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 $y \geq d$.* $\square$

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. $\blacksquare$

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.

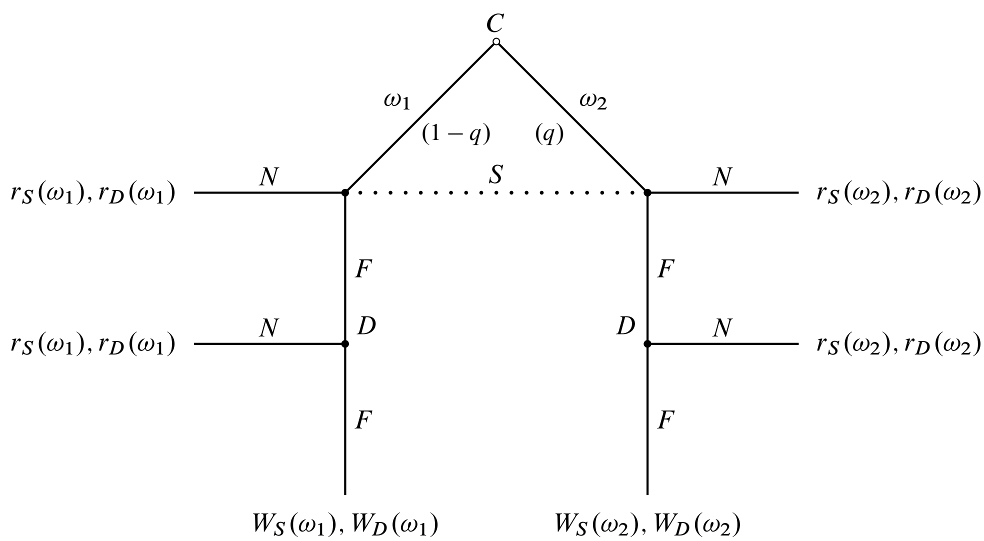

Figure 1: A game that conforms to all Fey & Ramsay assumptions.

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:

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

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.

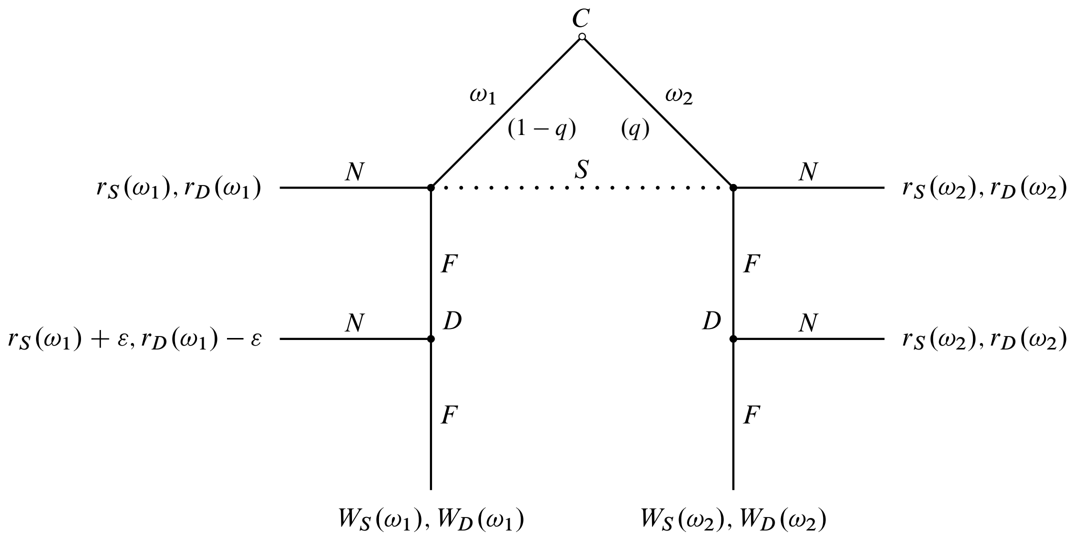

Figure 2: A game that conforms to all Fey & Ramsay assumptions except behavior-independent peace payoffs.

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.