Feigning Weakness

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

April 18, 2007

Abstract. All existing models of crisis bargaining lead to the same conclusion under asymmetric information: strong actors must convince the opponent that they are not bluffing. The usual solution is credible information revelation through some sort of costly signaling. However, these models ignore the problem that when war begins, strong actors may actually benefit from their opponent thinking they are weak. This creates contradictory incentives during the pre-war crisis: actors want to persuade the opponent of their strength to gain a better deal but, should war break out, they would rather have the opponent believe they are weak. I present an ultimatum crisis bargaining model that incorporates this feature and show that in equilibrium the strong actor will pretend to be weak during the bargaining phase with positive probability. The substantive implications are serious. If an actor fails to send a credible signal, then the traditional logic would lead one to infer that this actor is weak. If this results in a more intransigent bargaining stance, the crisis will be more likely to end in war. If the opponent has feigned weakness, then one will be at a serious disadvantage in this war.

*This research was supported by the National Science Foundation (grant SES-0518222). Prepared for presentations at the University of Washington, St. Louis, the University of California, Berkeley, the 2007 meeting of the Project on Polarization and Conflict (Gaillac, France) and the SGIR Pan-European IR Conference (Turin, Italy).

Logic, especially when human beings are involved, is often no more than a way to go wrong with confidence.

David Weber, *At All Costs*, 664

1 Introduction

If there is one conclusion that emerges from the theoretical literature on crisis bargaining, it is that actors must signal credibly their strength if they are to obtain concessions from their opponents. It never pays to pretend to be weak.

Loosely speaking, the logic goes as follows. What one expects from a negotiated settlement is related to what one expects from fighting: an actor whose expected payoff from war is relatively high will demand more at the bargaining table. Because actors are loath to make concessions unless they are convinced that they must, these negotiations take place in the shadow of power: to extract concessions, actors resort to threats to use force. Whether these are tacit or explicit, the result is a crisis. The threats and counter-threats are meant to manipulate the opponents' expectations about the likely outcome of a war and thereby influence the offers they make. The more confident actors are that their prospects in war are good, the more intransigent they will be. The purpose of crisis bargaining, then, is to disabuse them of their optimism.

But how does one persuade the opponent to give in? If I simply stated that I am strong (meaning my expected payoff from fighting is high) and I were believed, then presumably concessions would follow. But why would the opponent believe me? After all, since my claim would never be tested, I could have bluffed and just as easily made it if I were quite weak. But if I am, in fact, weak, then my opponent would not want to make the concessions I demand. Therefore, I must find a way to reveal my strength in a manner that my opponent would find convincing. In general, this means doing something that I could not, or would not, do if I were weak. In other words, in crisis bargaining, the fundamental problem is for a strong actor to send a credible signal. Such a signal must be costly or risky enough to discourage imitation by a weak type.

We have studied many mechanisms for doing so: sinking costs (Fearon 1997), tying hands (Fearon 1994), generating autonomous risks (Schelling 1960), mobilizing military forces (Slantchev 2005), and involving domestic (Schultz 1998) or foreign (Sartori 2002) political actors, among others. All of these enable the strong types to overcome some (but often not all) problems created by the potential for bluffing. All of them rely on the strong type incurring costs or running risks which are too high for the weak types to mimic even if doing so would persuade the opponent that they are strong. The general conclusion is that a strong actor can never profit from the opponent thinking him weak.

This article challenges this conclusion. Imagine a scenario in which war has just broken out and one side believes that the other is relatively weak. It is imperative to defeat the opponent at an acceptable cost and because the enemy if thought to be weak, one's effort will be much smaller than what one would have expended against a strong opponent. This now means that a strong enemy will be able to take advantage of this conservation of effort, and as a result enjoy a much higher expected payoff from war. In other words, it is not difficult to come up with a war scenario in which a strong actor can benefit from the opponent thinking him weak.

This creates contradictory incentives for such an actor in the run-up to the war. On one hand, in order to extract larger concessions he must persuade his opponent that he is strong during the crisis bargaining phase. On the other hand, if negotiations fail, then in order to obtain better chances of winning he would actually prefer that his opponent thinks that he is weak. Thus, he must somehow simultaneously signal strength and weakness. Achieving such a feat of obfuscation in a rationalist framework is not impossible. I modify Fearon's (1995) classic ultimatum game and show that in equilibrium the strong type pretends to be weak with positive probability. Contrary to the traditional conclusion from the crisis bargaining literature, the potential for bluffing is not so much of a problem here as the uncertainty strategically induced by the behavior of the strong type. The finding that sometimes a strong actor can be motivated to prevent credible revelation of information has serious empirical implications: the absence of a costly signal can no longer be interpreted as evidence of weakness. In practical terms this means that even if one's opponent did not opt for a show of force, one still cannot infer that it is safe to be intransigent. Failing to consider the possibility that the opponent has feigned weakness could lead to a war where one would find oneself damned by one's own beliefs.

The article is organized as follows. The next section lays out the crisis bargaining model with a stylized representation of war as simultaneous investment of costly effort. Section 3 shows that a strong actor can always benefit from the opponent thinking him weak in such a war. This establishes the incentive to feign weakness in the bargaining phase. Section 4 constructs an equilibrium of the ultimatum game in which the strong type pretends to be weak. Section 5 elaborates on the logic and substantive implications of the argument, and Section 6 concludes.

2 The Model

Two risk-neutral players, $i \in \{1, 2\}$ are disputing the two-way partition of a continuously divisible benefit represented by the interval $[0, 1]$. An agreement is a pair $(x, 1-x)$, where x is player 1's share and $1-x$ is player 2's share. The set of possible agreements is $\mathcal{X} = \{(x, 1-x) \in \mathcal{R}^2 : x \in [0, 1]\}$. The players have strictly opposed preferences with $u_1(x) = x$ and $u_2(x) = 1 - x$ for all $x \in \mathcal{X}$.[1] Player 1 begins by making a take-it-or-leave-it offer $x \in \mathcal{X}$ that player 2 can either accept or reject. If she accepts, the game ends with the agreement $(x, 1-x)$. If she rejects, the players engage in a costly contest (war). The contest is a simultaneous-move game in which each player chooses a level of effort $m_i \geq 0$ at cost $c_i > 0$. The probability of winning is determined probabilistically by the ratio contest-success function $\pi_i(m_1, m_2) = m_i/(m_1 + m_2)$ if $m_1 + m_2 > 0$ and $\pi_i = 1/2$ otherwise. The winner obtains the entire benefit, so player i's expected payoff from a contest is $\pi_i(m_1, m_2) - m_i/c_i$.

The game has two-sided incomplete information. Each player knows his own cost of effort, c_i, but is unsure about the opponent's cost. Specifically, player 1 believes that player 2 is strong, $\overline{c}_2$, with probability p and weak, $\underline{c}_2 < \overline{c}_2$ with probability $1 - p$. Player 2

[1] For ease of exposition, I will refer to player 1 as "he" and player 2 as "she."

believes that player 1 is strong, $\overline{c}_1$ with probability q and weak, $\underline{c}_1 < \overline{c}_1$, with probability $1 - q$. These beliefs are common knowledge. If the costs of effort are too high even for the strong type (that is, if $\overline{c}_i$ is too small), then war is prohibitively costly and the game will carry no risk of bargaining breakdown. Therefore, assume that the strong type's costs are at least somewhat lower than the costs of his weak opponent.

ASSUMPTION 1. The strong type's costs are not too high: $\overline{c}_j > \sqrt{\underline{c}_i \overline{c}_i}$.

This model has essentially the same crisis bargaining setup as the classic game in Fearon (1995). The crucial difference is it endogenizes the war payoff through the contest game. Since the strategies for the crisis bargaining game would have to form an equilibrium in the contest continuation game, I turn now to the analysis of the contest under uncertainty.

3 The Contest Endgame

Even though the model has two-sided asymmetric information, it turns out that all I need to establish the main claims in this article are the results from contests with complete information and with one-sided asymmetric information. To avoid clutter, I will present here only the results relevant to the full analysis.

3.1 Complete Information

In this case, the costs of effort are common knowledge. Players optimize very simply:

$$\max_{m_i} \left\{ \frac{m_i}{m_i + m_j} - \frac{m_i}{c_i} \right\}.$$

The first-order conditions for an interior equilibrium are:

$$\frac{m_2}{(m_1 + m_2)^2} - \frac{1}{c_1} = 0 \quad \text{and} \quad \frac{m_1}{(m_1 + m_2)^2} - \frac{1}{c_2} = 0.$$

These yield the best responses $m_1^*(m_2) = \sqrt{c_1 m_2} - m_2$ and $m_2^*(m_1) = \sqrt{c_2 m_1} - m_1$. Solving the system of equations then gives us the equilibrium effort levels:

$$m_1^* = c_2 \left(\frac{c_1}{c_1 + c_2} \right)^2 \quad \text{and} \quad m_2^* = c_1 \left(\frac{c_2}{c_1 + c_2} \right)^2.$$

The equilibrium expected probabilities of winning are:

$$\pi_1 = \frac{c_1}{c_1 + c_2} \quad \text{and} \quad \pi_2 = \frac{c_2}{c_1 + c_2} = 1 - \pi_1.$$

The equilibrium expected payoffs are:

$$W_1 = \left(\frac{c_1}{c_1 + c_2} \right)^2 \quad \text{and} \quad W_2 = \left(\frac{c_2}{c_1 + c_2} \right)^2. \tag{1}$$

Observe now that fighting is still inefficient: $W_1 + W_2 < 1 \Leftrightarrow 0 < 2c_1c_2$. Hence, players always have an incentive to negotiate a division of the good instead of fighting to win it all. Moreover, a mutually-acceptable peaceful division always exists. Endogenizing the contest does not change the results from the traditional model under complete information.

3.2 One-Sided Asymmetric Information

Suppose now that one player's cost of effort is common knowledge while his opponent's is privately known only to her. I shall derive the solution assuming that player 1 is the informed player (the other case is symmetric). Player 2 believes that player 1 is strong with probability $\hat{q}$ and weak with probability $1 - \hat{q}$. I am using $\hat{q}$ instead of q to keep in mind the distinction between these beliefs: $\hat{q}$ is the posterior belief that player 2 would form in equilibrium following player 1's initial demand; q is the prior with which she begins the game. In equilibrium, $\hat{q}$ will be common knowledge as well. Player 2's costs, $\overline{c}_2$, are common knowledge.

Since he knows his own cost, player 1 has a simple optimization problem:

$$\max_{m_1} \left\{ \frac{m_1}{m_1 + m_2} - \frac{m_1}{c_1} \right\},$$

for which we already know the solution:

$$m_1(m_2; c_1) = \max \left(\sqrt{c_1 m_2} - m_2, 0 \right). \tag{2}$$

This best response function is sufficient to eliminate some possible contests from consideration as equilibria.

LEMMA 1. *In equilibrium, either both types of the informed player participate in the contest, or only the strong type does.*

Proof. Since I am dealing with the case where player 1 is the informed party, I shall establish the claim for it. The proof for the case where player 2 is the informed party is symmetric. Let $m_2^* \geq 0$ denote player 2's expected equilibrium effort from player 1's perspective and let $m_1^*(c_1) = m_1(m_2^*; c_1)$ denote player 1's type-contingent equilibrium effort. Observe now that there can be no equilibrium in which player 1 makes no effort regardless of type. To establish that, suppose, to the contrary, that $m_1^*(\overline{c}_1) = m_1^*(\underline{c}_1) = 0$ in some equilibrium. Since $m_1(c_1) > 0$ whenever $c_1 > m_2^*$, this implies that $m_2^* \geq \overline{c}_1 > 0$. That is, player 2 is spending a strictly positive amount. However, this cannot be optimal because she can deviate to a lower effort and still win for sure given that neither type of player 1 is expending any effort. Therefore, in any equilibrium at least one type of player 1 must be exerting a strictly positive effort. Observe now that this cannot be the weak type by himself. Suppose, to the contrary, that $m_1^*(\underline{c}_1) > 0$ and $m_1^*(\overline{c}_1) = 0$ in some equilibrium. Since $m_1^*(\underline{c}_1) > 0$ implies that $m_2^* < \underline{c}_1$, it follows from $\underline{c}_1 < \overline{c}_1$ that $m_2^* < \overline{c}_1$, and so $m_1^*(\overline{c}_1) > 0$ as well, a contradiction. □

This result means that there are only two possibilities to consider: either both types of player 1 spend strictly positive effort (skirmish), or only the strong type does (war).[2]

[2]The fanciful names for these equilibria are meant as reminders that contests in which the weak type participate are lower in intensity than conflicts in which only the strong type participates.

3.2.1 The Skirmish Equilibrium

Since player 1's effort depends on his type, let $m_1(\underline{c}_1)$ denote an effort by the weak type, and $m_1(\overline{c}_1)$ denote an effort by the strong type. Because player 2 is unsure about player 1's type, her optimization problem is:

$$\max_{m_2} \left\{ \frac{\hat{q} m_2}{m_1(\overline{c}_1) + m_2} + \frac{(1-\hat{q}) m_2}{m_1(\underline{c}_1) + m_2} - \frac{m_2}{c_2} \right\},$$

for which the first-order condition is:

$$\frac{\hat{q} m_1(\overline{c}_1)}{(m_1(\overline{c}_1) + m_2)^2} + \frac{(1-\hat{q}) m_1(\underline{c}_1)}{(m_1(\underline{c}_1) + m_2)^2} - \frac{1}{c_2} = 0.$$

In equilibrium, player 2's optimal effort would have to be a best response to player 1's type-contingent efforts. This implies that the FOC must hold at his equilibrium effort levels. For convenience, let $\underline{m}_1 = m_1^*(m_2^*; \underline{c}_1)$ denote the equilibrium effort level of the weak type, and $\overline{m}_1 = m_1^*(m_2^*; \overline{c}_1)$ denote the equilibrium effort levels of the strong type. Since (2) implies that $(m_1^*(m_2^*; c_1) + m_2^*)^2 = c_1 m_2^*$, we now obtain:

$$m_2^* = \left(\frac{c_2}{\underline{c}_1 \overline{c}_1} \right) \left[\hat{q} \underline{c}_1 \overline{m}_1 + (1-\hat{q}) \overline{c}_1 \underline{m}_1 \right].$$

Using the definitions of $\underline{m}_1$ and $\overline{m}_1$ from (2), this yields:

$$m_2^* = \underline{c}_1 \overline{c}_1 \left[\frac{f(\hat{q})}{g(\hat{q}; c_2)} \right]^2, \tag{3}$$

where $f(\hat{q}) = \hat{q}\sqrt{\underline{c}_1} + (1-\hat{q})\sqrt{\overline{c}_1} > 0$ and $g(\hat{q}; c_2) = \underline{c}_1 \overline{c}_1 / c_2 + \hat{q} \underline{c}_1 + (1-\hat{q}) \overline{c}_1 > 0$. We can then write the type-contingent expected payoff for player 1 as:

$$W_1(\hat{q}; c_1) = \left(1 - \sqrt{\frac{m_2^*}{c_1}} \right)^2 = \left(1 - \frac{f(\hat{q})}{g(\hat{q}; c_2)} \sqrt{\frac{\underline{c}_1 \overline{c}_1}{c_1}} \right)^2, \tag{4}$$

and the expected payoff for player 2 as:

$$W_2(\hat{q}) = \left(\hat{q} \underline{c}_1 + (1-\hat{q}) \overline{c}_1 \right) \left[\frac{f(\hat{q})}{g(\hat{q}; c_2)} \right]^2. \tag{5}$$

In the skirmishing equilibrium, $\underline{m}_1 > 0$, which means that $m_2^* < \underline{c}_1$ is necessary for this equilibrium to exist. Using (3) then yields the necessary condition in terms of the belief about player 1's type:

$$\hat{q} < \frac{\overline{c}_1 \sqrt{\underline{c}_1}}{c_2 \left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1} \right)} \equiv q_s(c_2). \tag{6}$$

As evident from inspection of this expression, if $\underline{c}_1$ and $\overline{c}_1$ are close enough, the right-hand side will exceed 1, so this condition will not be binding; that is, it will be satisfied regardless of the value of $\hat{q}$. Conversely, the larger the disparity between the two types, the more demanding the condition becomes (because the ceiling on $\hat{q}$ is decreasing in the difference of the two cost terms). Intuitively, this makes sense: if the two costs are relatively close, the expected effort of player 1 from player 2's perspective will not depend much on which type she is actually facing (because their efforts will be quite similar). Conversely, when the costs are very different, so will the efforts be, and if the weak type is to expend positive effort, player 2's effort should not be too great, which in turn requires that she is relatively sure that she is facing the weak type (i.e., the probability of facing the strong type, $\hat{q}$, should be correspondingly low) for otherwise she would increase her own effort to cope with that, making competing unprofitable for the weak type.

3.2.2 The War Equilibrium

In this case, the weak type does not exert any effort in equilibrium, so $\underline{m}_1 = 0$. The strong type's optimal effort is still defined by (2). However, player 2's maximization problem is even simpler:

$$\max_{m_2} \left\{ \frac{\hat{q} m_2}{m_1(\overline{c}_1) + m_2} + (1 - \hat{q}) - \frac{m_2}{c_2} \right\}$$

because she will win outright for any $m_2 > 0$ if her opponent happens to be the weak type. The FOC for this is:

$$\frac{\hat{q} m_1(\overline{c}_1)}{(m_1(\overline{c}_1) + m_2)^2} - \frac{1}{c_2} = 0.$$

Using the definition of $\overline{m}_1$ from (2), which implies that $(\overline{m}_1 + m_2^*)^2 = \overline{c}_1 m_2^*$, we obtain:

$$m_2^* = \overline{c}_1 \left(\frac{\hat{q} c_2}{\overline{c}_1 + \hat{q} c_2} \right)^2. \tag{7}$$

Since the weak type must be willing to exert no effort, it follows that a necessary condition for this equilibrium is $m_2^* \geq \underline{c}_1$, which we obtain by setting $\underline{m}_1 \leq 0$ in (2). This now yields $\hat{q} \geq q_s(c_2)$, that is, exactly the converse of (6) as one would expect. This means that these two cases characterize the complete solution to the one-sided incomplete information problem for all values of $\hat{q}$: if $\hat{q} < q_s(c_2)$, then the skirmish equilibrium obtains; otherwise, the war equilibrium does.

We can now write the expected payoff for the strong type of player 1 (the weak type does not participate, so his payoff is 0) as:

$$W_1(\hat{q}; \overline{c}_1) = \left(\frac{\overline{c}_1}{\overline{c}_1 + \hat{q} c_2} \right)^2, \tag{8}$$

and the expected payoff for player 2 as:

$$W_2(\hat{q}) = 1 - \hat{q} + \hat{q} \left(\frac{\hat{q} c_2}{\overline{c}_1 + \hat{q} c_2} \right)^2. \tag{9}$$

We are now ready to establish some of the most important results for this paper.

3.3 Comparative Statics

As we shall see in the next section, in equilibrium, the uninformed player in the one-sided asymmetric information contests will always be the strong type. Therefore, the comparative statics will be established for that case.

LEMMA 2. *The uninformed player's equilibrium effort is increasing in her belief that her opponent is strong provided her costs of effort satisfy Assumption 1.*

Proof. Assume that player 2 is the strong type, so her costs of effort satisfy Assumption 1. Suppose the skirmish equilibrium, so (3) gives player 2's effort. Since both the numerator and the denominator in (3) are positive, it is easily established that:

$$\operatorname{sign} \frac{\partial m_2^*}{\partial \hat{q}} = \operatorname{sign}\left(\overline{c}_2 - \sqrt{\underline{c}_1 \overline{c}_1}\right) > 0,$$

where the inequality follows from Assumption 1. Suppose now the war equilibrium, so (7) gives player 2's effort. We now have:

$$\frac{\partial m_2^*}{\partial \hat{q}} = \frac{2\hat{q}\overline{c}_1^2 \overline{c}_2^2}{(\overline{c}_1 + \hat{q}\overline{c}_2)^3} > 0,$$

This exhausts the possible equilibria and establishes the claim. □

The next lemma shows that Sun Tzu's principle of feigning weakness can be derived as the result of optimal rational behavior in a contest under uncertainty.

LEMMA 3 (Sun Tzu). *The expected equilibrium payoff of an informed player who participates in the contest decreases in his opponent's belief that he is strong.*

Proof. Assume the skirmish equilibrium. Using the definition of player 1's expected payoff from (4), we get:

$$\frac{\partial W_1(c_1)}{\partial q} = -\left(\frac{\sqrt{c_1} - \sqrt{m_2^*}}{c_1 \sqrt{m_2^*}}\right) \frac{\partial m_2^*}{\partial \hat{q}} < 0.$$

The inequality obtains because in the skirmish equilibrium the bracketed term is positive by (6) (which implies $\underline{c}_1 > m_2^*$) and because m_2^* is increasing in $\hat{q}$ by Lemma 2. Since no step in the derivation depends on the type of player 1, the result holds for both types. The corresponding result in the war equilibrium is immediately evident from inspection of the strong type's expected payoff given by (8). Since this is the only type who participates in the contest in this equilibrium, this exhausts the possibilities and establishes the claim. □

The logic behind the lemma is straightforward. Player 2's equilibrium effort level is increasing in $\hat{q}$: the more pessimistic she is, the higher the effort she will exert. This leads player 1 to compensate by increasing his own effort, leading to an overall decrease in his expected payoff because of the higher costs he incurs in the process. This is not surprising, of course, but it does put player 1 in an interesting situation because he would strictly prefer player 2 to believe he is weak: she will invest less effort, and his expected payoff will increase.

LEMMA 4. *The expected equilibrium payoff of the informed strong type in the one-sided asymmetric information contest is strictly better than his payoff in the contest with complete information.*

Proof. We know from Lemma 3 that this type's expected payoff is strictly decreasing in $\hat{q}$. Therefore, it will be sufficient to show that as $\hat{q} \to 1$, his expected payoff under uncertainty converges to the complete-information payoff. Assume the skirmish equilibrium, so (4) gives the relevant payoff. Since $\lim_{\hat{q}\to 1} f(\hat{q}) = \sqrt{\underline{c}_1}$ and $\lim_{\hat{q}\to 1} g(\hat{q}; c_2) = \underline{c}_1\overline{c}_1/c_2 + \underline{c}_1$, we obtain:

$$\lim_{\hat{q}\to 1} W_1(\hat{q}; \overline{c}_1) = \left[1 - \frac{c_2\sqrt{\underline{c}_1}}{\underline{c}_1\overline{c}_1 + \underline{c}_1 c_2}\sqrt{\frac{\underline{c}_1\overline{c}_1}{\overline{c}_1}}\right]^2 = \left(\frac{\overline{c}_1}{\overline{c}_1 + c_2}\right)^2.$$

The analogous result is immediately obvious from inspection of (8). □

It is worth asking what happens when player 2's cost of effort is so high that Assumption 1 is not satisfied. Note that:

$$c_2 < \sqrt{\underline{c}_1\overline{c}_1} \Rightarrow q_s(c_2) = \frac{\overline{c}_1\sqrt{\underline{c}_1}}{c_2\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)} > \frac{\overline{c}_1\sqrt{\underline{c}_1}}{\sqrt{\underline{c}_1\overline{c}_1}\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)} > 1.$$

In other words, when Assumption 1 is not satisfied, the war equilibrium does not exist because $\hat{q} < q_s(c_2)$ is trivially true. It is then easy to show that m_2^* is *decreasing* in $\hat{q}$, which means that the payoff for the informed strong type of player 1 will *increase* in $\hat{q}$. As before, it converges to his full information payoff as $\hat{q} \to 1$, which implies that this type is doing strictly worse under asymmetric information. In this case, the strong type does have incentives to reveal his private information and the situation reduces to the familiar signaling setup. Because I am interested in investigating behavior when the strong type has contradictory incentives in the bargaining and contest phases, I will maintain Assumption 1 for the rest of this article. With these results in hand, it is now time to turn to player 1's ultimatum.

4 The Crisis Ultimatum

I will now construct an equilibrium in which the weak type makes a low-risk low-value demand and the strong type mixes between pooling on that demand with the weak type and separating to a high-risk high-value demand. In other words, in equilibrium the strong type pretends to be weak with positive probability. The weak type of player 2 accepts both demands, the strong type rejects the low-value demand with positive probability and the high-value demand with certainty. (This is why the low-value demand carries a lower risk of war from player 1's *ex ante* perspective.) The construction will also serve as a proof of the main result for this paper, and since it is illustrative, I will go through it in some detail to clarify the logic.

4.1 The Equilibrium Demands

To derive the optimal offers for player 1, I need to specify the sequentially rational strategies in the continuation games after receiving these offers. Let us begin with the low demand, $\underline{x}$. Since the strong type of player 2 is willing to mix, she must be indifferent between accepting $\underline{x}$ and the contest that would follow if she rejects. Since only the strong type rejects with positive probability, it follows that in this contest player 1 would know for sure that his opponent is strong. This means that the contest will be between a strong player 2 who is uncertain whether player 1 is weak or strong, and player 1 who knows that his opponent is strong. The strong player 2's optimal effort is then given by (3) if the contest admits the skirmish equilibrium and by (7) otherwise. I shall use $W_2(\hat{q};\overline{c}_2)$ to denote the expected payoff with the understanding that this notation refers to the appropriate equilibrium payoff. When it is necessary to be explicit about which equilibrium I am referring to, I shall use $W_2^s(\hat{q};\overline{c}_2)$ for the skirmish equilibrium, and $W_2^w(\hat{q};\overline{c}_2)$ for the war equilibrium. Since player 2 should have no incentive to deviate into the contest, it follows that $1-\underline{x} \geq W_2(\hat{q};\overline{c}_2)$. Since player 1 has no incentive to offer more than the absolute minimum necessary to obtain acceptance, it follows that in equilibrium,

$$\underline{x}(\hat{q}) = 1 - W_2(\hat{q};\overline{c}_2). \tag{10}$$

Clearly, if the strong type of player 2 accepts $\underline{x}(\hat{q})$ with positive probability, then the weak type will accept it for sure.

Turning now to the high demand $\overline{x}$, observe that only the strong type of player 2 is willing to reject this offer and only the strong type of player 1 is supposed to make it in equilibrium. Therefore, the contest that follows rejection is one of complete information between the two strong types. Since the weak type of player 2 must be willing to accept $\overline{x}$, it follows that she should not have incentives to deviate into this contest. Player 1, thinking that he is facing the strong type of player 2, will exert $\overline{m}_1 = \overline{c}_2(\overline{c}_1/(\overline{c}_1 + \overline{c}_2))^2$ effort. The weak type of player 2's optimal deviation is:

$$\operatorname{argmax} \left\{ \frac{m_2}{\overline{m}_1 + m_2} - \frac{m_2}{\underline{c}_2} \right\},$$

whose solution is $m_2' = \max\left(0, \sqrt{\underline{c}_2\overline{m}_1} - \overline{m}_1\right)$. The optimal deviation payoff for the weak type is therefore:

$$W_2' = \begin{cases} \left[1 - \left(\frac{\overline{c}_1}{\overline{c}_1+\overline{c}_2}\right)\sqrt{\frac{\overline{c}_2}{\underline{c}_2}}\right]^2 & \text{if } m_2' > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Note that this expected payoff is strictly worse than what the weak type would have obtained in the full information contest against a strong opponent.[3] The intuition behind this is

[3]This is clearly the case when $m_2' = 0$, so suppose $m_2' > 0$. To see that $W_2' < W_2(\overline{c}_1, \underline{c}_2)$ from (1), take the square roots to obtain:

$$1 - \left(\frac{\overline{c}_1}{\overline{c}_1 + \overline{c}_2}\right)\sqrt{\frac{\overline{c}_2}{\underline{c}_2}} < \frac{\underline{c}_2}{\overline{c}_1 + \underline{c}_2}.$$

Simplification yields $\overline{c}_1\left(\sqrt{\underline{c}_2} - \sqrt{\overline{c}_2}\right) + \sqrt{\underline{c}_2\overline{c}_2}\left(\sqrt{\overline{c}_2} - \sqrt{\underline{c}_2}\right) < 0$. Since $\overline{c}_2 > \underline{c}_2$, this simplifies further to $\sqrt{\underline{c}_2\overline{c}_2} < \overline{c}_1$, which holds by Assumption 1.

simple: when player 1 thinks he is fighting a strong opponent, he exerts more effort than he would have in a war against a weak opponent. Thus, the best a weak opponent can do in such an intense contest is strictly worse than what she could have done had player 1 correctly identified her strength. Since the weak player 2 accepts $\overline{x}$, it follows that $1-\overline{x} \geq W_2'$. Because player 1 has no incentive to offer anything more than that, it follows that in equilibrium:

$$\overline{x} = 1 - W_2'. \tag{11}$$

Clearly, the strong type of player 2 will reject this offer (her expected payoff from the contest is strictly better than the weak type's). Whereas $\overline{x}$ is entirely determined by the exogenous parameters, $\underline{x}(\hat{q})$ is endogenous. I now turn to deriving its value.

4.2 Player 2's Posterior Beliefs

To determine optimal play in the continuation games, it is necessary to specify how player 2 will update her beliefs upon seeing player 1's demand. First, I show that there are some (large) demands a strong player 2 must reject for sure, and other (small) demands that she must accept for sure regardless of her beliefs. For these demands, (10) is not binding. Second, I show that in the range of demands where her behavior is contingent on beliefs, we can always find $\hat{q}$ to satisfy (10).

The first step begins with the surprisingly involved proof of the intuitive claim that the strong player 2's payoff is decreasing in her belief that her opponent is strong.

LEMMA 5. *$W_2(\hat{q};\overline{c}_2)$ is continuous and strictly decreasing in $\hat{q}$.*

To complete the first step, we need to establish the largest and smallest possible payoffs the strong player 2 can expect. When $\hat{q}$ is small enough, the skirmish equilibrium is the one that exists. If $q_s(\overline{c}_2) \geq 1$, then the skirmish equilibrium exists for all admissible values of $\hat{q}$; otherwise, the war equilibrium will exist for all $\hat{q} \in [q_s(\overline{c}_2), 1]$. We now have:

$$\lim_{\hat{q}\to 0} W_2^s(\hat{q};\overline{c}_2) = \left(\frac{\overline{c}_2}{\underline{c}_1 + \overline{c}_2}\right)^2 = \overline{W}_2$$

$$\lim_{\hat{q}\to 1} W_2^s(\hat{q};\overline{c}_2) = \lim_{\hat{q}\to 1} W_2^w(\hat{q};\overline{c}_2) = \left(\frac{\overline{c}_2}{\overline{c}_1 + \overline{c}_2}\right)^2 = \underline{W}_2,$$

where $\underline{W}_2 = W_2(\overline{c}_1, \overline{c}_2) < \overline{W}_2 = W_2(\underline{c}_1, \overline{c}_2)$ denote the full information payoffs, as defined by (1), a strong player 2 can expect against a strong and a weak opponent, respectively. Not surprisingly, as uncertainty disappears, the expected payoffs converge to the full information payoffs with $W_2(0;\overline{c}_2) > W_2(1;\overline{c}_2)$. Note now that Lemma 5 also implies that $\overline{W}_2$ is *the best* payoff player 2 can ever hope to obtain in the contest endgame (a full information contest against a weak opponent), whereas $\underline{W}_2$ is *the worst* payoff she can expect (a full information contest against a strong opponent). Therefore, in equilibrium the strong type must reject any $1-x < \underline{W}_2$ and must accept any $1-x > \overline{W}_2$ regardless of her beliefs. Hence, the only possible demands that involve belief-contingent responses will be $x \in [x_1, x_2]$, where

$$x_1 = 1 - \overline{W}_2 < 1 - \underline{W}_2 = x_2. \tag{12}$$

This completes the first step. The following lemma takes care of the second.

LEMMA 6. *For any $x \in [x_1, x_2]$, there always exists a unique $\hat{q}(x) \in [0, 1]$ that satisfies* (10). *Moreover, $\hat{q}(x)$ is strictly increasing in x.*

We conclude that in any equilibrium, the strong player 2 will accept any $x < x_1$, will reject any $x > x_2$, and will be indifferent for between accepting and rejecting any $x \in [x_1, x_2]$ provided her posterior beliefs are $\hat{q}(x)$ as defined by (10). This suggests that the beliefs that can be used to support the equilibrium can be defined as follows:

DEFINITION 1. Let x_1 and x_2 be defined by (12). On and off the equilibrium path player 2's posterior beliefs are:

- if $x < x_1$, update to believe that player 1 is weak with certainty;
- if $x \in [x_1, x_2]$, update to believe that player 1 is strong with probability $\hat{q}(x)$ as defined by (10);
- if $x > x_2$, update to believe that player 1 is strong with certainty.

By Lemma 6, player 2's posterior belief that player 1 is strong is non-decreasing in the demand. In other words, *these beliefs are quite sensible: the larger the demand, the more likely is it to have been made by a strong player 1.* This equilibrium does not rely on strange off-the-path beliefs. Figure 1 shows how player 2's posterior belief, $\hat{q}$, varies as a function of x using Definition 1. In that figure, $x(\hat{q}_s)$ is the offer such that for all $x \leq x(\hat{q}_s)$, the skirmish equilibrium occurs in the contest endgame, and for all $x > x(\hat{q}_s)$, the war equilibrium occurs. The critical belief, $\hat{q}_s$, is, of course, specified in (6).

4.3 Admissible Priors for Player 1

Now that we have player 2's posterior beliefs as a function of player 1's demand, the next step in constructing the equilibrium is to determine some necessary conditions that must be met for it to exist. Since only the strong player 2 rejects $\overline{x}$ and she does that with certainty, the strong player 1's expected payoff from making the high-value demand is:

$$U_1(\overline{x}; \overline{c}_1) = p\underline{W}_1 + (1 - p)\overline{x}, \tag{13}$$

where $\underline{W}_1 = W_1(\overline{c}_1, \overline{c}_2)$ from (1) is his expected contest payoff against a strong player 2 under complete information. Since $\overline{x}$ is an optimal offer, one necessary condition for this equilibrium is that $U_1(\overline{x}; \overline{c}_1) \geq x_1$, or else the strong player 1 would deviate to an offer that player 2 is sure to accept. This requirement yields:

$$p \leq \frac{\overline{W}_2 - W_2'}{1 - \underline{W}_1 - W_2'} \equiv p_{\max}.$$

Furthermore, it is also necessary that $U_1(\overline{x}; \overline{c}_1) \leq x_2$. If this were not the case, then even if $r_2 = 0$ (and so a smaller offer carries no risk whatsoever) the strong player 1 would still strictly prefer to demand $\overline{x}$ to any $x < x_2$, which means that he would not be willing to mix between the high-value and the low-value demands. This requirement yields:

$$p \geq \frac{\underline{W}_2 - W_2'}{1 - \underline{W}_1 - W_2'} \equiv p_{\min}.$$

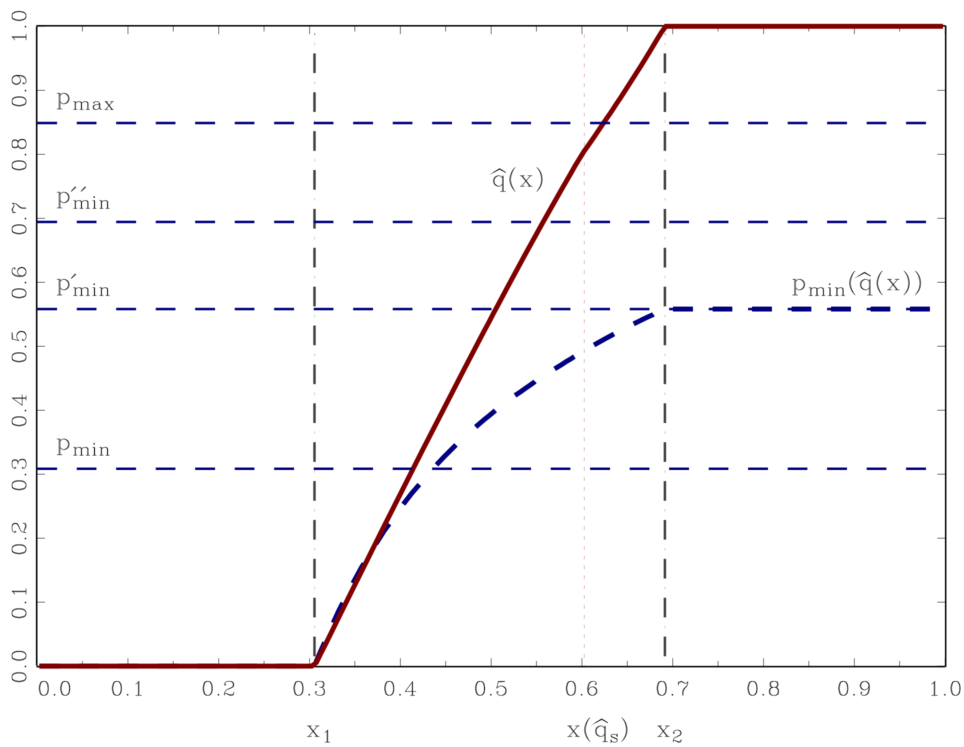

Figure 1: Posterior Beliefs and Restrictions on Priors.

Hence, in any equilibrium of the type I am interested in, it must be the case that $p \in [p_{\min}, p_{\max}]$. Since

$$p_{\max} - p_{\min} = \frac{\overline{W}_2 - \underline{W}_2}{1 - \underline{W}_1 - W_2'} > 0,$$

it is always possible to satisfy the two necessary conditions simultaneously so this interval always exists. Figure 1 shows these two bounds.

Turning to the incentives of the weak player 1, observe that player 1's expected payoff from making the low-value demand is:

$$U_1(\underline{x}(\hat{q}); c_1) = pr_2 W_1(\hat{q}; c_1) + (1 - pr_2)\underline{x}(\hat{q}). \tag{14}$$

Since all $x < x_1$ are accepted by the strong player 2, they will also be accepted by the weak type. Therefore, the risk of war after such offers is zero, making them tempting to the weak player 1. Since no beliefs player 2 may have will prevent her from accepting these offers, it has to be the case that the weak player 1 is unwilling to make them in equilibrium. The best deviation he can make is to largest such offer: if he will not deviate to $x = x_1$ when it is accepted, then he certainly would not deviate to any demand lower than that. Therefore, it is sufficient to establish that $U_1(\underline{x}(\hat{q}); \underline{c}_1) \geq x_1$, which holds whenever:

$$r_2 \leq \frac{\overline{W}_2 - W_2(\hat{q}; \overline{c}_2)}{p\left[1 - W_1(\hat{q}; \underline{c}_1) - W_2(\hat{q}; \overline{c}_2)\right]} \equiv \overline{r}_2(\hat{q}), \tag{15}$$

The probability with which the strong player 2 rejects the low-value offer must be sufficiently small (so the risk of war is low) to prevent the weak player 1 from deviating to the largest surely acceptable offer.[4] Observe now that if p itself is very low, then it can easily be the case that $\overline{r}_2(\hat{q}) > 1$ when $\hat{q}$ large enough. In this case, (15) is not binding for even at $r_2 = 1$ the risk of rejection will be sufficiently low because of p being small itself.

When (15) is satisfied, the weak player 1 will certainly not deviate to any acceptable offer. This implies that the strong player 1 would not deviate either: he would obtain the same payoff as the weak type from making such an offer but his equilibrium expected payoff is strictly larger. Therefore, if the weak player 1 does not like any $x < x_1$, then neither will the strong type.

The next restriction is not necessary for this equilibrium but will prove convenient for constructing it. There is quite a bit of latitude in specifying what r_2 must be in equilibrium. In fact, even restricting ourselves to the type of equilibria I am interested in still admits infinite solutions indexed by the precise value of r_2. All of these equilibria have the same form but differ in the low-value demand being made. Let r_2^* denote the equilibrium mixing probability and $q^* = \hat{q}(\underline{x})$ be the equilibrium posterior belief after the low-value demand. Although $r_2^* \leq \overline{r}_2(q^*)$ must hold, as we have seen this condition may not be particularly demanding if $\overline{r}_2(q^*) > 1$. It would be convenient if we could pin down some r_2^* in an non-arbitrary way. One such value is $r_2^* = \overline{r}_2(q^*)$ but we need to make sure this is a valid probability. For this to be the case, it will suffice to ensure that $\overline{r}_2(\hat{q}) < 1$ for all $\hat{q} \in [0, 1]$, or:

$$p > \frac{\overline{W}_2 - W_2(\hat{q};\overline{c}_2)}{1 - W_1(\hat{q};\underline{c}_1) - W_2(\hat{q};\overline{c}_2)} \equiv p_{\min}(\hat{q}).$$

Since $p_{\min}(\hat{q})$ is increasing in $\hat{q}$, it will be sufficient if it holds at $\hat{q} = 1$. Using $W_2(1;\overline{c}_2) = \underline{W}_2$, this yields:

$$p > \frac{\overline{W}_2 - \underline{W}_2}{1 - W_1(1;\underline{c}_1) - \underline{W}_2} \equiv p'_{\min}. \tag{16}$$

Hence, for any $p > \max\{p_{\min}, p'_{\min}\}$ and $p < p_{\max}$, $r_2(\hat{q}) < 1$ for all $\hat{q}$ and the strong player 1 would not deviate to an offer that player 2 is sure to accept. Since $\hat{q}(x)$ is a function x by (10), it follows that we can express $p_{\min}(\hat{q})$ as a function of x as well, as shown in Figure 1.

To see what this restriction on the prior gives us, note that using $r_2 = r_2(\hat{q})$ (which we can do since $p > p'_{\min}$ guarantees it to be a valid probability regardless of the value of $\hat{q}$) permits the elimination of the loose parameter r_2 from the optimization problem by replacing it with a function of $\hat{q}$ because:

$$pr_2(\hat{q}) = \frac{\overline{W}_2 - W_2(\hat{q};\overline{c}_2)}{1 - W_1(\hat{q};\underline{c}_1) - W_2(\hat{q};\overline{c}_2)} \equiv \alpha(\hat{q}).$$

This turns (14) itself into a function of $\hat{q}$:

$$U_1(\underline{x}(\hat{q});c_1) = \alpha(\hat{q})W_1(\hat{q};c_1) + (1 - \alpha(\hat{q}))(1 - W_2(\hat{q};\overline{c}_2)), \tag{17}$$

where we use the definition of $\underline{x}(\hat{q})$ from (10).

[4]Note that (15) only implicitly defines the maximum mixing probability because in equilibrium, $\underline{x}(\hat{q})$, and so the equilibrium value of $\hat{q}$, will depend on r_2.

4.4 The Optimal Feint

We now wish to find the equilibrium low-value demand $\underline{x}(\hat{q})$, which in turn will pin down the equilibrium probability with which the strong player 1 makes that demand (feigns weakness). Since the strong type of player 2 must be willing to mix upon observing this demand, it follows that $\hat{q}$ must satisfy (10). That is, the strong player 1 must pick the appropriate mixing probability to induce this belief. Let $r_1 \in (0, 1)$ denote the equilibrium probability with which the strong player 1 offers $\underline{x}$. Bayes rule then requires that player 2's posterior belief is:

$$\hat{q} = \frac{qr_1}{qr_1 + 1 - q} < q.$$

Note that upon observing $\underline{x}$, player 2's belief that player 1 is strong will be lower than her prior. This now means that if player 1 wishes to induce the belief $\hat{q}$, then the strong type would have to make the low offer with probability:

$$r_1(\hat{q}) = \frac{\hat{q}(1-q)}{q(1-\hat{q})}. \tag{18}$$

The strong player 1 will be willing to mix only if he is indifferent between the expected payoffs from the two offers:

$$U_1(\underline{x}(\hat{q}); \overline{c}_1) = U_1(\overline{x}; \overline{c}_1),$$

or, using the definitions in (13) and (17), whenever:

$$\alpha(\hat{q})W_1(\hat{q}; \overline{c}_1) + (1 - \alpha(\hat{q}))(1 - W_2(\hat{q}; \overline{c}_2)) = p\underline{W}_1 + (1-p)(1 - W_2'). \tag{19}$$

Observe now that this equation is solely in terms of $\hat{q}$. To establish the existence of a solution, note that since $W_2(0; \overline{c}_2) = \overline{W}_2 \Rightarrow \alpha(0) = 0$, it follows that

$$U_1(\underline{x}(0); \overline{c}_1) = 1 - \overline{W}_2 < U_1(\overline{x}; \overline{c}_1),$$

where the inequality follows from $p < p_{\max}$. Since $U_1(\underline{x}(\hat{q}); \overline{c}_1)$ is continuous in $\hat{q}$, the intermediate value theorem will guarantee a solution to (19) provided that $U_1(\underline{x}(1); \overline{c}_1) > U_1(\overline{x}; \overline{c}_1)$. Observe now that $\alpha(1) = p'_{\min}$, with $W_1(1; \overline{c}_1) = \underline{W}_1$ (by Lemma 2) and $W_2(1; \overline{c}_2) = \underline{W}_2$, which yields:

$$U_1(\underline{x}(1); \overline{c}_1) = p'_{\min}\underline{W}_1 + (1 - p'_{\min})x_2.$$

The sufficient condition for a solution then is:

$$p > \frac{1 - W_2' - \left[p'_{\min}\underline{W}_1 + (1 - p'_{\min})x_2\right]}{1 - \underline{W}_1 - W_2'} \equiv p''_{\min}. \tag{20}$$

Since $p''_{\min} > p'_{\min} \Leftrightarrow \underline{W}_2 > W_2'$, it follows that if (20) is satisfied, then so will (16). Furthermore, $p''_{\min} > p_{\min} \Leftrightarrow p'_{\min}(1 - \underline{W}_1 - \underline{W}_2) > 0$, so whenever (20) is satisfied, $p > p_{\min}$ as well, ensuring that the other necessary condition is met. Finally, since $p''_{\min} < p_{\max} \Leftrightarrow W_1(1; \underline{c}_1) < \underline{W}_1$, it follows that the interval $[p''_{\min}, p_{\max}]$ always exists, as illustrated in Figure 1. Therefore, if $p \in [p''_{\min}, p_{\max}]$, then (19) will have a solution and all

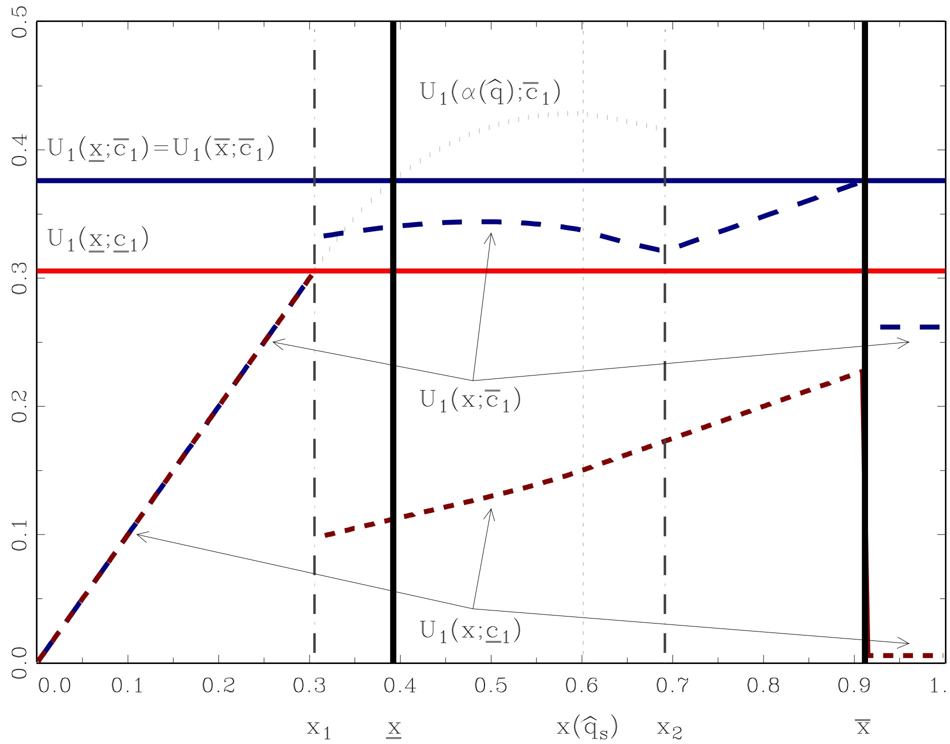

Figure 2: Equilibrium and Deviation Payoffs for Player 1.

necessary conditions for the existence of this equilibrium that we identified before will be met. To ensure that there would be no profitable deviations for the strong player 1, let:

$$\hat{x} = \text{argmax}_x \; pW_1(\hat{q}(x); \overline{c}_1) + (1-p)x \quad \text{subject to } x = 1 - W_2(\hat{q}(x); \overline{c}_2),$$

and observe that there will be no profitable deviation to any $x \in [x_1, x_2]$ as long as $pW_1(\hat{q}(\hat{x}); \overline{c}_1) + (1-p)\hat{x} \leq U_1(\overline{x}; \overline{c}_1)$, or:

$$p \leq \frac{\overline{x} - \hat{x}}{\overline{x} - \hat{x} + W_1(\hat{q}(\hat{x}); \overline{c}_1) - \underline{W}_1} \equiv p'_{\max}. \tag{21}$$

Let $p''_{\max} = \min(p_{\max}, p'_{\max})$. Finally, since $r_1(\hat{q}) < 1$ requires $\hat{q} < q$, it follows that the equilibrium can only be constructed if the solution to (19) is smaller than the prior belief. Although the result holds for cases when the war equilibrium does not exist, $q_s > 1$, I will state it for the case where $q_s < 1$ because it is easier to prove.

PROPOSITION 1. *Assume $q_s < 1$ and $p \in [p''_{min}, p''_{max}]$, and let q^* be the solution to* (19)*. Let $r_1^* = r_1(q^*)$ from* (18)*, $r_2^* = r_2(q^*)$ from* (15)*, with $\underline{x} = \underline{x}(q^*)$ from* (10) *and $\overline{x} > \underline{x}$ from* (11)*. If $q^* < q$, then the following strategies constitute a perfect Bayesian equilibrium of the crisis bargaining game:*

- *If player 1 is weak, he demands $\underline{x}$, and if he is strong, he demands $\underline{x}$ with probability r_1^* and $\overline{x}$ with probability $1 - r_1^*$.*

- *If player 2 is weak, she accepts both demands, and if she is strong, she rejects* $\underline{x}$ *with probability* r_2^* *and rejects* $\overline{x}$ *with certainty. Off the equilibrium path, player 2 accepts* $x \leq x_1$ *regardless of her type, accepts* $x \in (x_1, \underline{x}) \cup (\underline{x}, \overline{x})$ *if she is weak but rejects it if she is strong, and rejects any* $x > \overline{x}$ *regardless of her type.*

In the contest endgame, both choose the equilibrium strategies given their beliefs. On and off the path, beliefs are specified by Definition 1.

Figure 2 shows how the construction given in Proposition 1 works for a particular numerical example (all numbers rounded to two decimal points).[5] The equilibrium demands are $\underline{x} = 0.39$ and $\overline{x} = 0.91$. The risk of war after the low-value demand is $p \times r_2^* = (0.75)(0.31) = 0.23$ and $p = 0.75$ after the high-value demand. The weak player 1 makes the low-value low-risk demand, while the strong player 1 feigns weakness with probability $r_1^* = 0.11$ and makes the separating high-value high-risk demand with probability 0.89. Upon observing the low-value demand, player 2 updates to believe that player 1 is strong with probability $q^* = 0.25$.

Figure 2 plots the optimal deviation payoffs for each type of player 1 given the beliefs in Definition 1 and shown in Figure 1. For example, for any $x \leq x_1$, both types get the same payoffs because all such demands are accepted with certainty. Observe how the construction ensures that (a) the weak type is indifferent between $\underline{x}(q^*)$ and the highest no-risk demand x_1, and (b) the strong type is indifferent between $\overline{x}$ and the highest demand that the weak player 2 would accept. With slight abuse of notation, $U_1(\alpha(\hat{q}); \overline{c}_1)$ denotes what player 1 would have received by demanding $x \in [x_1, x_2]$ if player 2 rejected that with probability $\alpha(\hat{q})$; that is, this is the left-hand side of (19). Of course, $\underline{x}(q^*)$ is the demand at which this intersects the constant function $U_1(\overline{x}; \overline{c}_1)$. What is the benefit of the feint for the strong player 1? If war occurs, his expected payoff is 0.32 after the low-value demand and only 0.20 after the high-value demand. The strong player 1's two options are: (1) demand a lot (great if player 2 concedes) but run a high risk of a war in which player 2 exerts maximum effort, or (2) demand less (not so great if player 2 concedes) but run a lower risk of war and take advantage of player 2's beliefs there.

It is worth emphasizing that the conditions stated in Proposition 1 are sufficient for the feint equilibrium to exist and are actually much more stringent than necessary. For example, $q_s < 1$ can be easily relaxed and the admissible interval on p widened considerably. However, since my purpose is to establish the existence of this equilibrium, and not to investigate the contours of the parameter set that admits that existence, Proposition 1 is sufficient.

One intriguing question arises when we consider the *ex ante* equilibrium payoffs for the strong types of the two players. We already know that $U_1^*(\overline{c}_1) = U_1(\underline{x}; \overline{c}_1) = U_1(\overline{x}; \overline{c}_1) = 0.38$. Recall that because the strong player 2 is willing to mix after the low-value demand, her expected payoffs from accepting it and rejecting it must be equal. This implies $U_2(\underline{x}; \overline{c}_2) = 1 - \underline{x}$. Since she is certain to reject the high-value demand and will know that her opponent is strong, it follows that $U_2(\overline{x}; \overline{c}_2) = W_2(1; \overline{c}_2) = \underline{W}_2$. She expects the low-value demand to be made with probability $y = qr_1^* + (1-q)(1) = 0.33$, and the high-value

[5]The parameters are $\underline{c}_1 = 1$, $\overline{c}_1 = 4$, $\underline{c}_2 = 2$, and $\underline{c}_2 = 5$. The prior beliefs are $p = q = 0.75$. These imply $p_{\max} = 0.85$, $p'_{\max} = 0.82$, $p'_{\min} = 0.56$, and $p''_{\min} = 0.70$. Since $p \in [0.70, 0.82]$, the sufficient conditions are met.

demand to be made with the complementary probability $(1-y) = q(1-r_1^*)+(1-q)(0) = 0.67$. Therefore, her equilibrium expected payoff is $U_2^*(\overline{c}_2) = y(1-\underline{x}) + (1-y)\underline{W}_2 = 0.41$. Observe now that $U_1^*(\overline{c}_1)+U_2^*(\overline{c}_2) = 0.79 < 1$. In other words, there are negotiated deals that are better for both players, even if they are strong, than pursuing the equilibrium strategies. Clearly, any $x \in [U_1^*(\overline{c}_1), 1 - U_2^*(\overline{c}_2)]$ is mutually preferable to the equilibrium payoffs. So the question is, why don't players settle on some $x \in [0.38, 0.59]$ instead of running positive risks of war in equilibrium? After all, such a deal should be acceptable because it makes both players better off even *ex ante*, not just *ex post*, as in the usual bargaining failure puzzle.

To see why no such deal can be worked out, suppose player 1 actually proposes something in this interval, say $x = 0.50$. Player 2 would update to believe that he is strong with probability $\hat{q}(0.5) = 0.55$. In this situation, the strong player 2's expected payoff from war will be, as we should expect by the construction of $\hat{q}(x)$, $W_2(\hat{q}(0.5); \overline{c}_2) = 0.50$. Therefore, if this proposal is made, the strong player 2's payoff will be 0.50. This is better than $U_2^*(\overline{c}_2)$, so she will be quite disposed to such a deal. Turning now to the strong player 1, note that his expected war payoff against a strong player 2 will be $W_1(\hat{q}(0.5); \overline{c}_1) = 0.29$. Since player 1 thinks player 2 is weak with probability $1 - p = 0.25$ and expects her to accept x, his overall expected payoff is $U_1(0.5; \overline{c}_2) = (0.75)(0.29) + (0.25)(0.50) = 0.34$. Unfortunately, this is worse than $U_1^*(\overline{c}_1)$, so the strong type of player 1 is not disposed to make this offer. In general, none of the offers in the interval identified above can make both players better off in expectations. Therefore, either player 1 will be unwilling to make such offers or player 2 will be unwilling to accept them.

5 Discussion

5.1 Traditional Crisis Bargaining

Since the present model is a modification of the original take-it-or-leave-it (TILI) Fearon (1995) model, a comparison with it is the natural place to begin the discussion.

5.1.1 Fearon's Basic Model

Consider Fearon's model under complete information and let π denote the probability that player 1 wins. His expected payoff from war is $F_1 = \pi - k_1$, where $k_1 > 0$ is his cost of fighting. Analogously, player 2's expected payoff from war is then $F_2 = 1-\pi-k_2$, where $k_2 > 0$ is her cost of fighting. Since $F_1 + F_2 = 1-(k_1+k_2) < 1$, it follows immediately that war is inefficient and that there exist peaceful settlements that both players strictly prefer to war. In particular, player 1 would prefer any $x \geq F_1$ and player 2 would prefer any $1-x \geq F_2$. Therefore, $x \in [F_1, 1-F_2]$ is a negotiated settlement that both players prefer to war. Player 2 cannot credibly threaten to reject any offer $x \leq 1-F_2$. Since player 1 has no reason to offer anything more than the minimum that player 2 would accept, it follows that in the unique subgame-perfect equilibrium he demands $x = 1 - F_2$. The game ends peacefully with player 1 extracting the entire bargaining surplus (not surprising considering that the TILI protocol gives him the bargaining leverage).

Turning now to the present model, recall that under complete information, players expect to get W_1 and W_2 as given by (1). Since $W_1 + W_2 < 1$, it follows that war remains inefficient in our new setup. In particular, any $x \in [W_1, 1 - W_2]$ is mutually preferable to fighting. Furthermore, with complete information player 2 cannot credibly threaten to reject any $1 - x \geq W_2$. Since player 1 has no incentive to offer anything less than the minimum player 2 would accept, it follows that in the unique subgame-perfect equilibrium he demands $x = 1 - W_2$. The game ends peacefully with player 1 (again) extracting the entire bargaining surplus.

Under complete information, both models make essentially the same prediction: in the unique equilibrium, player 1 offers player 2 her expected payoff from fighting, and she accepts. Therefore, treating the conflict endgame as simultaneous expenditures of effort does not alter the fundamental logic of the *bargaining puzzle:* since fighting is costly, there always exist deals that both players prefer to war, and with complete information, the TILI crisis game always ends with a peaceful settlement whether the contest is modeled explicitly or not.

Consider now the incomplete information scenario. Player 1 believes that player 2 is strong, $\underline{k}_2$, with probability p and weak, $\overline{k}_2 > \underline{k}_2$ with probability $1 - p$.[6] Uncertainty about player 1's costs is inconsequential in this model because player 2's expected payoff from war is independent of player 1's costs. When she responds, player 2 needs only compare her fighting payoff with what's being offered, and in neither of these calculations do the costs of player 1 enter in any way. Therefore, the game with one-sided asymmetric information will have the same solution as the game with two-sided uncertainty (abstracting from off-the-path beliefs in the latter case).

Let $F_2(k_2) = 1 - \pi - k_2$ denote player 2's expected payoff from war when her costs are k_2. Subgame perfection implies that player 2 with costs k_2 would accept any $1 - x \geq F_2(k_2)$. The weak player 2 accepts $x \leq 1 - F_2(\overline{k}_2) \equiv \overline{x}$, while the strong player 2 accepts any $x \leq 1 - F_2(\underline{k}_2) \equiv \underline{x}$. Because F_2 is decreasing in k_2 and $\underline{k}_2 < \overline{k}_2$, it follows that $\overline{x} > \underline{x}$.

With two types of player 2, there are only two offers that player 1 would ever make in equilibrium: if he demands $\underline{x}$, then player 2 would accept for sure regardless of her type, so $U_1(\underline{x}) = \underline{x}$; if he demands $\overline{x}$, then player 2 would accept if she is weak but reject if she is strong, so $U_1(\overline{x}) = pF_1 + (1 - p)\overline{x}$. There is no reason for player 1 to offer $x < \underline{x}$ because all such offers will be accepted by player 2 regardless of her type, leaving him worse off. All $x \in (\underline{x}, \overline{x})$ are accepted by the weak type but rejected by the strong, so player 1's expected payoff would be $pF_1 + (1 - p)x < U_1(\overline{x})$. Finally, all $x > \overline{x}$ are rejected by player 2 regardless of her type yielding player 1 the war payoff $F_1 < \underline{x}$ where the inequality follows from $k_i > 0$. Therefore, he would prefer to make an offer that is certain to be accepted rather than one that is certain to be rejected.

Player 1 will make the high-risk high-value demand, $\overline{x}$, if $U_1(\overline{x}) \geq U_1(\underline{x})$, which will be satisfied whenever $k_1 \leq [(1 - p)\overline{k}_2 - \underline{k}_2]/p \equiv k_1^*(p)$. Since $k_1 > 0$, a necessary condition for player 1 to risk war in equilibrium if $(1 - p)\overline{k}_2 - \underline{k}_2 > 0$, or if $p < (\overline{k}_2 - \underline{k}_2)/\overline{k}_2 \equiv \overline{p}$. If $p < \overline{p}$, then there always exist values of k_1 that admit the equilibrium in which player

[6] Note that here the costs do not enter inversely as they do in my model. This has no bearing on the results. We only have to keep in mind that whereas both models define a strong player as one who has low costs of war, in my model this is accomplished with *higher* values of c_i while in the traditional setup it is accomplished with *lower* values of k_i. In Fearon's model a strong type has $\underline{k}_i < \overline{k}_i$.

1 demands $\overline{x}$ and risks war. Substantively, this means that player 1 will take the chance on an offer that carries a risk of rejection only if he believes that this risk is not too large. Observe that the uncertainty about player 1's type is completely irrelevant here even though the choice of optimal offer does depend on his costs: if these are sufficiently low, then player 1 makes the risk-return trade-off (Powell 1999); otherwise, he makes the offer that is guaranteed to be acceptable. This equilibrium is a far cry from the one we found for the present model.

Player 1 can never benefit from a feint in the original model under asymmetric information. Even if we assume that player 2 is uncertain about his costs, she will still behave just in the way we derived above. For example, suppose that $\underline{k}_1 < k_1^*(p)$ but $\overline{k}_1 > k_1^*(p)$. In the unique (up to a specification of off-the-path beliefs) perfect Bayesian equilibrium, the strong player 1 makes the high-value high-risk demand $\overline{x}$, and the weak player 1 makes the low-value no-risk demand $\underline{x}$. This is a separating equilibrium, so all information gets revealed by player 1's optimal strategy. Since player 2's best response is not conditional on her beliefs, not only will *any* beliefs support this equilibrium but player 1 has no incentive to manipulate them. Therefore, he has no reason to feign weakness if strong.

5.1.2 Signaling Strength in Crises

Although Fearon's original model was limited to the TILI bargaining protocol, we know that the basic logic extends to an infinite-horizon alternating-offers protocol provided players can initiate war only when responding to offers (Powell 1996), and, under certain conditions, to a setup where they can initiate war after their own offers are rejected (Leventoğlu and Tarar 2005). Although the risk-return trade-off does not necessarily emerge in the latter scenario, it is not difficult to see that there will be no incentives for feints there either. The basic logic is that a dissatisfied player must credibly reveal his type to get the opponent to make an acceptable offer. Whereas in Powell's (1996) model this is impossible, and so the risk of war is irreducible, in Leventoğlu and Tarar's (2005) model the player can use the costs of delay to signal his type without necessarily risking war. At any rate, in both models strong types always want to convince the opponent of their strength and cannot benefit from pretending to be weak. In fact, this conclusion is much more general and is not limited to our standard crisis bargaining models.

Banks (1990) studies a wide class of crisis bargaining games in which one player's type affects the expected payoffs from war for both. Although he conducted his analysis assuming one-sided asymmetric information, his main conclusions extend to the case of two-sided uncertainty. The present model falls into this category, and it is therefore not surprising that the equilibrium solution exhibits the properties identified by Banks (1990) as common to all games in this class. In particular, the strong type of the informed player obtains a better negotiated settlement but runs a higher risk of war than the weak type. In our case, if the crisis is peacefully resolved, the strong player 1 will get either $\underline{x}$ or $\overline{x}$. Since he makes the low-value demand with probability r_1^*, his ex ante's payoff conditional on peace is $r_1^*\underline{x} + (1-r_1^*)\overline{x} > \underline{x}$, which is what the weak type expects. The strong type is also running larger risks: whereas the weak type faces a probability of war pr_2^*, the strong type faces the same probability if he makes the low-value demand and p if he makes the high-value demand, so his ex ante risk is $r_1^* pr_2^* + (1 - r_1^*)p > pr_2^*$. Finally, the strong type does expect a higher overall payoff than the weak type.

Incentive-compatibility requires that in equilibrium no type should be able to benefit by mimicking the other type's strategy instead of playing its own. The strong type's strategy, if it reaches a peaceful resolution of the crisis, is quite tempting to the weak type: he would get at least $\underline{x}$, his equilibrium payoff, or even $\overline{x}$ if his bluff succeeds. As usual, it is the higher risk of war generated by the strong type's strategy that keeps the weak type from attempting to use it. The possible gain of $\overline{x}$ is just not worth the significantly higher risk of war that this demand produces. Hence, the risk inherent in demanding $\overline{x}$ is a signal, which is credible because only the strong player 1 is willing to send it even when it is certain to be believed. Naturally, because the signal is credible, it rationalizes player 2's inference that her opponent is strong upon seeing it.

This means that just as Fearon claimed, war can occur because players have private information and incentives to misrepresent it. The strong player 1 must send a credible signal to obtain a better negotiated deal, which means he must do something the weak player 1 is unwilling to do. In our setup, the only way to discourage the weak player from mimicry is to reveal willingness to run a high risk of war. Therefore, the strong player 1's optimal strategy must necessarily involve a risk of war, leading to bargaining failure.

This now suggests that the strong player 1 should be keenly interested in other ways he can signal credibly his type, preferably ones that do not require him to risk war so badly. There are several costly signaling mechanisms for doing this: sinking costs (Fearon 1997), tying hands (Fearon 1994), involving the opposition (Schultz 1998), or militarizing the crisis (Slantchev 2005). Regardless of the signaling method, the fundamental logic is always the same: the strong type imposes costs upon himself that are high enough to deter the weak type from bluffing by pretending to be strong. For example, if the action is inherently costly regardless of the crisis outcome, then the strong type will choose an action that incurs so much costs that even a credible signal would yield no benefit to the weak type after these costs are subtracted. Only when there is gain in establishing perfect credibility would the strong type permit bluffing, as can sometimes happen when the action also undermines the opponent's war payoff (as it would in a militarized escalation). However, since bluffing is inconsequential when this is the case, the credibility problem is moot (Slantchev 2005). If, on the other hand, the action involves costs that are only incurred if the player fails to carry out his threat (e.g., audience costs), then the risk of war remains irreducible. The credible commitment to fight reveals willingness to run the risk that this will not be enough to get player 2 to back down. It is this willingness that allows the strong type to signal his private information. With all these mechanisms the primary concern of the strong type is to eliminate bluffing by the weak type whenever that would hurt his position (is it usually does). In other words, *the strong type always wants to find a way to reveal its strength,* leaving no incentive to feign weakness at all.

One interesting exception is investigated by Kurizaki (2007) who analyzes a model in which player 1 can decide whether to make his threat public (so whoever backs down incurs audience costs) or keep it private (so backing down is costless). In the private threat equilibrium, the strong player 1 is indifferent between going public and staying private, whereas the weak type always threatens in private. The logic, however, has nothing to do with feigning weakness in Sun Tzu's sense. From the strong type's perspective, the probability of war is the same whether he goes public or stays private, and his payoff from war is independent of the type of threat he makes. The risk of war is the same because player 2 resists a public threat with the same probability she resists a private threat, and a strong type always fights when resisted. To see why player 2 resists both threats with the same probability, note that capitulation is costlier in public, which implies the risk of war she will be willing to run in public to avoid capitulating is much higher than the risk she will be willing to run in private. This means that whereas she has to be sure that her opponent is strong after a public threat, she does not have to be sure after a private threat in order to resist with the same probability. In other words, she will tolerate some bluffing by player 1, but only in private. Therefore, even though the strong type permits bluffing by not fully separating into a public threat, the underlying logic is quite different from what happens in the present model. It is worth noting that whenever he can benefit from eliminating bluffs, the strong type will always do so by going public. There is never any gain in getting player 2 to think that he is weak.

This entire discussion culminates into a simple point: *in all our existing crisis bargaining models the strong players always want to convince their opponent of their type, so all the action is in sending credible signals that the weak types are unwilling or unable to mimic. Strong types never have any incentive to pretend to be weak.* One immediate inference would be that in a crisis in which some costly or risky action is available to one's opponent but he refuses to take it, it must follow that this opponent is weak. The consensus among the models would lead to policy prescriptions based on the logic of this inference. However, if we allow for the possibility that the opponent could benefit from one thinking that he is weak when fighting breaks out, then such a conclusion can no longer be assured. After all, an opponent foregoing the fully revealing action could be a strong type pretending to be weak. Any policy prescription should be cognizant of that possibility.

This further suggests something quite troubling: in the traditional setup, a strong player always wants to find ways to signal credibly. Hence, the persistence of asymmetric information in the presence of means to ensure this can be quite puzzling. Recall that military preparations can be used as a credible signal (Slantchev 2005) and imagine the following scenario. One player conceals his mobilization from his opponent, which causes the opponent to infer that he is weak, which makes her less accommodating than she otherwise would have been, leading to war between the two. This is not a far-fetched scenario. In fact, this is precisely what happened in September 1950 when Mao was trying to deter the American forces from crossing the 38th parallel but hid the extent of China's mobilization (Whiting 1960, Stueck 2002). If we try to use any of the traditional bargaining models, then we must resort to some fancy extra-model arguments to explain why Mao did not send a strong a signal as he could have. The logic of the present model can illuminate this decision nicely.

More generally, *the model suggests that private information can remain private not for lack of means to reveal it but because the only type who can afford to send the credible signal may have no incentive to do so.* It is this intentional and strategic concealment of information that is so troubling for it means that when war is caused by private information, finding the means to reveal it credibly will not be sufficient to avoid the trap; not when players have incentives to obfuscate inferences that would have benefitted them in other situations.

It is perhaps worth noting that feigning weakness is not something one sees in signaling games in general because the incentives required to induce such behavior are quite specific. I am aware of one sole exception to this, an article on jump-bidding in auctions by Hörner and Sahuguet (2007). In their model, player 1 can make an opening bid, which player 2 must cover if she wants to get to the auction. When the cost of the initial bid is small, the unique perfect sequential equilibrium exhibits both bluffing (moderately strong types open with a bid but then lose in the auction) and sandbagging (strong types randomize between making no opening bid, just like the weak types, and making the opening bid and bidding hard in the auction). Sandbagging is similar to feigning weakness: if he behaves like the weak type, the strong player 1 can benefit from player 2's optimistic beliefs during the auction stage because she will bid less than she otherwise would have. That this sort of behavior can emerge in an auction with pre-payments as well as in a crisis bargaining game suggests that it is not an artifact of the model presented here.

5.2 Mutual Optimism and the "Stinging Ice of Reality"

Blainey's (1988) argument that mutual optimism is a major cause of war is very influential. In one form or another, the idea that "war is usually the outcome of a diplomatic crisis which cannot be solved because both sides have conflicting estimates of their bargaining power" (114) can be found in work that usually shares little else. For example, Van Evera (1999) offers "false optimism" as the first hypothesis about factors that increase the probability of war, Jervis (1976) enumerates a variety of cognitive and psychological factors that may lead to overconfidence. On the rationalist side, Wittman (1979) was the first to formalize the notion and Wagner (2000) placed it in an explicit bargaining context that extended past the outbreak of war. Of course, Fearon's (1995) classic explanation is fundamentally about bargaining breakdown caused by inconsistent expectations.

What is mutual optimism? As Blainey puts it, "it is doubtful if there was any war, since 1700, in which initial hopes were low on both sides. On the eve of many wars both nations or alliances expected the campaign to be short and victorious...This recurring optimism is a vital prelude to war" (53). Although Slantchev (2003) has argued that belief in ultimate victory is not necessary in order to overestimate one's bargaining position sufficiently to push one into war, the basic idea that war occurs because opponents "agree to disagree" about their expected payoffs from war remains.

Recently, Fey and Ramsay (2005) have criticized what they call the "mutual optimism hypothesis" on the grounds that rational actors cannot know that war is inefficient and still be willing to forego negotiations for costly fighting. As discussed by Slantchev and Tarar (2007), there are numerous problems with that application of Aumann's (1976) celebrated result to crisis bargaining. My purpose here is to argue that a fundamental assumption implicit in that argument is likely to be violated. As Fey and Ramsay (2005) note, it is not necessary that both sides believe that they are more likely to win than to lose. For mutual optimism to exist, all that is required is that their subjective estimates of the probability of winning are incompatible with the fact that the "true" probabilities must sum to 1 (ignoring ties). In other words, if π_i is the underlying "true" probability that player i will win the war and $\hat{\pi}_i$ is i's estimate of that probability, then $\pi_1 + \pi_2 = 1$ but $\hat{\pi}_1 + \hat{\pi}_2 > 1$. The problem with the mutual optimism hypothesis, they say, is that in equilibrium that player 1 knows that his opponent will be willing to fight only if she believes that he is likely to lose. Knowing this, player 1 will revise his belief about $\hat{\pi}_1$ and as a result, whenever he chooses to fight, his willingness to do so is now a signal to player 2 that she should revise her belief about $\hat{\pi}_2$ in turn. This should unravel the conjectures to the point where players cannot retain their initial incompatible expectations and therefore cannot go to war because of them.

I shall leave aside the strange requirement that negotiated outcomes are independent of crisis behavior (after all, it is precisely to influence these outcomes that actors engage in costly signaling). Let me simply note that Fey and Ramsay's (2005) argument inherently assumes that there is a "true" probability of winning that is exogenous to the strategies players pursue during the war.[7] However, there is no such thing in the present model. Instead, the probability of winning is endogenous to the model given the technology assumed in the contest success function. In other words, the willingness to go to war does not reveal one's subjective belief about the value of an underlying "true" parameter but what one thinks one's opponent will do when war breaks out. As I have demonstrated in this article, there are situations in which one would rather obfuscate that precise inference in order to take advantage of the resulting belief once fighting begins.

Let us see how the endogenous estimates of the probability of winning can be incompatible. Consider the contest endgame with one-sided asymmetric information. In the skirmish equilibrium, the expected probabilities of winning are:

$$\pi_1(\hat{q};\overline{c}_1) = 1 - \frac{f(\hat{q})\sqrt{\underline{c}_1}}{g(\hat{q};c_2)}, \quad \pi_1(\hat{q};\underline{c}_1) = 1 - \frac{f(\hat{q})\sqrt{\overline{c}_1}}{g(\hat{q};c_2)}, \quad \text{and} \quad \pi_2(\hat{q}) = \frac{f(\hat{q})^2}{g(\hat{q};c_2)}$$

Observe now that $\pi_1(\hat{q};\overline{c}_1) + \pi_2(\hat{q}) > 1 \Leftrightarrow f(\hat{q}) > \sqrt{\underline{c}_1}$. That is, the expectations of the strong type of player 1 and his uninformed opponent are incompatible. In other words, this is a case of mutual optimism where players disagree on the probability of winning and this disagreement cannot be reconciled unlike the disagreement with the weak type of player 1, where $\pi_1(\hat{q};\underline{c}_1) + \pi_2(\hat{q}) < 1$. In the war equilibrium, the expected probabilities of winning are:

$$\pi_1(\hat{q};\overline{c}_1) = \frac{\overline{c}_1}{\overline{c}_1 + \hat{q}c_2} \quad \text{and} \quad \pi_2(\hat{q}) = \frac{(1-\hat{q})\overline{c}_1 + \hat{q}c_2}{\overline{c}_1 + \hat{q}c_2}$$

As in the skirmish equilibrium, these expectations are incompatible: it is easily verified that $\pi_1(\hat{q};\overline{c}_1) + \pi_2(\hat{q}) > 1$ for any $\hat{q} < 1$.

Although the original uncertainty is over the costs of fighting, the contest endgame endogenizes the probability of winning (and the entire expected payoff from war), which causes players to be uncertain about who is likely to emerge victorious in war. Disagreement about the probability of winning can be caused by asymmetric information about costs, which translate into uncertainty about the opponent's wartime behavior, which then leads to possibly conflicting estimates about the expected payoffs from war. We need not assume incompatible beliefs over a primitive "true" underlying probability.

What makes matters worse is that one may have incentives to manipulate the opponent's beliefs to induce lower effort. This means that when war breaks out it may very well be the

[7] Of course, it would be wrong to blame Fey and Ramsay (2005) for that particular assumption; it is easily the most prevalent method of modeling war in the literature. Slantchev (2005) provides an explicit critique and the burgeoning research on war as a bargaining process dispenses with it implicitly.

case that both sides are confident of victory but if one (or both) of them bases its confidence on the opponent being mistaken, then one cannot use the opponent's willingness to go to war to revise downward one's own expectations. In other words, when you have gone to great lengths to convince the opponent to be optimistic, you cannot very well use that optimism as evidence that your own assessment is faulty.

It is perhaps worth noting that Blainey himself was very careful to note that optimism is based on conflicting expectations which are at least partially endogenous to one's wartime behavior:

> This optimism does not arise from a mathematical assessment. It does not simply represent one nation's careful calculation that its military and economic capacities exceed those of the potential enemy.... [The rival expectations are] also influenced by relative assessments of each other's ability to attract allies, their ability to finance a war, their internal stability and national morale, their qualities of civilian leadership and their performance in recent wars (53–54).

Unfortunately, sometimes it is impossible to reconcile these incompatible expectations without actual fighting: "The start of war is... marked by conflicting expectations of what that war will be like. War itself then provides the stinging ice of reality" (56).

6 Conclusion

The theoretical literature on crisis bargaining unequivocally states that among the primary concerns actors have during a crisis is their ability to convince their opponents of their resolve to fight unless concessions are forthcoming. Much effort has gone into finding means through which such credible signaling can be effected. Despite the variety of the suggested mechanisms, the fundamental logic remains the same: a strong actor must do something that he would have been unable or unwilling to do if he were weak. Such a signal invariably involves incurring prohibitive costs or running serious risks with the end result of discouraging bluffing. I have argued that sometimes bluffing is not the problem. When a strong actor can benefit from the opponent (incorrectly) believing that he is weak, then the regular signaling logic does not follow: such an actor can feign weakness despite having means of revealing his strength. While it remains to be seen just how widespread such incentives are empirically, the substantive implications are quite serious for one can no longer safely infer that one's opponent is weak when that opponent fails to demonstrate his resolve. In particular, this suggests that if the traditional logic is followed, then one may blunder straight into disaster by being lured into a war through one's own bargaining intransigence induced by one's logical inference.

A Proofs

Proof of Lemma 5. I need to prove the claim for both types of equilibria. As evident from (5) and (9), $W_2(\hat{q};\overline{c}_2)$ is continuous within each equilibrium type, it will suffice to show that it is continuous at q_s, the point at which the equilibrium switch occurs. Letting

$\overline{q}_s = q_s(c_2)$ from (6), we obtain:

$$
\begin{aligned}
W_2^s(\overline{q}_s; \overline{c}_2) = \left(\overline{q}_s \underline{c}_1 + (1-\overline{q}_s)\overline{c}_1\right)\left[\frac{f(\overline{q}_s)}{g(\overline{q}_s;\overline{c}_2)}\right]^2 \\
= 1 - \frac{\underline{c}_1 + \sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2} \\
= 1 - \overline{q}_s + \overline{q}_s\left(\frac{\overline{q}_s\overline{c}_2}{\overline{c}_1 + \overline{q}_s\overline{c}_2}\right)^2 = W_2^w(\overline{q}_s;\overline{c}_2).
\end{aligned}
$$

That is, at the point of the equilibrium switch, the expected payoffs are the same, so the function is continuous. Turning now to monotonicity, assume the war equilibrium, so the expected payoff is defined by (9):

$$\frac{dW_2^w(\hat{q};\overline{c}_2)}{d\hat{q}} = -\frac{\overline{c}_1^2(\overline{c}_1 + 3\hat{q}c_2)}{(\overline{c}_1 + \hat{q}c_2)^3} < 0.$$

Assume the skirmish equilibrium, so the expected payoff is defined by (5):

$$\frac{dW_2^s(\hat{q};\overline{c}_2)}{d\hat{q}} = \frac{g'f^2}{g^2} + \frac{2f\left(\hat{q}\underline{c}_1 + (1-\hat{q})\overline{c}_1\right)}{g^3}\left(f'g - g'f\right) < 0,$$

We now have $f > 0$, $g > 0$, $f' < 0$, and $g' < 0$. Also,

$$f'g - g'f = \sqrt{\underline{c}_1\overline{c}_1}\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)\left(1 - \frac{\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2}\right) > 0,$$

where the inequality follows from Assumption 1 which implies that last bracketed term is positive. I need to show that:

$$gg'f + 2(\hat{q}\underline{c}_1 + (1-\hat{q})\overline{c}_1)(f'g - g'f) < 0.$$

Substituting the definitions of the various short-cuts produces:

$$
\begin{aligned}
2\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)\left(\overline{c}_1 - \hat{q}(\overline{c}_1 - \underline{c}_1)\right)\left(\sqrt{\underline{c}_1\overline{c}_1} - \frac{\underline{c}_1\overline{c}_1}{\overline{c}_2}\right) \\
< (\overline{c}_1 - \underline{c}_1)\left(\sqrt{\overline{c}_1} - \hat{q}\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)\right)\left(\frac{\underline{c}_1\overline{c}_1}{\overline{c}_2} + \overline{c}_1 - \hat{q}(\overline{c}_1 - \underline{c}_1)\right)
\end{aligned}
$$

Using $\overline{c}_1 - \underline{c}_1 = \left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)$, dividing through by $\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1} > 0$, noting that $\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)\left(\sqrt{\overline{c}_1} - \hat{q}\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)\right) = \overline{c}_1 + \sqrt{\underline{c}_1\overline{c}_1} - \hat{q}\left(\overline{c}_1 - \underline{c}_1\right)$, and letting $y = \overline{c}_1 - \hat{q}(\overline{c}_1 - \underline{c}_1) > 0$, we can simplify this to:

$$2y\left(\sqrt{\underline{c}_1\overline{c}_1} - \frac{\underline{c}_1\overline{c}_1}{\overline{c}_2}\right) < \left(\frac{\underline{c}_1\overline{c}_1}{\overline{c}_2} + y\right)\left(\sqrt{\underline{c}_1\overline{c}_1} + y\right).$$

Further simplification yields:

$$y\sqrt{\underline{c}_1\overline{c}_1} - \frac{3y\underline{c}_1\overline{c}_1}{\overline{c}_2} < y^2 + \frac{\underline{c}_1\overline{c}_1\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2},$$

and a division through by $y\sqrt{\underline{c}_1\overline{c}_1} > 0$ finally yields:

$$\frac{3\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2} + \frac{y}{\sqrt{\underline{c}_1\overline{c}_1}} + \frac{\underline{c}_1\overline{c}_1}{y\overline{c}_2} > 1. \tag{22}$$

Taking the derivative of the left-hand side with respect to $\hat{q}$ yields:

$$\frac{y'}{\sqrt{\underline{c}_1\overline{c}_1}} - \frac{y'\underline{c}_1\overline{c}_1}{y^2\overline{c}_2},$$

where $y' = \underline{c}_1 - \overline{c}_1 < 0$. Since the second derivative is:

$$\frac{2(y')^2\underline{c}_1\overline{c}_1}{y^3\overline{c}_2} > 0,$$

it follows that the left-hand side is minimized at:

$$y^* = \sqrt{\frac{\underline{c}_1\overline{c}_1\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2}}. \tag{23}$$

Note now that $y \in [\underline{c}_1, \overline{c}_1]$ for $\hat{q} \in [0, 1]$. There are three possible cases:

1. $y^* \geq \overline{c}_1$ reduces to $\underline{c}_1\sqrt{\underline{c}_1\overline{c}_1} > \overline{c}_1\overline{c}_2$, a contradiction under Assumption 1. Therefore, this case cannot occur.

2. $y^* \leq \underline{c}_1$, which means that $y = \underline{c}_1$ minimizes the left-hand side of (22). This implies that (23) reduces to $\overline{c}_2 > \frac{\overline{c}_1\sqrt{\underline{c}_1\overline{c}_1}}{\underline{c}_1} = \overline{c}_1\sqrt{\frac{\overline{c}_1}{\underline{c}_1}}$. We now have two subcases:

   - $q_s \geq 1$, so (6) is not binding, and $\hat{q} = 1$ does achieve the minimum, and we can rewrite (22) as:

   $$\frac{3\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2} + \sqrt{\frac{\underline{c}_1}{\overline{c}_1}} + \frac{\overline{c}_1}{\overline{c}_2} = \frac{\left(3\sqrt{\underline{c}_1} + \sqrt{\overline{c}_1}\right)\sqrt{\overline{c}_1}}{\overline{c}_2} + \sqrt{\frac{\underline{c}_1}{\overline{c}_1}}$$

   and since $q_s \geq 1 \Rightarrow \overline{c}_2 \leq \frac{\overline{c}_1\sqrt{\underline{c}_1}}{\sqrt{\overline{c}_1}-\sqrt{\underline{c}_1}}$,

   $$\begin{aligned} &> \frac{\left(3\sqrt{\underline{c}_1} + \sqrt{\overline{c}_1}\right)\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)}{\sqrt{\underline{c}_1\overline{c}_1}} + \sqrt{\frac{\underline{c}_1}{\overline{c}_1}} \\ &= \frac{\left(3\sqrt{\underline{c}_1} + \sqrt{\overline{c}_1}\right)\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right) + \underline{c}_1}{\sqrt{\underline{c}_1\overline{c}_1}} \end{aligned}$$

   The inequality in (22) then reduces to:

   $$\begin{aligned} \left(3\sqrt{\underline{c}_1} + \sqrt{\overline{c}_1}\right)\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right) + \underline{c}_1 &> \sqrt{\underline{c}_1\overline{c}_1} \\ 3\sqrt{\underline{c}_1\overline{c}_1} - 3\underline{c}_1 + \overline{c}_1 - \sqrt{\underline{c}_1\overline{c}_1} + \underline{c}_1 &> \sqrt{\underline{c}_1\overline{c}_1} \\ \left(\sqrt{\underline{c}_1\overline{c}_1} - \underline{c}_1\right) + (\overline{c}_1 - \underline{c}_1) &> 0, \end{aligned}$$

   so the claim holds.

- $q_s < 1$, so (6) is binding and the minimum cannot be achieved because $\hat{q} < 1$ is required. This now means that the minimum in the allowed range must be at $\hat{q} = q_s$, which implies $y = \overline{c}_1 - q_s(\overline{c}_1 - \underline{c}_1) = \overline{c}_1 - \overline{c}_1\sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)/\overline{c}_2$. Using this value of y, we rewrite (22) as:

$$\frac{3\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2} + \sqrt{\frac{\overline{c}_1}{\underline{c}_1}} - \frac{\overline{c}_1\sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)}{\overline{c}_2\sqrt{\underline{c}_1\overline{c}_1}} + \frac{\underline{c}_1}{\overline{c}_2 - \sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)}$$
$$= \frac{2\sqrt{\underline{c}_1\overline{c}_1} - \overline{c}_1}{\overline{c}_2} + \sqrt{\frac{\overline{c}_1}{\underline{c}_1}} + \frac{\underline{c}_1}{\overline{c}_2 - \sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)} > 1.$$

Since $y > 0 \Rightarrow \overline{c}_2 > \sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)$, the third term is always positive. If the first term is positive too, then the inequality holds because $\sqrt{\overline{c}_1/\underline{c}_1} > 1$. If the first term is negative, we can rewrite it as:

$$-\frac{\overline{c}_1 - 2\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2} > -\frac{\overline{c}_1 - 2\sqrt{\underline{c}_1\overline{c}_1}}{\sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right)} = -\left(\frac{\sqrt{\overline{c}_1} - 2\sqrt{\underline{c}_1}}{\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}}\right)\sqrt{\frac{\overline{c}_1}{\underline{c}_1}}.$$

Adding this to the second term then gives us:

$$\left(1 - \frac{\sqrt{\overline{c}_1} - 2\sqrt{\underline{c}_1}}{\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}}\right)\sqrt{\frac{\overline{c}_1}{\underline{c}_1}} = \frac{3\sqrt{\underline{c}_1}}{\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}}\sqrt{\frac{\overline{c}_1}{\underline{c}_1}} = \frac{3\sqrt{\overline{c}_1}}{\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}} > 1,$$

so the inequality is satisfied in this case as well.

3. $y^* \in (\underline{c}_1, \overline{c}_1)$, so we can substitute $y = y^*$ in (22) to obtain:

$$\frac{3\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2} + 2\sqrt{\frac{\sqrt{\underline{c}_1\overline{c}_1}}{\overline{c}_2}} > 1 \Rightarrow 3\sqrt{\underline{c}_1\overline{c}_1} + 2\sqrt{\underline{c}_2\sqrt{\underline{c}_1\overline{c}_1}} > \overline{c}_2,$$

which defines the quadratic:

$$-\overline{c}_2 + 2(\underline{c}_1\overline{c}_1)^{\frac{1}{4}}\sqrt{\overline{c}_2} + 3\sqrt{\underline{c}_1\overline{c}_2} > 0,$$

whose discriminant is $16\sqrt{\underline{c}_1\overline{c}_1} > 0$, so there are two solutions, $-(\underline{c}_1\overline{c}_1)^{\frac{1}{4}}$ and $3(\underline{c}_1\overline{c}_1)^{\frac{1}{4}}$. Because the coefficient on the squared term is negative, the inequality holds for all values between the two roots, or, moving back to $\overline{c}_2$, for all

$$\overline{c}_2 \in \left[\sqrt{\underline{c}_1\overline{c}_1}, 9\sqrt{\underline{c}_1\overline{c}_1}\right].$$

Since $\overline{c}_2 > \sqrt{\underline{c}_1\overline{c}_1}$ by Assumption 1, we only need to verify that $\overline{c}_2$ is not so large as to fall out of this range. To do this, I shall show that at the upper bound of the range, $\hat{q} > q_s$, violating the necessary condition for the skirmish equilibrium. Therefore, when the skirmish equilibrium exists, $\overline{c}_2$ will always fall in the range that satisfies the inequality. So let $\overline{c}_2 = 9\sqrt{\underline{c}_1\overline{c}_2}$, and note that we now have:

$$y^* = \frac{\sqrt{\underline{c}_1\overline{c}_1}}{3} \Rightarrow \hat{q} = \frac{\overline{c} - \frac{\sqrt{\underline{c}_1\overline{c}_1}}{3}}{\overline{c}_1 - \underline{c}_1} = \frac{3\overline{c}_1 - \sqrt{\underline{c}_1\overline{c}_1}}{3(\overline{c}_1 - \underline{c}_1)},$$

and

$$q_s(\overline{c}_2) = \frac{\sqrt{\overline{c}_1}}{9\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right)}.$$

Using $\overline{c}_1 - \underline{c}_1 = \left(\sqrt{\overline{c}_1}\right)^2 - \left(\sqrt{\underline{c}_1}\right)^2$, we can simplify $\hat{q} > q_s(\overline{c}_2)$ as:

$$9\left(3\overline{c}_1 - \sqrt{\underline{c}_1\overline{c}_1}\right) > 3\sqrt{\overline{c}_1}\left(\sqrt{\overline{c}_1} + \sqrt{\underline{c}_1}\right) \Leftrightarrow 4\overline{c}_1 - \sqrt{\underline{c}_1\overline{c}_1} > 0,$$

which holds because $\overline{c}_1 > \sqrt{\underline{c}_1\overline{c}_1}$. This means that for all values of $\overline{c}_2$ such that $y^* \in (\underline{c}_1, \overline{c}_1)$ and $\hat{q} < q_s(\overline{c}_2)$, so that the skirmish equilibrium exists, the claim holds.

This exhausts all the possibilities and establishes the claim. □

Proof of Lemma 6. The argument in the text establishes the range of $W_2(\hat{q};\overline{c}_2)$. By Lemma 5, $W_2(\hat{q};\overline{c}_2)$ is continuous, and by the intermediate value theorem this implies that for any $y \in [\underline{W}_2, \overline{W}_2]$, there exists $\hat{q}$ such that $W_2(\hat{q};\overline{c}_2) = y$. Moreover, since Lemma 5 also proves that $W_2(\hat{q};\overline{c}_2)$ is strictly decreasing, this $\hat{q}(y) = W_2^{-1}(y)$ is unique and itself strictly decreasing in y. Letting $x = 1 - y$ and noting that x is decreasing in y establishes the claim. □

CLAIM 1. $W_1(\hat{q};\underline{c}_1) + W_2(\hat{q};\overline{c}_2) < 1.$

Proof. By Lemma 5, $W_2(\hat{q};\overline{c}_2)$ attains a maximum at $W_2(0;\overline{c}_2) = \overline{W}_2$. In the war equilibrium, $W_1(\hat{q};\underline{c}_1) = 0$, and since $\overline{W}_2 < 1$, the claim holds. I now show that in the skirmish equilibrium, $W_1(\hat{q};\underline{c}_1)$ is strictly decreasing in $\hat{q}$:

$$\frac{dW_1(\hat{q};\underline{c}_1)}{d\hat{q}} = -2\sqrt{\overline{c}_1}\left(1 - \frac{f\sqrt{\overline{c}_1}}{g}\right)\left(\frac{f'g - g'f}{g^2}\right) < 0,$$

where the inequality follows from $f'g - g'f > 0$, as established in the proof of Lemma 5 and the first bracketed term also being positive. To see that latter claim:

$$1 - \frac{f\sqrt{\overline{c}_1}}{g} > 0 \Leftrightarrow \frac{\underline{c}_1\overline{c}_1}{\overline{c}_2} > \hat{q}\sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right).$$

In the skirmish equilibrium, $\hat{q} < q_s$ as defined in (6). Hence, it will be sufficient to establish that:

$$\frac{\underline{c}_1\overline{c}_1}{\overline{c}_2} \geq q_s\sqrt{\underline{c}_1}\left(\sqrt{\overline{c}_1} - \sqrt{\underline{c}_1}\right) = \frac{\underline{c}_1\overline{c}_1}{\overline{c}_2},$$

so the claim holds and the payoff is strictly decreasing. This now implies that it attains a maximum at $\hat{q} = 0$. We now have:

$$\lim_{\hat{q}\to 0} W_1(\hat{q};\underline{c}_1) = \left(\frac{\overline{c}_1}{\overline{c}_1 + \overline{c}_2}\right)^2 = \overline{W}_1.$$

That is, the payoff approaches that of the strong type under full information, as defined by (1). But as we have already seen, $\overline{W}_1 + \overline{W}_2 < 1$, which establishes the claim. □

Proof of Proposition 1. Since $p \in [p''_{\min}, p_{\max}]$, it follows that (19) has a solution, denoted by q^*. This implies that $U_1(\underline{x}; \overline{c}_1) = U_1(\overline{x}; \overline{c}_1)$, so the strong player 1 is willing to mix between the two equilibrium demands. In particular, he can mix with $r_1^* < 1$ (where the inequality follows from $q^* < q$), which ensures that Bayes rule produces q^* such that the strong player 2 is indifferent between accepting $\underline{x}$ and rejecting it (that is, (10) is satisfied). This implies that she is willing to mix, and in particular she can do so with probability r_2^*. Player 2's strategies are sequentially rational given her beliefs. In particular, since the strong type is indifferent between acceptance and rejection for any $x \in [x_1, x_2]$ with $x \neq \underline{x}$ given $\hat{q}$, she can reject it with certainty. Because the strong type is indifferent, the weak type strictly prefers to accept (her payoff from rejecting is strictly worse). Furthermore, $\overline{x}$ is the maximum demand that the weak type would accept by construction. Consider now possible deviations by the strong player 1. For any $x \leq x_1$, $p < p_{\max}$ makes such deviation unprofitable. For any $x \in (x_1, x_2)$ with $x \neq \underline{x}$, $p < p'_{\max}$ makes such deviation unprofitable. For any $x \in [x_2, \overline{x})$, $\hat{q} = 1$ and the demand is accepted by the weak player 2 but rejected by the strong. This implies that the expected payoff from such a deviation is strictly increasing in x, making any such deviation unprofitable. Any $x > \overline{x}$ results in certain war and is thus unprofitable. Consider now possible deviations by the weak player 1. No deviation to $x \leq x_1$, which is surely accepted by player 2, is profitable because r_2^* satisfies (15). Consider now $x \in [x_1, \overline{x}]$. It will suffice to show that $x_1 > U_1(\overline{x}; \underline{c}_1) = pW_1(1; \underline{c}_1) + (1 - p)\overline{x}$, or:

$$p > \frac{\overline{W}_2 - W_2'}{1 - W_1(1; \underline{c}_1) - W_2'} \equiv \hat{p}.$$

Since $p > p''_{\min}$, this inequality will hold if $p''_{\min} > \hat{p}$. Noting that $\underline{W}_1 > W_1(1; \underline{c}_1)$, $\underline{W}_2 > W_2'$, $W_1(1; \underline{c}_1) + \underline{W}_2 < 1$, $\underline{W}_1 + W_2' < 1$, and $W_1(1; \underline{c}_1) + W_2' < 1$, this requirement reduces to $W_1(1; \underline{c}_1) + \overline{W}_2 < 1$. Because $q_s < 1$, at $\hat{q} = 1$, the war equilibrium obtains, and in it, $W_1(1; \underline{c}_1) = 0$, so this always holds.Finally, any $x > \overline{x}$ results in a certain war and is thus unprofitable. □

References

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

Banks, Jeffrey S. 1990. "Equilibrium Behavior in Crisis Bargaining Games." *American Journal of Political Science* 34 (August): 599–614.

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.

Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41 (February): 68–90.

Fey, Mark, and Kristopher W. Ramsay. 2005. "Mutual Optimism and War." Department of Political Science, University of Rochester.

Hörner, Johannes, and Nicholas Sahuguet. 2007. "Costly Signalling in Auctions." *The Review of Economic Studies* 74 (1): 173–206.

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

Kurizaki, Shuhei. 2007. "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy." *American Political Science Review* 101 (August).

Leventoğlu, Bahar, and Ahmer Tarar. 2005. "War and Incomplete Information." Manuscript, Departments of Political Science, Duke Univeristy and Texas A&M University.

Powell, Robert. 1996. "Bargaining in the Shadow of Power." *Games and Economic Behavior* 15 (August): 255–89.

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

Sartori, Anne E. 2002. "The Might of the Pen: A Reputational Theory of Communication in International Disputes." *International Organization* 56 (February): 121–149.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge: Harvard University Press.

Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review* 92 (December): 829–44.

Slantchev, Branislav L. 2003. "The Principle of Convergence in Wartime Negotiations." *American Political Science Review* 47 (December): 621–32.

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

Slantchev, Branislav L., and Ahmer Tarar. 2007. "War Is Not a Bet: Mutual Optimism as a Cause of War." Manuscript, Department of Political Science, University of California, San Diego.

Stueck, William. 2002. *Rethinking the Korean War*. Princeton: Princeton University Press.

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

Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44 (July): 469–84.

Whiting, Allen S. 1960. *China Crosses the Yalu: The Decision to Enter the Korean War*. New York: The Macmillan Company.

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