war.

References

Fudenberg, Drew, and Jean Tirole. 1991. *Game Theory*. Cambridge: The M.I.T. Press.