Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict

Mark Fey University of Rochester
Kristopher W. Ramsay Princeton University

We study two different varieties of uncertainty that countries can face in international crises and establish general results about the relationship between these sources of uncertainty and the possibility of peaceful resolution of conflict. Among our results, we show that under some weak conditions, there is no equilibrium of any crisis bargaining game that has voluntary agreements and zero probability of costly war. We also show that while uncertainty about the other side's cost of war may be relatively benign in peace negotiations, uncertainty about the other side's strength in war makes it much more difficult to guarantee peaceful outcomes. Along the way, we are able to assess the degree to which particular modeling assumptions found in the existing literature drive the well-known relationship between uncertainty, the incentive to misrepresent, and costly war. We find that while the theoretical connection between war and uncertainty is quite robust to relaxing many modeling assumptions, whether uncertainty is about costs or the probability of victory remains important.

A central concern in international relations is understanding the obstacles that prevent countries from reaching mutually beneficial settlements in times of conflict. Why do countries engage in lengthy wars when the two sides would be better off if they could settle their dispute without war? As an answer, the conflict literature has long held that uncertainty and the resulting incentive to misrepresent private information together are a central cause of conflict among states (Blainey 1988; Fearon 1995; Waltz 1979; Wittman 1979). One central finding of this theoretical literature is that informational differences regarding aspects of the bargaining process play a key role in determining bargaining and war behavior (Fearon 1995; Powell 1999; Schultz 1998; Slantchev 2003b; Smith and Stam 2006; Wagner 2000). The common theme of this work is that, because of private information, the two sides in a conflict may be unable to identify suitable settlement terms without learning more about each other. But because of incentives to misrepresent their private information, there is no way to credibly signal this information short of war.

One of the limitations of this body of theory, as it exists today, is that it resembles more a collection of theoretical anecdotes than a systematic body of organized reasoning linking uncertainty to the risk of costly war. The formal literature on international conflict contains a wide variety of modeling approaches. For example, one class of models in this literature is characterized by the assumption that actors are uncertain about the costs of war, but these models vary widely in their description of a crisis. Differences arise over the kinds of strategies available to decision makers and the timing of interactions. Some scholars model crisis bargaining as a war of attrition in continuous time (Fearon 1994), others use infinite horizon, alternating-offers bargaining as their model (Powell 1996), and still others use discrete choice crisis bargaining games with entry (Schultz 1999). Alternatively, sometimes a reader finds consistency in game form, but differences in information structures. One such example is illustrated by the ultimatum crisis bargaining model, which has been considered with uncertainty about costs by some (Fearon 1995), while the same bargaining model

Mark Fey is Associate Professor of Political Science, 109E Harkness Hall, University of Rochester, Rochester, NY 14627 (mark.fey@rochester.edu). Kristopher W. Ramsay is Assistant Professor of Politics, 033 Corwin Hall, Princeton University, Princeton, NJ 08544 (kramsay@princeton.edu).

We have benefited from comments received at the University of Rochester's Peter D. Watson Center, Princeton's International Relations Colloquium, the UCSD IR Seminar, and from participants at various professional conferences. We particularly thank Scott Ashworth, Dan Bernhardt, Josh Clinton, Tom Christensen, John Duggan, Jay Lyall, Adam Meirowitz, Andy Moravcsik, Tom Palfrey, Tom Romer, Curt Signorino, Randy Stone, and Johannes Urpelainen. We would also like to thank three anonymous reviewers for helpful comments. Any remaining errors are our own.

American Journal of Political Science, Vol. 55, No. 1, January 2011, Pp. 149–169

 DOI: 10.1111/j.1540-5907.2010.00486.x

has been considered with uncertainty regarding the probability of victory in war by others (Reed 2003). Thus, our collective knowledge regarding the relationship between uncertainty, the incentive to misrepresent, and war is entangled with countless other assumptions about the type of uncertainty, the timing of actions, the bargaining protocol, and various other assumptions made for either practical or substantive reasons.

While this diversity of models is not necessarily a cause for alarm, with some regularity we discover that central conclusions reached from the study of one particular model are overturned when new game forms are considered. For example, Leventoğlu and Tarar (2008) show that small modifications to the timing structure of the alternating-offers bargaining game can generate equilibria without the "risk-reward trade-off," Fey and Ramsay (2007) give a formulation of the mutual optimism problem with private information about the probability of winning a war but show there is no war in any equilibrium, and Slantchev (2003a) points to situations where rational players fight costly wars in a bargaining model with no asymmetric information at all. In this article, we address this problem by taking a systematic approach to analyzing the effects of uncertainty on the risk of war in crisis bargaining situations and focus on the two prominent forms of uncertainty faced by decision makers in a crisis. We first consider the case where decision makers face uncertainty about the costs of war or the level of resolve of their opponent (Fearon 1994; Powell 1996; Schultz 1999). Alternatively, decision makers could have uncertainty and private information about the probability of victory in war (Smith and Stam 2006; Wagner 2000; Wittman 1979). Our analysis shows how these types of uncertainty influence the crisis bargaining problem.[1]

Assessing the effect of uncertainty, even if systematically organized into categories based on its source, is not an easy task. In particular, there is a fundamental problem with using game-theoretic models like those found in the literature to formulate general claims about the role of uncertainty in international conflict. The root of this problem is the fact that the equilibria in any specific game are typically sensitive to the particular details of the game form. That is, it is typically not known how far a result that holds in a specific extensive form generalizes to different extensive forms. This problem is magnified in the study of international conflict. Unlike the study of elections, say, where candidates must first declare their candidacy, then run their campaigns, which are followed by all voters voting simultaneously on election day, there is no "natural" game form for crisis bargaining. In a crisis, who gets to make proposals? Can a state start a war directly after its initial proposal is rejected? Can a state start a war even if its proposal was accepted? Is bargaining restricted to being bilateral or can mediators be used? Where would such a mediator fit into the process? For questions like these there are no clear answers, and this lack of a "natural" game form limits the applicability of results derived from any particular choice of game form.

With this in mind, we employ a methodological approach from economic theory called Bayesian mechanism design. This approach makes it possible to analyze the outcomes of bargaining games even when the precise procedures used by the parties are unclear. Indeed, by using a tool known as the revelation principle along with the constraints implied by a situation where any agreements must be voluntary, we find that there are some general facts about crisis bargaining that can be characterized without reference to a specific game form. That is, we present results that are "game-free" in the sense of Banks (1990) and are not a consequence of particular modeling choices.

This approach accomplishes three things. First, it allows us to cut through the clutter of potentially endless variation in the extensive form of crisis bargaining models and establish the existence of some truly general results for the crisis bargaining problem. Second, this approach allows us to focus on the fundamental role that private information plays in explaining why crisis bargaining can break down even when countries have a common interest in avoiding war. Without such a general analysis, it would be difficult to compare the effects of different types of information structures, as any results would confound the effects of uncertainty and the effects of game form. Third, as emphasized by Fearon (1995), just as important as this uncertainty is the resulting incentive for countries to misrepresent their private information. In the mechanism design framework, this concern is captured by an "incentive-compatibility" requirement that turns out to play a crucial role in our analysis. Indeed, working with these incentive-compatibility constraints shows that the incentive to misrepresent plays an indispensable role in the connection between uncertainty and war. This is most evident in our result that peaceful equilibria can only occur when such incentives are absent.

Our analysis shows that the link between uncertainty and war in the bargaining problem depends in important ways on the kind of uncertainty states face. We first consider the case in which states are uncertain about their opponent's costs of fighting, but there is no uncertainty about the probability of success in war. For this

[1]A recent paper by Wittman (2009) also makes a distinction regarding an important difference in the strategic problem when uncertainty is about the expected outcome of war rather than costs.

commonly assumed type of uncertainty, we show that in any equilibrium of any crisis bargaining game both the expected probability of war and the expected utility of each side are weakly decreasing in the cost of war. We then show that there exist crisis bargaining games in which war never occurs and such games have a payoff structure that is necessarily insensitive to the private information of each country. The second kind of uncertainty we investigate is one in which the two sides' values for war are interdependent. Such interdependence arises naturally in situations where uncertainty is about relative power, but not costs. We show that when these costs are low, private information and the incentive to misrepresent generate a positive probability of war in every crisis bargaining game with voluntary agreements. We then describe how external subsidies from a third country or an international organization can be used to peacefully resolve such a conflict and again offer a characterization of peaceful game forms. This implies that while uncertainty about the other side's cost of war may be relatively benign in peace negotiations, uncertainty about the other side's strength in war makes it much more difficult to guarantee peaceful outcomes.

The approach that we employ to establish these results builds on the mechanism design literature in economics, particularly on the classic paper of Myerson and Satterthwaite (1983) on bilateral bargaining.[2] However, our application to international conflict differs from the standard economic literature in several important respects. First, it is a central facet of international conflict that bargaining occurs between "sovereign states with no system of law enforceable among them, with each state judging its grievances and ambitions according to the dictates of its own reason or desire" (Waltz 2001, 159). That is, a country can, at any time, choose to use force and there is no way a country can commit to not do so. We incorporate this in our formal analysis by requiring that any agreements must be voluntary. Thus, in the international arena, as no enforcement is possible, the type of binding contracts that are implicitly or explicitly assumed in the standard mechanism design literature do not exist. To be precise, a standard assumption in the mechanism design literature is that agents cannot back out of a contract that they previously agreed to. Thus, participation decisions are generally assumed to occur at the interim stage, before the terms of the agreement are finalized. In contrast, because the lack of enforceable contracts is a fundamental fact in international relations, we suppose that either side can veto a potential agreement at any stage.[3]

A second important element of international conflict is that countries can learn about their opponents through the process of negotiation (Pillar 1983; Schelling 1960; Slantchev 2003b). Specifically, we assume that a country's final decision about whether or not to use force should incorporate whatever additional information it has inferred about the opposing country through the negotiation process.[4] A third important difference from the standard mechanism design model of trade in economics is that war is always costly, and it is therefore common knowledge that there is some peaceful settlement available that both sides would strictly prefer. This is unlike the standard model of bilateral trade due to Myerson and Satterthwaite (1983) in which trade is sometimes inefficient.[5]

In the literature on international conflict, the paper that is closest to ours is Banks (1990). This paper uses a game-free approach similar to ours but only considers the case of one-sided incomplete information. Banks shows that with this information structure, all equilibria of all bargaining games are monotonic in the sense that the stronger the informed country is, the higher its payoff from peaceful settlement. We show that these results only partially extend to two-sided uncertainty. In addition, because Banks only considers one-sided incomplete information, his analysis cannot evaluate the different consequences of uncertainty about costs versus uncertainty about relative power.

In the next section, we outline the method we use to generate our results about the relationship between uncertainty and war. The third section details our model, defines what is meant by a crisis bargaining game, and explains how our analysis uses the revelation principle. The fourth section contains our formal results, and the fifth section provides a discussion of the implications of our findings for theories of bargaining, war, and institutional design. The final section concludes.

[2] Several recent surveys of the vast literature on mechanism design in economics include Jackson (2003), Myerson (2008), and Baliga and Sjöström (2008).

[3] However, see Cramton and Palfrey (1995) for a model in which players can opt out of mechanism. This paper differs from our project, though, because we permit actors to opt out *after* the settlement is generated, instead of only beforehand.

[4] In the mechanism design literature, similar ideas have been discussed by Matthews and Postlewaite (1989), and Forges (1999) as an *ex post* participation constraint, and Compte and Jehiel (2009) and Fey and Ramsay (2009) as a veto constraint.

[5] More specifically, in the standard model of bilateral trade, efficient trade always occurs if and only if there is common knowledge of gains from trade.

A Method for Game-Free Analysis

It is perhaps self-evident that in formal models, the results that are obtained depend on the assumptions that are made in formulating the model. While this fact is well understood, all assumptions are not equal and there is little discussion by practitioners of formal theory regarding how sensitive their results can be to the details of the assumed game form. Often the predictions of our models depend crucially on the precise specification of the game we choose. In the game-theoretic literature on bargaining, for example, a number of variations of the standard alternating-offers model due to Rubinstein (1982) have been studied.[6] Taken together, these variations demonstrate that important features of the equilibrium outcome are often highly sensitive to the exact specification of the bargaining procedure. For example, a seemingly minor change in who makes the first offer can have a significant effect on the distributional outcome. Other variations, including when disagreement leads to a costly inside option, when players can opt out of bargaining, or when players cannot commit to forego renegotiating their proposal after their opponent accepts, also can have significant effects on equilibrium outcomes.

The bottom line is that consumers of current theoretical models are necessarily left unsure about how robust existing findings are and how much our theoretical expectations depend on the analysis of a specific game form.[7] Indeed, Powell emphasizes the importance of "the potential sensitivity of informational accounts of war to the bargaining environment—to the sources of uncertainty and the ability to resolve that uncertainty" and calls for "robustness checks for a particular formalization of the bargaining environment" (2004). In order to address these concerns, we make use of a methodological approach from economic theory called Bayesian mechanism design. This approach enables us to analyze the outcomes of bargaining games while leaving the precise procedures used by the parties unspecified. In particular, we may ask, what possible outcomes could occur *for all possible bargaining procedures that could be used?* This seems, at first glance, to be an intractable question. It is not even apparent how one might categorize all the different kinds of bargaining procedures that could be used.

So how is Bayesian mechanism design able to generate game-free results? The answer is that, through the use of a powerful result known as the revelation principle, we are able to reduce the scope of our analysis from the class of all possible Bayesian games to the much smaller class of "incentive-compatible direct mechanisms." In essence, the revelation principle allows us to include the strategic calculations and incentives to misrepresent of the bargainers as part of the direct mechanism. More specifically, the revelation principle states that the outcome of *any* equilibrium of *any* Bayesian game is also the outcome of some equivalent incentive-compatible direct mechanism.[8]

The important implication of this observation is that *any* outcome achievable via *any* equilibrium, under *any* bargaining procedure, must be attainable as the equilibrium outcome of an "information revelation" game in which each player finds it optimal to truthfully reveal his information, given the conjecture that all other players will truthfully reveal their information as well. This is what is referred to as an incentive-compatible direct mechanism. The revelation principle thus implies that if all incentive-compatible direct mechanisms have some property, then every equilibrium of every game form has this property. More importantly for our purposes, if *no* incentive-compatible direct mechanism has some property, then *no* equilibrium of *any* game form has this property.[9] In this way, the revelation principle enables us to use direct mechanisms as a powerful tool for analyzing strategic behavior in a wide variety of settings.

Formalizing the Approach to Crisis Bargaining

As is customary within much of the conflict literature, consider a situation where two states are involved in a dispute which may lead to war. We conceptualize the conflict as occurring over a divisible item of unit size, such as an area of territory or an allocation of resources. The expected payoff to war depends on the probability that a country will win, the utility of victory and defeat, and the inefficiencies present in fighting. We normalize the utility of countries to be 1 for victory in war and 0 for defeat, and we suppose there is a cost $c_i \geq 0$ for country i fighting a war. Thus, if p_i is the probability that country i wins the war, the expected utility for country i of going to war is simply $w_i = p_i - c_i$.

[6] For a survey, see the textbook by Muthoo (1999).

[7] Obviously, there is a trade-off between the ability to make specific predictions and the generality of results. For a discussion of this trade-off, see Banks (1990).

[8] A formal description of and intuition for the revelation principle are given below. See Chwe (1999) and Baron (2000) for other applications of the revelation principle in political science.

[9] We will invoke this version of the revelation principle to prove the impossibility of peaceful resolution of disputes, regardless of the game form.

At the outset, each state has private information about its ability to contest a war. That is, each state has private information regarding its chance of prevailing in a war or the costs of conducting a military campaign. For example, a state could have unique knowledge about its relative value for the issue of dispute (captured by the relative cost of fighting c_i) or the strength and capabilities of its military force (reflected in the probability of victory p_i). Formally, we think of each country as having a variety of possible types, where country i's *type* $t_i \in T_i \subseteq \mathbb{R}$ represents its private information. The countries have a common prior about the joint distribution of types, given by $F(t)$ for type pair $t = (t_1, t_2) \in T = T_1 \times T_2$. In general, we will denote country i's war payoff for a type profile t by $w_i(t)$.[10]

The two countries can attempt to avoid war by resolving their dispute through some peaceful process, which may include direct negotiations, bargaining, threats, mediation by a third party, or some other interaction. Whatever settlement procedure is available in a given instance could then, in principle, be described (abstractly) by a *game form G* which is composed of a set of *actions* for each country, A_1 and A_2, and an *outcome function* $g(a_1, a_2)$ for $a_1 \in A_1$ and $a_2 \in A_2$. It is worth emphasizing that this game form can be anything from a simple strategic form game to an arbitrarily complicated extensive form. We denote a pair of actions (a_1, a_2) by $a \in A = A_1 \times A_2$. Thus, a game form defines the actions available to the countries (e.g., what negotiation tactic to use, etc.) and how those actions interact to determine outcomes.

A *crisis bargaining game* is a game form in which the final outcome is either a peaceful settlement or an impasse that leads to war.[11] Thus, we can decompose the outcome function $g(a)$ of a crisis bargaining game into two parts: the probability of war $\pi^g(a)$, and, in the case of a settlement, the value of the settlement to country 1, $v^g(a)$. We will assume that any potential settlement is efficient and therefore the value of the settlement to country 2 is given by $1 - v^g(a)$. We sometimes write $v_i^g(a)$ for the value of the settlement to country i. With this structure, it is easy to see that the payoff to country i of an action profile a is given by

$$u_i(a, t) = \pi^g(a) \cdot w_i(t) + (1 - \pi^g(a))v_i^g(a). \quad (1)$$

In words, the payoff to country i is the probability of a war times the payoff of war plus the probability of a peaceful settlement times the value of a peaceful settlement for an action profile a. As in the canonical take-it-or-leave-it bargaining game studied by Fearon (1995), the private information of the two sides only affects payoffs in the case of war. That is, the terms of a peaceful settlement are determined by the actions of the two sides in the bargaining game and, in the case of peace, it is only these terms that determine the countries' payoffs and not their private information about their war-fighting ability.

Of course, a country's private information can affect its behavior at the bargaining table. Formally, the actions taken in the game can depend on the countries' types, and we reflect this fact by defining a strategy for country i by a function $s_i : T_i \to A_i$. The set of all possible strategies for state i is S_i and we let $(s_1, s_2) = s \in S = S_1 \times S_2$. The equilibrium concept we employ is Bayesian-Nash equilibrium. In particular, a strategy profile s^* is a Bayesian-Nash equilibrium if each type of each player is playing a best response to the strategies used by the other players. For a given equilibrium s^* of G, define $U_i(t_i)$ to be the expected utility of this equilibrium for a type t_i of country i. Analytically, we let

$$U_i(t_i) = \int_{T_j} u_i(s^*(t_i, t_j), t)\, dF(t_j \mid t_i), \quad (2)$$

where $i \neq j$ and $u_i(a, t)$ is given by equation (1).[12]

In the remainder of this section, we formalize the method of game-free analysis and discuss how to apply this method to crisis bargaining games. We begin by linking the game form and the information structure described above in the following way. Fix an equilibrium s^* to the overall game. For a type pair $t = (t_1, t_2)$, this equilibrium generates an equilibrium probability of war $\pi(t) = \pi^g(s^*(t))$ and an equilibrium value of settlement to country i, $v_i(t) = v_i^g(s^*(t))$. As peaceful settlements are efficient, the equilibrium value of settlement to country 2 satisfies $v_2(t) = 1 - v_1(t)$. The functions $\pi(t)$ and $v_i(t)$ form what is called an *equivalent direct mechanism*, which can be understood as nothing more than a new game in which each country's available actions are limited to reporting one of its possible types and outcomes are determined using $\pi(t)$ and $v(t)$. If it is an equilibrium for all types to "tell the truth" by sending a report equal to their type in this game, then we say that the direct mechanism is *incentive-compatible*.

With these definitions it is now possible to formally state the revelation principle.

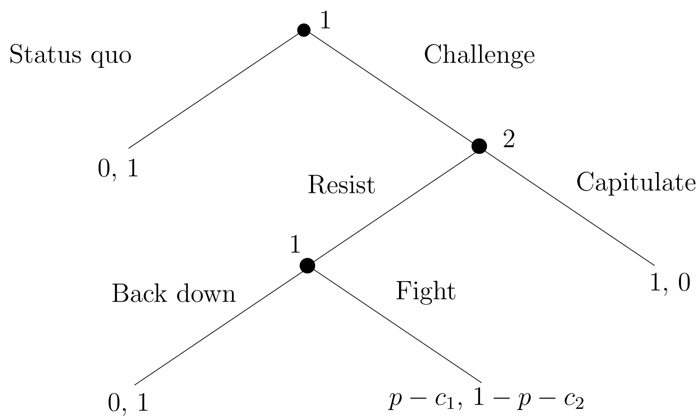
Result 1 (Myerson, 1979). *If s^* is a Bayesian-Nash equilibrium of the crisis bargaining game form G, then there*

[10]For simplicity we assume that $F(t)$ has full support. That is, the support of $F(t)$ is equal to T.

[11]Here a settlement can involve either a peaceful redistribution of whatever is being contested or continuation of the status quo.

[12]Here, and throughout the text, we suppose that all the functions we consider are measurable and thus we take all integrals to be Lebesgue integrals.

FIGURE 1 Example of the Revelation Principle: A Simple Crisis Game



exists an incentive-compatible direct mechanism yielding the same outcome.

To understand the intuition of the revelation principle, consider the simple extensive form game in Figure 1. This game has two players, country 1 and country 2, and is similar in nature to the crisis subgame in Bueno de Mesquita and Lalman (1992). To start, country 1 decides whether to accept the status quo or make a challenge. If country 1 challenges, then country 2 can capitulate, giving in to country 1's demand, or resist. If country 2 resists, country 1 decides whether to repeal its challenge and back down, or to fight—leading to war. This describes the outcome function g of this game. If a war occurs, then country 1 wins with probability $p_1 = p$ and country 2 wins with probability $p_2 = 1 - p$. Each country also pays a cost of war. In this example, there is two-sided uncertainty: each country's cost of war is private information and can take on one of two values, c_l or c_h, with $c_h > c_l > 0$. For each country, the prior probability that $c_i = c_h$ is given by π. For simplicity, we focus on the case in which c_l is small and c_h is large. Specifically, we assume that $c_l < \min\{p, 1 - p\}$ and $c_h > \max\{(1 - p(1 - \pi))/(1 - \pi), p\}$.

Now consider the following pair of strategies for the two countries. For country 1, both types make a challenge and if country 2 resists, then the high-cost type backs down and the low-cost type fights. For country 2, after country 1 challenges, the high-cost type capitulates and the low-cost type resists. Note that as described above, these strategies describe what actions each type of each country takes. It is easy to verify that under the assumptions on the parameters given above, this pair of strategies is an equilibrium of this game. Thus we denote this pair of strategies by $(s_1^*(t_1), s_2^*(t_2))$.

We now construct the equivalent direct mechanism for this equilibrium. As each country has two types, a high-cost type and a low-cost type, in the direct mechanism each country simultaneously chooses to report one of these two types, which we denote H and L. As the direct mechanism is just a strategic form game formed by these simultaneous choices, the only other thing we need to specify is what outcomes are assigned to these choices. This is accomplished via an outcome function for the direct mechanism equal to $g(s^*(t_1, t_2))$, for all the possible reported type profiles (t_1, t_2). In order to see how this works in our example, consider the profile (H, H) in which both countries report they are the high-cost type. The direct mechanism assigns the outcome "Capitulation" to this combination. Why? Because in the equilibrium of this example the high-cost type of country 1 challenges the status quo and the high-cost type of country 2 capitulates to this challenge. In other words, the direct mechanism assigns the outcome that would occur if the specified pair of types plays the equilibrium strategy in the original game. In the same way, the direct mechanism assigns the outcome "War" to the profile (L, L), "Capitulation" to the profile (L, H) and "Back Down" to the profile (H, L). Using the payoffs from the game tree, we can complete the direct mechanism as given in Figure 2. In this case, because each side has two possible types, the direct mechanism is just a 2×2 strategic form Bayesian game.

The revelation principle is a statement about equilibria of a direct mechanism. Specifically, the principle says that in the direct mechanism we construct for

FIGURE 2 Equivalent Direct Mechanism

		Country 2 L	Country 2 H
Country 1	L	$p - c_1, 1 - p - c_2$	$1, 0$
	H	$0, 1$	$1, 0$

a given equilibrium of a given game, truth-telling is a Bayesian-Nash equilibrium, and it leads to the same outcomes as the original equilibrium in the original game. To see that this is true in our example, consider the direct mechanism given in Figure 2. To show that every type reporting truthfully is an equilibrium, consider country 1, assuming that country 2 is reporting its type truthfully. Then for the low-cost type of country 1, reporting L gives expected utility $(1-\pi)(p-c_l)+\pi(1)$, while reporting H gives expected utility $(1-\pi)(0)+\pi(1)$. Thus, reporting L is optimal if $c_l < p$. Similarly, for the high-cost type of country 1, reporting H is optimal if $c_h > p$. As both of these conditions hold given our assumptions on the parameters, it is optimal for country 1 to report its type truthfully, given that country 2 is doing so. In a similar manner, it is easy to check that truth-telling is optimal for country 2, given this behavior by country 1 in the direct mechanism. Finally, it is clear that the outcome in the direct mechanism is the same as the outcome in the equilibrium of the original game by the way that the direct mechanism is constructed.

The construction of the direct mechanism for this example offers insight into how we use the revelation principle in this article. As mentioned above, because each country has two possible types, the direct mechanism for the equilibrium in our example is a 2×2 strategic form Bayesian game. But as the method of construction makes clear, as long as each country has two possible types, the direct mechanism of *any* equilibrium of *any* extensive form, no matter how complicated, will be a 2×2 strategic form Bayesian game. Thus, the revelation principle allows us to make statements about *all* equilibria of *all* extensive forms, no matter how complicated, by analyzing the set of possible direct mechanisms, which in this case is just the set of 2×2 strategic form Bayesian games with the given outcomes. In this way, we can use the revelation principle to learn about arbitrarily complex games by studying the much simpler class of direct mechanisms. Of course, we do not limit ourselves to models with only two types. Also, we will most often make use of the contrapositive of the revelation principle which says that if some distribution of outcomes—say, always peace—is not an equilibrium to any direct mechanism, then it cannot be an equilibrium to any other extensive form game either.

Thus, by using the revelation principle, we can study equilibria that satisfy incentive-compatibility constraints in direct mechanisms and use our findings to establish general results about properties of equilibria in *all* possible crisis bargaining game forms. Before doing so though, we discuss three important qualitative properties of crisis bargaining that play a significant role in our later analysis.

First, given the anarchic nature of the international system, any peaceful agreement reached must be voluntary. In particular, we suppose that a country always has the option of rejecting a proposed settlement $v^g(a)$ if it thinks it will be better off by resolving the crisis by force. The form of this rejection can literally be the use of military force or it can be an escalation of the crisis which either leads to war or to the other side capitulating. Whichever is true, this feature of the bargaining context ensures each side can act in such a way that the outcome is either war or a voluntary settlement that that side prefers to war. Formally, we say a crisis bargaining game form has *voluntary agreements* if, for $i \neq j$, there exists an action $\tilde{a}_i \in A_i$ such that either $\pi^g(\tilde{a}_i, a_j) = 1$ for all $a_j \in A_j$ or for all $t \in T$, $v_i^g(\tilde{a}_i, a_j) \geq w_i(t)$.[13] Put simply, in a crisis bargaining game form with voluntary agreements, there is no way to force a country to accept an agreement that makes it worse off than it would expect to be by going to war.

A second important observation about crisis bargaining is that the process of bargaining has the potential to reveal, to a greater or lesser extent, the private information of the bargainers. Of course, a country should incorporate this additional information into its decision whether to reject a proposed settlement in favor of using force. To capture this formally, we let $\mu_i(v_i, t_i)$ be country i's updated belief about the type of country j after observing the settlement offer v_i.[14] As is standard, we assume that this belief is formed via Bayes' Rule, whenever possible.

[13]For an example of such an action, consider the take-it-or-leave-it bargaining game of Fearon (1995). For country 1, the action $\tilde{a}_1$ consists of demanding the whole pie. This demand will either be rejected, leading to war, or accepted, which results in a payoff that is higher than country 1's war payoff. Likewise, for country 2 in this game, rejecting any offer leads to war, which thus satisfies the definition. So this game has voluntary agreements.

[14]Although additional information may be revealed by the specific actions taken by countries in a specific game, we focus on the information revealed by the final settlement offer because this information will be available in all crisis bargaining games. Incorporating the additional information possible revealed in a specific game would only strengthen our results.

Combining this updating with our assumption on voluntary agreements, it is easy to show the following result is a consequence of the revelation principle.[15]

Result 2. *Suppose that s^* is an equilibrium of a crisis bargaining game form that has voluntary agreements. Then there exists an incentive-compatible direct mechanism such that $v_i(t) \geq E[w_i(t) \mid \mu_i(v_i, t_i)]$ for all $t \in T$ such that $\pi(t) \neq 1$.*

The intuition for this result is straightforward. Referring back to equation (1), we see that in a game form that has voluntary agreements, if $\pi^g(s^*(t)) \neq 1$, then this condition simplifies to the requirement that the equilibrium settlement value to country i must be at least as big as the expected war payoff to country i. That is, when faced with a settlement offer v_i, all types of country i have the option of playing $\tilde{a}_i$ and receiving a payoff of $E[w_i(t) \mid \mu_i(v_i, t_i)]$. Therefore if s^* is an equilibrium and $\pi(t) = \pi^g(s^*(t)) < 1$ for some $t = (t_1, t_2)$, it must be that this deviation is not profitable, which is true only if $v_i(t) \geq E[w_i(t) \mid \mu_i(v_i, t_i)]$. In other words, if agreements are voluntary and the unilateral use of force is always an option, any negotiated settlement must give each country a payoff at least as large as the payoff that it expects to get from settling the dispute by force, given what it has inferred about its opponent as a consequence of the negotiations. As an example, refer back to Figure 1. This game satisfies our voluntary agreements condition. Therefore, it is easy to verify that, given our assumptions on c_h, both players receive at least their war payoff in each cell of the direct mechanism in Figure 2.

The third and final observation that we incorporate into our analysis is the simple fact that war is costly. Because of this, we are interested in whether private information makes war unavoidable or whether there can be cases in which countries always arrive at peaceful settlements. Formally, an equilibrium s^* of a crisis bargaining game form is *always peaceful* if $\pi^g(s^*(t)) = 0$ for all $t \in T$. In other words, an always peaceful equilibrium of a game form is one in which no possible pair of types ever ends up abandoning a peaceful resolution of the dispute and resorting to force. Adding this requirement to Result 2 gives our final result.

Result 3. *Suppose that s^* is an always peaceful equilibrium of a crisis bargaining game form that has voluntary agreements. Then there exists an incentive-compatible direct mechanism such that for all $t \in T$, $\pi(t) = 0$ and $v_i(t) \geq E[w_i(t) \mid \mu_i(v_i, t_i)]$.*

[15]This result contains our version of what is usually known as the individual rationality constraint or the participation constraint.

The importance of this result is that if we can show that for a given information structure there is no incentive-compatible direct mechanism with these properties, then *no* always peaceful equilibria exist in *any* crisis bargaining game form that has voluntary agreements.

We conclude this section by presenting a lemma that will be useful in deriving the general results presented in the next section. This lemma is a direct consequence of the revelation principle applied to crisis bargaining games with independent types. Formally, we say that types are independent if there exist distributions F_1 and F_2 on T_1 and T_2, respectively, such that $F(t) = F_1(t_1)F_2(t_2)$ for every $t \in T$.

Lemma 1. *Suppose types are independent and $T_i = [\underline{t}_i, \bar{t}_i]$. Suppose also that $w_i(t)$ is continuously differentiable in t_1 and t_2. Let G be any crisis bargaining game form and let s^* be any equilibrium of G. Then $U_i(t_i)$ is continuous, is differentiable almost everywhere, and can be expressed as*

$$U_i(t_i) = U_i(\underline{t}_i) + \int_{\underline{t}_i}^{t_i} \int_{T_j} \pi(s, t_j) \times \frac{\partial w_i}{\partial t_i}(s, t_j)\, dF_j(t_j)\, dF_i(s),$$

where $\pi(t) = \pi^g(s^*(t))$.

Proof. Suppose that s^* is an equilibrium of a crisis bargaining game form. Then by Result 1, there exists an incentive-compatible direct mechanism yielding the same outcome as s^*. This direct mechanism is given by $\pi(t) = \pi^g(s^*(t))$ and $v_i(t) = v_i^g(s^*(t))$. Therefore the expected utility of type t_i from equation (2) can be written

$$U_i(t_i) = \int_{T_j} \pi(t_i, t_j) w_i(t_i, t_j) + (1 - \pi(t_i, t_j)) v_i(t_i, t_j)\, dF_j(t_j)$$

and the expected utility of falsely reporting a type $\tilde{t}_i$ is

$$U_i(\tilde{t}_i \mid t_i) = \int_{T_j} \pi(\tilde{t}_i, t_j) w_i(t_i, t_j) + (1 - \pi(\tilde{t}_i, t_j)) v_i(\tilde{t}_i, t_j)\, dF_j(t_j).$$

Clearly, $U_i(t_i) = U_i(t_i \mid t_i)$. Moreover, incentive-compatibility requires that for all $\tilde{t}_i \in T_i$,

$$U_i(t_i \mid t_i) \geq U_i(\tilde{t}_i \mid t_i),$$

which is the same as

$$U_i(t_i) = U_i(t_i \mid t_i) = \max_{\tilde{t}_i \in T_i} U_i(\tilde{t}_i \mid t_i).$$

Now, viewing $U_i(\tilde{t}_i \mid t_i)$ solely as a function of t_i, because w_i is differentiable in t_i, it follows that $U_i(\tilde{t}_i \mid t_i)$

is differentiable in t_i and because w_i is continuously differentiable on a compact set, this derivative is bounded. This implies that $U_i(\tilde{t}_i \mid t_i)$ is absolutely continuous in t_i. Therefore, the Envelope Theorem (Milgrom and Segal 2002, Theorem 2) gives us that $U_i(t_i)$ is absolutely continuous and

$$\frac{dU_i(t_i)}{dt_i} = \left.\frac{\partial U_i(\tilde{t}_i \mid t_i)}{\partial t_i}\right|_{\tilde{t}_i = t_i}$$
$$= \left.\int_{T_j} \pi(\tilde{t}_i, t_j) \frac{\partial w_i}{\partial t_i}(t_i, t_j)\, dF_j(t_j)\right|_{\tilde{t}_i = t_i}$$

at every point of differentiability. Evaluating this expression at $\tilde{t}_i = t_i$ yields the desired result. □

Results

In this section we establish several general results that hold for *every* crisis bargaining game that has voluntary agreements. Before doing so, though, we must identify the appropriate information structure for our analysis. Historically, the conflict literature has held that uncertainty is a central cause of conflict among countries, but has varied with respect to the type of uncertainty it views as important. Some scholars focus on uncertainty about the relative strength of the countries (Blainey 1988; Organski and Kugler 1980) while others concentrate on uncertainty about the costs of conflict or the resolve of countries (Kydd 2003; Morrow 1985; Ramsay 2004; Schultz 1998).

In particular, we focus on two kinds of uncertainty: uncertainty about the costs of war and uncertainty about the distribution of power. These various sources of uncertainty correspond to significantly different informational structures with important implications for strategic interaction in a crisis. Uncertainty about costs implies that the values of war to each country are independent, in the sense that one country's realized preference for fighting does not directly affect another's utility for fighting. Alternatively, the international system may produce uncertainty about the relative strength and the probability of victory. In this situation, the countries' values for war are interdependent, although information remains uncorrelated.

Uncertainty about the Cost of Conflict

We first consider the case in which each country is uncertain about the other's cost of fighting, which is also often interpreted as a country's level of resolve (Fearon 1994; Schultz 2001) or preference for fighting.[16] From a theoretical perspective, this is a situation with uncertainty about independent private values for war. That is, the realization of one state's resolve, or relative cost of war, does not directly influence the utility of the other side. A classic example of crisis bargaining with uncertainty regarding resolve occurred between Germany and the United Kingdom in the 1930s. While relatively well informed about the distribution of power in Europe and Germany's military armaments, the British Foreign Office was unsure about Hitler's intentions and how far he would go to achieve his territorial goals (Yarhi-Milo 2009).

To model uncertainty about costs formally, we suppose that the probability that country 1 wins a war, p, is common knowledge but there is uncertainty regarding each country's cost for fighting, c_i. In this setting, suppose country i's type, $c_i \in [\underline{c}_i, \bar{c}_i] = C_i$, is its cost of war, which is distributed according to a cumulative distribution function $F_i(c_i)$, with support C_i.[17] War is costly, so we assume that $\bar{c}_i > \underline{c}_i \geq 0$. Denote a pair of types $(c_1, c_2) = c \in C = C_1 \times C_2$.

We begin our analysis of this information structure by identifying some qualitative features of equilibrium play in crisis bargaining games with uncertain costs. Often, such qualitative features are central components in the analysis of a particular model. For example, in the classic take-it-or-leave-it bargaining model of conflict (Fearon 1995), it is possible to show that the expected probability of war is weakly decreasing in the cost of war for both countries. The following result demonstrates that this qualitative feature of this specific model is completely general—it will hold in any equilibrium of any crisis bargaining game with uncertainty about the cost of conflict. The result also shows that a similar conclusion can be reached about the overall expected utility of each country. For country i, define the expected probability of war, given strategy profile s, by

$$\Pi(c_i) = E\pi^g(c_i) = \int_{C_j} \pi^g(s(c_i, c_j))\, dF_j(c_j).$$

Also, recall that the expected utility of this strategy profile is given by

[16]It is often the case that uncertainty about costs is said to also model uncertainty about "resolve," where resolve is loosely defined as how much one side "cares" about the issue or item of dispute. The equivalence comes from the fact that, if costs are known, but the value of the prize is unknown, we can normalize the value of the prize, generating a new "relative cost" c/v_i which is a privately known value.

[17]Because the private information of country i is the cost of war, here we use the notation c_i rather than t_i for the type of country i.

$$U_i(c_i) = \int_{C_j} u_i(s(c_i, c_j), c_i)\, dF_j(c_j).$$

We then have the following proposition.

Proposition 1. *Suppose costs c_i are private information, but each country's probability of winning a war is common knowledge. Let G be any crisis bargaining game form and let s^* be any equilibrium of G. Then $\Pi(c_i)$ and $U_i(c_i)$ are both weakly decreasing in c_i and differentiable almost everywhere, and $U_i(c_i)$ is continuous in c_i.*

Proof. Suppose that s^* is an equilibrium of a crisis bargaining game form. Then by Result 1, there exists an incentive-compatible direct mechanism yielding the same outcome as s^*. This direct mechanism is given by $\pi(c) = \pi^g(s^*(c))$ and $v_i(c) = v_i^g(s^*(c))$. Therefore the expected utility of type c_i from equation (2) can be written

$$U_i(c_i) = \int_{C_j} \pi(c_i, c_j)(p_i - c_i) + (1 - \pi(c_i, c_j))v_i(c_i, c_j)\, dF_j(c_j)$$

and the expected utility of falsely reporting a type $\tilde{c}_i$ is

$$U_i(\tilde{c}_i \mid c_i) = \int_{C_j} \pi(\tilde{c}_i, c_j)(p_i - c_i) + (1 - \pi(\tilde{c}_i, c_j))v_i(\tilde{c}_i, c_j)\, dF_j(c_j).$$

Incentive-compatibility requires that for all $c_i, \tilde{c}_i \in C_i$,

$$U_i(c_i \mid c_i) \geq U_i(\tilde{c}_i \mid c_i) \quad \text{and} \quad U_i(\tilde{c}_i \mid \tilde{c}_i) \geq U_i(c_i \mid \tilde{c}_i).$$

Adding these two inequalities and simplifying yields

$$\int_{C_j} \pi(c_i, c_j)(p_i - c_i) - \int_{C_j} \pi(c_i, c_j)(p_i - \tilde{c}_i) \geq \int_{C_j} \pi(\tilde{c}_i, c_j)(p_i - c_i) - \int_{C_j} \pi(\tilde{c}_i, c_j)(p_i - \tilde{c}_i)$$

$$\int_{C_j} \pi(c_i, c_j)(\tilde{c}_i - c_i)\, dF_j(c_j) \geq \int_{C_j} \pi(\tilde{c}_i, c_j)(\tilde{c}_i - c_i)\, dF_j(c_j).$$

From this, it follows that if $\tilde{c}_i > c_i$, then $\Pi(c_i) \geq \Pi(\tilde{c}_i)$ and so $\Pi(c_i)$ is weakly decreasing. In addition, as it is monotonic on a closed interval, it is differentiable almost everywhere.

We now focus on $U_i(c_i)$. As $w_i = p_i - c_i$ is continuously differentiable in c_i, it follows from Lemma 1 that $U_i(c_i)$ is absolutely continuous and therefore continuous and differentiable almost everywhere. To show that $U_i(c_i)$ is weakly decreasing, take $c_i < \tilde{c}_i$ and apply Lemma 1 to get

$$U_i(\tilde{c}_i) - U_i(c_i) = \int_{c_i}^{\tilde{c}_i} \int_{C_j} (-\pi(s, c_j))\, dF_j(c_j)\, ds.$$

As $\pi(c)$ is always nonnegative, it follows that $U_i(c_i) \geq U_i(\tilde{c}_i)$. □

Thus, monotonicity of utility and the probability of war are general qualitative characteristics of any equilibrium of any game with uncertainty about the costs of war.[18] It is then interesting to observe that this monotonicity is borne out in a number of well-known, and quite different, models of crisis bargaining. For example, in the alternating-offers bargaining model of Powell (1996), it is straightforward to deduce from his Lemma 3 and Proposition 1 that the probability of war is deceasing in each player's cost of war. Similarly, the war of attrition model in Fearon (1994) has a monotone probability of war. From Proposition 2 in Fearon (1994, 584) we can see that a country with a low cost of war chooses to wait to some time t and then fight, while a country with a high cost waits for some time and then quits. The probability of war goes from positive in the first case to zero in the second. Finally, we can consider the model in Schultz (1999). In this model the probability of war is also a step function, which is positive when a country has a low cost of war but is zero when costs are large.

These types of comparative statics represent testable predictions that are robust to the particular formalization of crisis bargaining, as advocated by Powell (2004). Moreover, as this result depends only on the incentive compatibility constraints inherent in the revelation principle, it holds for all crisis bargaining games, even those without voluntary agreements. In this way, we see that the monotonicities identified in the proposition are purely a consequence of the incentives to misrepresent private information that are fundamental to the crisis bargaining environment.

In addition, this result shows that in any equilibrium of any crisis bargaining game, two types of a country with similar costs must receive similar expected utility in equilibrium. Thus a small change in the private information of the country leads to a small change in its expected payoff. This is a feature that is commonly found in "cut-point" equilibria of Bayesian games, but it is present here with no explicit assumption on the form of equilibrium strategies.

Of course, in a peaceful equilibrium the probability of war is constant at zero, which satisfies Proposition 1

[18] See Banks (1990) for additional monotonicity results that hold in Bayesian games with one-sided incomplete information. Unfortunately, these additional results do not extend to the case of two-sided incomplete information studied here.

in a trivial way. We now turn to inquiring about the characteristics of such equilibria. We begin with a simple existence result.

Proposition 2. *If costs c_i are private information, but each country's probability of winning a war is common knowledge, then there exists a crisis bargaining game form that has voluntary agreements in which an always peaceful equilibrium exists.*

Proof. To prove this result, it is enough to give an example of a game form that has voluntary agreements and that has an always peaceful equilibrium. The following very simple example shows that this is indeed the case. Consider a game form that has voluntary agreements such that $g_i^v(a) = p_i$ for all $a \in A$ and $\pi^g(a^*) = 0$ for some $a^* \in A$. In this case, the strategy profile $s^*(c) = a^*$ is an equilibrium because if either side deviates and starts a war, then both sides are worse off and if either side deviates to a different peaceful action, then the settlement amount does not change. Moreover, this equilibrium is always peaceful by construction. □

This proposition shows that there exists at least one game form that has voluntary agreements and that can eliminate the possibility of war. The game form described in the proof has a particularly simple payoff structure in which the potential settlement is always equal to a country's (commonly known) likelihood of success from war. One example of such a game form would be a direct "arbitration" game in which an arbitrator would present both sides with a take-it-or-leave-it offer. Since this agreement would make both sides better off, regardless of their costs for fighting, it would provide a rational and preferable alternative to war.[19] While a single example is sufficient to prove existence, it would be incorrect to conclude that this arbitration game is the only model with uncertainty about costs and voluntary agreements that has a peaceful equilibrium. For example, one kind of equilibrium studied in the alternating-offers model of Leventoğlu and Tarar (2008) has uncertainty about costs and zero probability of war. Interestingly, Leventoğlu and Tarar observe that a puzzling feature of the standard alternating-offers bargaining model of war is that a country cannot go to war in a period in which it makes a proposal, and when such behavior is allowed peaceful equilibria emerge.

At some level, though, the mere existence of game forms with peaceful equilibria is not a very satisfying result. It only tells us that a peaceful equilibrium is theoretically possible; it is not a general result about the possibility of war that applies to all possible crisis bargaining games. Put another way, Proposition 2 establishes that the common understanding in the literature that incomplete information about costs leads to the risk of war is not *always* true; there exist instances where it does not hold. However, without a better characterization of equilibrium incentives, we do not know if this case is just an isolated exception or if peaceful equilibria are the norm and the conventional view on the link between incomplete information and war must be rethought.

In light of this, we present a general characterization of peaceful equilibria in all possible crisis bargaining games that have voluntary agreements. For country i, define the expected settlement value from action a_i, given strategy s_j, by

$$Ev_i^g(a_i \mid s_j) = \int_{C_j} v_i^g(a_i, s_j(c_j))\, dF_j(c_j).$$

The following proposition gives a necessary condition for the existence of an always peaceful equilibrium.

Proposition 3. *Suppose costs c_i are private information, with $C_i = [0, \bar{c}_i]$, but each country's probability of winning a war is common knowledge. Let G be any crisis bargaining game form that has voluntary agreements. Then an equilibrium s^* of G is always peaceful only if, for $i = 1, 2$, (1) $Ev_i^g(s_i^*(c_i) \mid s_j^*) = p_i$ for all $c_i \in C_i$, (2) $Ev_i^g(a_i \mid s_j^*) \leq p_i$ for all $a_i \in A_i$ and (3) $v_i^g(s^*(c)) \geq p_i - c_i$ for all $c \in C$.*

Proof. Suppose that s^* is an always peaceful equilibrium of a crisis bargaining game form that has voluntary agreements. Then by Result 3, there exists an incentive-compatible direct mechanism such that $\pi(c) = 0$ and

$$v_i(c) \geq E[w_i(c) \mid \mu_i(v_i, c_i)] = p_i - c_i$$

for all $c_i \in C_i$ and $i = 1, 2$. Here, $E[w_i(c) \mid \mu_i(v_i, c_i)] = p_i - c_i$ because the value of war to country i does not depend on the value of c_j. This direct mechanism is given by $\pi(c) = \pi^g(s^*(c))$ and $v_i(c) = v_i^g(s^*(c))$.

Because $\pi(c) = 0$, applying Lemma 1 yields $U_i(c_i) = U_i(0)$ for all c_i and therefore $U_i(c_i) = \int_{C_j} v_i(c_i, c_j)\, dF_j(c_j)$ is constant in c_i. In addition, because the constraint $v_i(c) \geq p_i - c_i$ must hold for $c_i = 0$, we have

$$\int_{C_j} v_i(c_i, c_j)\, dF_j(c_j) = \int_{C_j} v_i(0, c_j)\, dF_j(c_j) \geq p_i,$$

for all $c_i \in C_i$, $i = 1, 2$. To show that this expression must hold with equality, suppose not. Then

[19]In this sense, our result lends hope to institutional designers that the problem of war is solvable, at least in the case of uncertainty about costs or "resolve."

$$\begin{aligned}
&\int_{C_1}\int_{C_2}[v_1(c)+v_2(c)]\,dF_2(c_2)\,dF_1(c_1)\\
&\quad=\int_{C_1}\int_{C_2}v_1(c)\,dF_2(c_2)\,dF_1(c_1)\\
&\qquad+\int_{C_2}\int_{C_1}v_2(c)\,dF_1(c_1)\,dF_2(c_2)\\
&\quad>\int_{C_2}p\,dF_2(c_2)+\int_{C_1}(1-p)\,dF_1(c_1)\\
&\quad=p+(1-p)=1.
\end{aligned}$$

But as peaceful settlements are efficient, $v_1(c)+v_2(c)=1$ for all $c \in C$ and therefore we have a contradiction. This proves statement (1) of the proposition. The remaining two statements follow directly from the assumption that s^* is an equilibrium and thus no profitable deviation is possible. □

This proposition shows that the simple payoff structure used in the proof of Proposition 2 is, in fact, a general property of peaceful equilibria. Every peaceful equilibrium must have this same simple payoff structure in which the terms that a country accepts are completely insensitive to the costs or resolve of the country. High-cost and low-cost countries (and all types in between) must receive the same expected settlement value. The most natural example of a strategy profile that would generate such insensitivity is a (completely) pooling strategy, in which all types of a country choose the same action.[20] A general lesson from this result, then, is that peaceful equilibria must be "simple" equilibria that do not depend on the private information of the countries. That is, an always peaceful agreement necessarily depends only on publicly observable information and cannot vary in response to the privately known costs or resolve of countries.

We can better understand the general idea of this characterization of peaceful equilibria by observing that it is the generalized expression of the well-known incentive to misrepresent private information. It is this incentive that prevents countries from resolving their uncertainty by way of "cheap talk" communication prior to bargaining (Fearon 1995). Proposition 3 shows that this same incentive forces peaceful equilibria to be completely insensitive to the private information that countries possess. This follows because if a peaceful equilibria did provide different expected settlements to different types, the type getting the worse settlement would have an incentive to mimic the behavior of the type getting the better settlement. Such a deviation is profitable because in a peaceful equilibrium the discipline generated by the risk of war does not exist, which means we could not have had an equilibrium that responds to the countries' private information.

It is immediately clear from this general lesson that because many examples of crisis bargaining games in the literature have equilibria that do not conform to this simple payoff structure, these examples support the common understanding that incomplete information about costs leads to war. Indeed, given the severity of the necessary condition that strong types receive the same expected settlement as weak types, it does seem intuitively clear that many game forms will fail this necessary condition. In fact, we can strengthen this observation into the following necessary and sufficient condition:

Corollary 1. *Suppose costs c_i are private information, but each country's probability of winning a war is common knowledge. Let G be any crisis bargaining game form that has voluntary agreements and let s^* be any equilibrium of G. Then $U_i(c_i) \neq U_i(\tilde{c}_i)$ for some $c_i, \tilde{c}_i \in C_i$ if and only if there is a positive probability of war in equilibrium.*

Proof. The "if" direction follows directly from the argument in the proof of Proposition 3 that $U_i(c_i)$ is constant in c_i if the equilibrium is peaceful. To show the other direction, suppose $U_i(c_i) = U_i(\tilde{c}_i)$ for all $c_i, \tilde{c}_i \in C_i$. Then by Lemma 1,

$$U_i(\tilde{c}_i) - U_i(c_i) = \int_{c_i}^{\tilde{c}_i}\int_{C_j}(-\pi(s, c_j))\,dF_j(c_j)\,ds = 0,$$

for all $c_i, \tilde{c}_i \in C_i$. This implies that $\pi(c)$ is zero almost everywhere. The result follows. □

In other words, this corollary establishes that two types can have different expected utilities if and only if war occurs with positive probability. If we interpret private information about costs as a country's "resolve," then this result can be stated more clearly in the following form. Crisis bargaining games in which more resolved countries expect to be better off than less resolved countries in equilibrium are those with a positive probability of war. Put another way, this result shows that the *only* way that countries stand to gain from their resolve is through running the risk of war, as in the well-known "risk-reward" trade-off.

As an illustration of Proposition 3 and Corollary 1, we briefly return to the model of Leventoğlu and Tarar (2008). As mentioned above, their analysis considers an extension of the alternating-offers bargaining model with

[20]It would be very useful if this result implied that completely pooling equilibria were the only type of peaceful equilibrium possible. However, it can be shown that there also exist nonpooling strategies with the property that, in expectation, the payoff to settlement is constant and that therefore satisfy the conditions given in the proposition. Details of this example are available from the authors.

one-sided uncertainty about costs that allows both countries to choose war in any period. In the always peaceful equilibrium to their game, either both types accept the initial offer or both types reject this initial offer and make the same counteroffer, which is accepted. In either case, the payoff of two types is the same, as required by Corollary 1.

This model also allows us to demonstrate how Proposition 3 and Corollary 1 can be used to characterize equilibria outcomes of games. At the end of their paper, Leventoğlu and Tarar consider what would happen if their game had two-sided incomplete information with a continuum of types. While they do not attempt to solve this version of their model, they conjecture that a peaceful equilibrium would exist and suggest that the agreement would likely be on the division $(p, 1-p)$, perhaps after some delay. Our results allow us to address these conjectures directly, without actually solving the model. Suppose that the two countries' costs are drawn from the interval $[0, \bar{c}_i]$ in this model. Then Proposition 3 tells us that if a peaceful equilibrium like Leventoğlu and Tarar describe exists, the expected value of settlement for every type pair of the two countries must be $(p, 1-p)$. Moreover, as these expected values sum to one for every type pair and future payoffs are discounted, it must be that in any peaceful equilibrium the game ends with an accepted offer in the first round, with probability one. Thus we can confirm the conjecture that the peaceful settlement will be $(p, 1-p)$ and, additionally, demonstrate that this settlement will occur without delay in the bargaining.[21] In this way, the results in this section can offer significant insight into equilibrium behavior of complex games with general uncertainty without actually doing the hard work of solving these games.

Uncertainty about Relative Strength

While the previous section dealt with the case of uncertainty about the costs of conflict, in this section we deal with a second source of uncertainty that has received significant attention in the literature—uncertainty about the distribution of power and the likelihood of success in war. In this case, countries are assumed to be informed about their opponent's relative cost of fighting, but are uncertain about the likely outcome of conflict. In particular, countries have private information about the quality of their military or their combat strategy that leads each side to hold private beliefs about what will happen as a result of fighting a war. This is a situation with interdependent values. That is, each country's utility for conflict is not only dependent on its own type but also depends on the type of its opponent. A nice example of this class of problems is the run-up to the 1973 Arab-Israeli War. In this instance, any questions about Israeli resolve, by and large, had been laid to rest in the 1967 war; there was little doubt that Israel would defend herself resolutely if attacked. But even knowing this fact, Egypt and Syria were uncertain about how effectively the Israeli Defense Force would respond to a surprise two-front war. In particular, the Syrian leadership thought their military build up put them in a strong position to retake the Golan Heights (Yarhi-Milo 2009).

To model this, we assume that the costs of engaging in a war, c_1 and c_2, are common knowledge, but both countries have private information regarding the probability of winning. We implement this in our framework by supposing that country i's type, $t_i \in [\underline{t}_i, \bar{t}_i] = T_i$, is independently distributed according to a distribution function F_i and the probability that country 1 prevails in a war, $p(t_1, t_2)$, is a function of both types. We assume that the types can be ordered such that the probability of victory is monotone in the countries' types. That is, higher types have a greater chance of winning, all other things being equal. Formally, this assumption is that $t_1 > t_1'$ implies $p(t_1, t_2) \geq p(t_1', t_2)$, for all $t_2 \in T_2$. Likewise, we assume that p is monotonically decreasing in t_2. Also, to ensure there is uncertainty about the distribution of power, we assume that p is not everywhere constant. In this way, the type t_i reflects the "strength" of country i and thus the probability of victory depends on the relative strength of the two combatants.

Our first result deals with how expected outcomes must vary as the strength of a country varies. Recall from the previous section that the expected utility of type t_i of country i is denoted $U_i(t_i)$ and the expected probability of war is denoted $\Pi(t_i)$.

Proposition 4. *Suppose costs are common knowledge but each country is uncertain about the probability of winning a war. Let G be any crisis bargaining game form and let s^* be any equilibrium of G. Then $U_i(t_i)$ is weakly increasing in t_i and differentiable almost everywhere. If, in addition, $p(t)$ is continuously differentiable, then $U_i(t_i)$ is continuous in t_i.*

If $p(t_1, t_2)$ is strictly increasing in t_1 and strictly decreasing in t_2, then there exist $t_i^p, t_i^w \in T_i$, with $t_i^p \leq t_i^w$ such that $\Pi(t_i) = 0$ for all $t_i < t_i^p$, $0 < \Pi(t_i) < 1$ for all $t_i \in (t_i^p, t_i^w)$, and $\Pi(t_i) = 1$ for all $t_i > t_i^w$.

[21] It is worth noting that Leventoğlu and Tarar (2008) consider the case where costs are bounded away from zero, i.e., c_i is in the interval $[\underline{c}_i, \bar{c}_i]$, $\underline{c}_i > 0$. Our characterization will hold as long as there is positive density, although it can be arbitrarily small, on costs in the interval $[0, \underline{c}_i]$.

Proof. Suppose that s^* is an equilibrium of a crisis bargaining game form. Then by Result 1, there exists an incentive-compatible direct mechanism yielding the same outcome as s^*. This direct mechanism is given by $\pi(t) = \pi^g(s^*(t))$ and $v_i(t) = v_i^g(s^*(t))$.

As in the proof of Proposition 1, we can write the expected utility of type t_i reporting a type $\tilde{t}_i$ as

$$U_i(\tilde{t}_i \mid t_i) = \int_{T_j} \pi(\tilde{t}_i, t_j)(p_i(t_i, t_j) - c_i) + (1 - \pi(\tilde{t}_i, t_j))v_i(\pi(\tilde{t}_i, t_j))\, dF_j(t_j).$$

Incentive-compatibility requires that for all $t_i, \tilde{t}_i \in T_i$, $U_i(t_i \mid t_i) \geq U_i(\tilde{t}_i \mid t_i)$.

To show that $U_i(t_i)$ is weakly increasing, fix $\hat{t}_i$ and t_i such that $\hat{t}_i < t_i$. Then

$$U_i(t_i) - U_i(\hat{t}_i) = U_i(t_i \mid t_i) - U_i(\hat{t}_i \mid \hat{t}_i) = U_i(t_i \mid t_i) - U_i(\hat{t}_i \mid t_i) + U_i(\hat{t}_i \mid t_i) - U_i(\hat{t}_i \mid \hat{t}_i).$$

The above incentive-compatibility condition gives $U_i(t_i \mid t_i) - U_i(\hat{t}_i \mid t_i) \geq 0$. In addition, we have

$$\begin{aligned} U_i(\hat{t}_i \mid t_i) - U_i(\hat{t}_i \mid \hat{t}_i) &= \int_{T_j} \pi(\hat{t}_i, t_j)(p_i(t_i, t_j) - c_i) \\ &\quad -\pi(\hat{t}_i, t_j)(p_i(\hat{t}_i, t_j) - c_i)\, dF_j(t_j) \\ &= \int_{T_j} \pi(\hat{t}_i, t_j)(p_i(t_i, t_j) - p_i(\hat{t}_i, t_j))\, dF_j(t_j). \end{aligned}$$

By the monotonicity of p_i, the integrand is always nonnegative. Therefore, $U_i(\hat{t}_i \mid t_i) - U_i(\hat{t}_i \mid \hat{t}_i) \geq 0$. This means that $U_i(t_i) \geq U_i(\hat{t}_i)$, which proves that $U_i(t_i)$ is weakly increasing. It is differentiable almost everywhere because it is a monotonic function on a closed interval. If, in addition, $p(t)$ is continuously differentiable, then $w_i(t) = p_i(t) - c_i$ is continuously differentiable, and $U_i(t_i)$ is continuous by Lemma 1.

We now prove the final part of the proposition. By the same argument as in the proof of Proposition 1, incentive compatibility requires that

$$\int_{T_j} (\pi(t_i, t_j) - \pi(\hat{t}_i, t_j))(p_i(t_i, t_j) - p_i(\hat{t}_i, t_j))dF_j(t_j) \geq 0.$$

If we consider $\hat{t}_i < t_i$ and assume p_i is strictly monotonic in t_i, then $p_i(t_i, t_j) - p_i(\hat{t}_i, t_j) > 0$ for all t_j. If $\Pi(t_i) = 0$, then the above inequality implies $\Pi(\hat{t}_i) = 0$. Likewise, if $\Pi(\hat{t}_i) = 1$, then the above inequality implies $\Pi(t_i) = 1$. Now define t_i^p to be the supremum of the set $\{t_i \in T_i \mid \Pi(t_i) = 0\}$ (and $t_i^p = \underline{t}_i$ if this set is empty) and t_i^w to be the infimum of the set $\{t_i \in T_i \mid \Pi(t_i) = 1\}$ (and $t_i^w = \bar{t}_i$ if this set is empty). The result follows. □

This proposition shows that, regardless of the game form, a stronger type of a country is never worse off than a weaker type of that country. Thus the monotonicity result for expected utility given in Proposition 1 with uncertainty about costs continues to hold with uncertainty about relative power. With the added assumption that the probability of winning does not jump discontinuously with type, we find that the continuity result of Proposition 1 also carries over to this setting. However, the earlier result on the monotonicity of the expected probability of war does not fully carry over. Instead, when the probability of winning is strictly monotonic in type, the weakest types have zero expected probability of war, moderate types have an expected probability of war strictly between zero and one, and the strongest types have war with certainty.[22] Thus, expected probability of war displays a limited kind of monotonicity in this environment.

One important consequence of focusing on uncertainty over relative strength is that the process of bargaining can reveal important clues as to the likely strength of the two countries. In particular, when a country receives a settlement offer, it can update its prior about the private information of the opposing state by inferring what must be true of the other state in order to generate the received offer. Recall that $\mu_i(v_i, t_i)$ is country i's updated belief about the type of country j after observing the settlement offer v_i. Let $V_1(t_1, v) = \{t_2 \mid v_1(t_1, t_2) = v\}$ be the set of possible types of country 2 that a given type t_1 of country 1 would think are possible after observing a settlement v.[23] In this setting, then,

$$E[w_1(t) \mid \mu_1(v_1, t_1)] = E[p(t_1, t_2) \mid V_1(t_1, v_1)] - c_1, \tag{3}$$

and a similar expression holds for country 2. The right-hand side of this inequality is simply the updated expected utility for war incorporating the inference about the types of the other country from the observed settlement offer.

For convenience, we use the following notation in stating our results:

$$P_1(t_1) = \int_{T_2} p(t_1, y)dF_2(y) \quad \text{and}$$

$$P_2(t_2) = \int_{T_1} p(x, t_2)dF_1(x).$$

In words, $P_1(t_1)$ is the expected probability of winning a war for type t_1 of state 1 and $P_2(t_2)$ is the expected probability of *losing* a war for type t_2 of state 2.

[22]Note, however, that some of these sets of types could be empty.

[23]In general, this conditional expectation must be defined abstractly. But this abstract definition simplifies in many cases. For example, if $V_1(t_1, v_1)$ is an interval, then

$$E[p(t_1, y) \mid V_1(t_1, v_1)] = \frac{\int_{V_1(t_1,v_1)} p(t_1, y)\, dF_2(y)}{\int_{V_1(t_1,v_1)} dF_2(y)}.$$

Let $\bar{c} = P_1(\bar{t}_1) - P_2(\bar{t}_2)$. It follows from the monotonicity of p that $\bar{c} > 0$. Our next result shows that if the costs of war are less than $\bar{c}$, then no matter what bargaining procedure is used, there is a positive probability of war during a crisis.

Proposition 5. *Suppose costs are common knowledge but each country is uncertain about the probability of winning a war. If $c_1 + c_2 < \bar{c}$, then no always peaceful equilibrium exists in any crisis bargaining game form that has voluntary agreements.*

Proof. The method of proof is by contradiction. We begin by supposing there is an always peaceful equilibrium of a crisis bargaining game form that has voluntary agreements. By Result 3, there exists an incentive-compatible direct mechanism such that $\pi(t) = 0$ and $v_i(t) \geq E[w_i(t) \mid \mu_i(v_i, t_i)]$ for all t and $i = 1, 2$. Because $\pi(t) = 0$, the expected utility of country 1 with true type t_1 reporting type $\hat{t}_1$ is

$$U_1(\hat{t}_1 \mid t_1) = \int_{T_2} v_1(\hat{t}_1, y)\, dF_2(y). \tag{4}$$

The incentive-compatibility condition is then

$$U_1(t_1 \mid t_1) \geq U_1(\hat{t}_1 \mid t_1) \quad \text{for all} \quad t_1, \hat{t}_1 \in T_1.$$

However, from equation (4), it is clear that $U_1(\hat{t}_1 \mid t_1)$ does not depend on t_1. Therefore, the only way the incentive-compatibility condition can be satisfied for all t_1 and $\hat{t}_1$ is if $U_1(\hat{t}_1 \mid t_1)$ is a constant, for all t_1, and $\hat{t}_1$.[24] We write $\bar{U}_1$ for this constant.

Turning now to the condition that $v_i(t) \geq E[w_i(t) \mid \mu_i(v_i, t_i)]$ for all t and $i = 1, 2$, we use equation (3) evaluated at the type pair $t_1 = \bar{t}_1$ and $t_2 = \bar{t}_2$ to get

$$v_1(\bar{t}_1, t_2) \geq E[p(\bar{t}_1, t_2) \mid V_1(\bar{t}_1, v_1)] - c_1.$$

Taking expectations of both sides, we get

$$E[v_1(\bar{t}_1, t_2)] \geq E[E[p(\bar{t}_1, t_2) \mid V_1(\bar{t}_1, v_1)]] - c_1.$$

By the law of iterated expectations, this expression is equivalent to

$$\bar{U}_1 = \int_{T_2} v_1(\bar{t}_1, t_2)\, dF_2(t_2) \geq \int_{T_2} p(\bar{t}_1, t_2)\, dF_2(t_2) - c_1. \tag{5}$$

By a similar argument, it follows that

$$\begin{aligned} \bar{U}_2 &= \int_{T_1} v_2(t_1, \bar{t}_2)\, dF_1(t_1) \\ &\geq \int_{T_1} [1 - p(t_1, \bar{t}_2)]\, dF_1(t_1) - c_2. \end{aligned} \tag{6}$$

We next show that $\bar{U}_1 + \bar{U}_2 = 1$. Starting with the fact that $v_1(t_1, t_2) + v_2(t_1, t_2) = 1$ for all pairs (t_1, t_2), it follows that

$$\begin{aligned} &\int_{T_1}\int_{T_2} [v_1(t) + v_2(t)]\, dF_2(t_2)\, dF_1(t_1) = 1 \\ &\int_{T_1}\int_{T_2} v_1(t)\, dF_2(t_2)\, dF_1(t_1) \\ &\quad + \int_{T_2}\int_{T_1} v_2(t)\, dF_1(t_1)\, dF_2(t_2) = 1 \\ &\int_{T_1} \bar{U}_1\, dF_1(t_1) + \int_{T_2} \bar{U}_2\, dF_2(t_2) = 1 \\ &\bar{U}_1 + \bar{U}_2 = 1. \end{aligned}$$

Therefore, adding inequalities (5) and (6) yields

$$\begin{aligned} 1 \geq &\int_{T_2} p(\bar{t}_1, y)\, dF_2(y) - c_1 + 1 \\ &- \int_{T_1} p(x, \bar{t}_2)\, dF_1(x) - c_2, \end{aligned}$$

from which it follows that

$$c_1 + c_2 \geq P_1(\bar{t}_1) - P_2(\bar{t}_2) = \bar{c}.$$

This contradicts the supposition that $c_1 + c_2 < \bar{c}$ and thus proves the proposition. □

This proposition shows that there is a range of costs such that the countries' private information about the probability of winning is always an obstacle to peace in any interaction with voluntary agreements. Phrased as a general result, Proposition 5 shows that small costs of war is a sufficient condition for the impossibility of completely peaceful settlements in crisis bargaining games with voluntary agreements. We thus can view this result as extending the claim of Fearon (1995) that when costs are low, private information leads to a positive probability of war from the case of uncertainty about costs to the case of uncertainty about relative power. The reason that the war problem is particularly hard to solve in games with interdependent types is because, in such games, the strongest type of a given country knows that every type of their opponent cannot be a stronger adversary. So, from the strongest type's perspective, war is a relatively attractive option and therefore in order to persuade the strongest type to forego this option, such types must get relatively generous terms from peaceful settlements. This in turn creates an incentive for weaker types to pretend that they are strong in order to secure this high payoff. The only way a peaceful settlement can exist with these incentives is if the settlement gives *every* type of a country the same payoff as the *strongest* type of the country. But

[24]This result would follow immediately from Lemma 1 if we made the additional assumption that $p_i(t)$ was continuously differentiable in t_i.

when the total costs of war are less than the critical value $\bar{c}$, such settlements are impossible.

As an example of this sufficient condition, consider the model of Reed (2003). In this model, country 1 makes an offer to country 2, who either accepts the offer or goes to war. This model has one-sided uncertainty: country 2 knows the probability of winning the war, while country 1 does not. In fact, in Reed's model, country 2's type is exactly the probability p that 1 wins the war. In our notation, then, country 2's type $t_2 \in [a, b] \subseteq [0, 1]$ and $p(t_2) = t_2$. Therefore, $P_1 = \int_a^b y\, dF_2(y)$ and $P_2(t_2) = t_2$, because country 2 has no uncertainty. Thus, the sufficient condition in Proposition 5 states that if $c_1 + c_2 < \bar{c} = \int_a^b y\, dF_2(y) - a$, then war must occur in any equilibrium.[25] To understand what this condition means in Reed's model, first note that in order to avoid war, country 1 must make an offer that the strongest type of country 2 will accept. Thus, country 1 must make an offer that gives at least $1 - a - c_2$ in order to ensure peace. Such an offer gives country 1 a payoff of at most $a + c_2$. However, suppose country 1 makes an offer that all types of country 2 reject. This leads to war with certainty, so country 1 has an expected payoff of $\int_a^b y\, dF_2(y) - c_1$. Thus, peace is not possible if $a + c_2 < \int_a^b y\, dF_2(y) - c_1$, which is exactly the sufficient condition in Proposition 5. To be clear, this does not show that such an extreme offer is optimal, only that country 1 will not make an offer that avoids war entirely.

The intuition of Proposition 5 further suggests that when costs are large enough, peaceful settlements may be possible. This is exactly what we show in the next result. Specifically, we show that if the total costs of war are greater than the critical value $\bar{c}$, it is possible to find *some* game form that always avoids war.

Proposition 6. *Suppose costs are common knowledge but each country is uncertain about the probability of winning a war. If $c_1 + c_2 \geq \bar{c}$, then there exists a crisis bargaining game form that has voluntary agreements in which an always peaceful equilibrium exists.*

Proof. As with Proposition 2, it is enough to give an example of a crisis bargaining game form that has voluntary agreements which has an always peaceful equilibrium. Consider a crisis bargaining game form that has voluntary agreements such that $v^g(a) = P_1(\bar{t}_1) - c_1$ for all $a \in A$ and $\pi^g(a^*) = 0$ for some $a^* \in A$. By construction, the strategy profile $s^*(t) = a^*$ is always peaceful. To show that it is an equilibrium, first note that if either side deviates to a different peaceful action, then the settlement amount does not change. If country 1 takes an action that leads to war, its expected payoff from war, given its type t_1, is

$$\int_{T_2} p(t_1, y)\, dF_2(y) - c_1 \leq \int_{T_2} p(\bar{t}_1, y)\, dF_2(y) - c_1 = P_1(\bar{t}_1) - c_1 = g_v(a^*),$$

where the first inequality follows from the monotonicity of p. Therefore, no type of country 1 will deviate. A similar argument shows that country 2 will not deviate and start a war. Therefore $s^*(t)$ is an equilibrium and the proof is complete. □

The proof of Proposition 6 presents a crisis bargaining game form with voluntary agreements that has a peaceful equilibrium when $c_1 + c_2 \geq \bar{c}$. Thus, the proposition demonstrates the existence of a peaceful game form. It should be emphasized that the result does not claim that *all* game forms will have such a peaceful equilibrium when $c_1 + c_2 \geq \bar{c}$. As an example, consider the model of Reed (2003) described above and assume that F_2 is the uniform distribution on $[a, b]$. In this case, it is easy to check that $\bar{c} = (b - a)/2$. However, it is also easy to verify that the peaceful equilibrium in which country 1 makes the offer $x = a + c_2$ that is accepted by all types of country 2 exists only when $c_1 + c_2 \geq b - a$. Thus, Reed's model is an example in which war still occurs in equilibrium for values of $c_1 + c_2 > \bar{c}$.

Another important feature of Proposition 6 is that it implies that a trivial sufficient condition for the existence of a peaceful game form is that $c_1 + c_2 \geq 1$. Yet, much of the time this condition is unlikely to hold, i.e., it is rarely the case that a real-world conflict generates relative costs that are greater than the total value of the object of dispute. However, if there is a third party that is willing to provide a sufficiently large subsidy to the peace process, such as an international organization or a superpower, a peaceful settlement is possible. Permitting this possibility, we can say that, if a third party provides a subsidy

$$\phi \geq P_1(\bar{t}_1) - P_2(\bar{t}_2) - (c_1 + c_2),$$

then there exists a crisis bargaining game form that has voluntary agreements in which an always peaceful equilibrium exists. In this circumstance, the subsidy amount is the minimum amount that will ensure that there is an agreement that both sides will prefer to unilaterally starting a war. Thus, in a world with a large powerful country willing to provide sufficient subsidies, the occurrence of war as a consequence of private information about relative power can be avoided.[26] Unfortunately for

[25] In this case, $\bar{t}_2 = a$ because this is the strongest type of country 2.

[26] This corollary supports the argument that if there is a global hegemon, then the international system is likely to be more

small $(c_1 + c_2)$, depending on the technology of war $p(t)$ and the distribution of types, the size of the subsidy can approach the value of the whole prize.

We conclude our analysis of this case by presenting a characterization of peaceful equilibrium that is similar to the results in the previous section.

Proposition 7. *Suppose costs are common knowledge but each country is uncertain about the probability of winning a war. Suppose also that $p(t)$ is continuously differentiable, strictly increasing in t_1, and strictly decreasing in t_2. Let G be any crisis bargaining game form that has voluntary agreements and let s^* be any equilibrium of G. Then $U_i(t_i) \neq U_i(\tilde{t}_i)$ for some $t_i, \tilde{t}_i \in T_i$ if and only if there is a positive probability of war in equilibrium.*

Proof. The proof of Proposition 5 shows that $U_i(t_i)$ is constant in any peaceful equilibrium. For the other direction, suppose $U_i(t_i) = U_i(\tilde{t}_i)$ for all $t_i, \tilde{t}_i \in T_i$. Then by Lemma 1,

$$U_i(\tilde{t}_i) - U_i(t_i) = \int_{t_i}^{\tilde{t}_i} \int_{T_j} \pi(s, t_j) \frac{\partial p_i}{\partial t_i}(s, t_j)\, dF_j(t_j)\, ds = 0,$$

for all $t_i, \tilde{t}_i \in T_i$. As p_i is strictly increasing in t_i, $\partial p_i/\partial t_i > 0$, and so $\pi(t)$ is zero almost everywhere. The result follows. □

This proposition tells us that, as in the previous section, every peaceful equilibrium of every crisis bargaining game with uncertainty about relative power must give every type the same expected settlement. As with the case with uncertainty about costs, we can use this result to shed new light on strategic problems without having to solve complicated games. For example, consider the crisis bargaining game in Smith (1998). The extensive form of this game is quite involved. To start, a leader from a "defender" country makes a cheap talk announcement of their intent to defend a client state if that state is attacked. Next, an aggressor state chooses to attack the client or not. If the aggressor attacks, then the client can retaliate or capitulate. If the client retaliates, the defender chooses to intervene or not. If the defender intervenes, he changes the probability that the client wins the war. After the crisis is complete, the defender's citizens decide to keep or replace their leader, given the cheap talk message, the play of the crisis, and the outcome of the war.

One of the main conclusions of Smith's paper is that the defender's message "can only be informative if war occurs with positive probability." To prove this, Smith uses several lemmas,[27] but we can derive this result directly from Proposition 7.[28] In a peaceful equilibrium, all types of the attacker must get the same expected utility. This means that either the client retaliates with positive probability when attacked and all types of the attacker stay out, or, if the client never retaliates, all types of the attacker attack and get the capitulation utility. In either case, the outcome of the crisis does not vary. Therefore, for the defender country, the only way for its utility to vary is by the probability that the government is replaced. But again by Proposition 7, all types of the defender must have the same expected utility. Therefore, all types must have the same probability of replacement. But this implies that the cheap talk message must be uninformative, because otherwise differences in the types would be revealed and the less competent types would be removed. Therefore, messages must be uninformative if the probability of war is zero. In this way, we can derive useful results from specific models using our propositions without actually solving for the equilibria in those models.

Discussion

As we have emphasized, one strength of the game-free method of analysis presented here is the ability it provides to establish results that hold across a wide variety of strategic interactions between countries. That is, our results apply to any direct bargaining process in which the countries can make offers and counteroffers to each other, with or without communication, as well as any arbitration mechanism in which the parties communicate to a formal institution and this institution decides how the dispute will be settled. In this way, we have been able to give a clear picture of the general consequences of private information and the incentives to misrepresent this information on the possibility and form of war and peace. In addition, given our interest in understanding how different kinds of uncertainty may imply different general tendencies for strategic behavior, our mechanism design approach allows us to clearly and cleanly explore the effects of various types of uncertainty across many different game forms.

Using our techniques, we have established several general results. The first group of results concerns the monotonicity properties of equilibria. We show that

peaceful, given that the hegemon is willing to pay the cost (Gilpin 1981; Keohane 1984; Kindleberger 1973).

[27] Notably, Smith's Lemma 1 proves that expected utility is weakly increasing in type, which is exactly our Proposition 4.

[28] Although Smith's model has three players, it is easy to extend our results to this case.

crisis bargaining games with two-sided private information have some, but not all, of the monotonicity properties described by Banks (1990) for games with one-sided private information. Proposition 1 shows that countries' expected probability of war and expected utility are both (weakly) decreasing in their costs. So the general intuition that low-cost types have higher expected utilities from crisis bargaining, but also face greater risk of war, carries over from the setting with one-sided incomplete information. Unlike Banks (1990), however, we are unable to guarantee monotonicity in expected settlements with two-sided incomplete information about costs. Similarly, Proposition 4 shows that with uncertainty about strength, the monotonicity result for utilities also holds in every crisis bargaining game. For expected probability of war, however, we are only able to establish a limited kind of monotonicity, unlike Banks (1990).

A second interesting general result occurs in the interdependent case, where countries are uncertain about relative power. Here we find that the conditions that determine whether or not peaceful equilibria can exist need not depend on any aspect of the game form. As seen in Propositions 5 and 6, the nonexistence of peaceful equilibria depends only on the costs of fighting and the form of the uncertainty about relative power, and not on the details of the bargaining. When the conditions for peaceful equilibria are met, then a simple mechanism that gives both sides at least the best they could hope for from war is a simple solution. Otherwise, when such an arrangement is not possible, there is no bargaining process, no matter how complicated, that ensures peace.

A final interesting result of our analysis is that the effect of private information and the incentive to misrepresent varies as we vary the structure of the decision maker's uncertainty. This is true even though in each framework we consider there is common knowledge that a settlement exists that both sides prefer to war. The fact that in each framework the conditions on the game form that lead to war are different makes clear that the type of uncertainty that exists in the international environment can have important implications for players' strategies and the probability of war. At the same time, our focus on the informational roots of conflict is not meant to suggest that this is the only potential cause of war. To be clear, this article does not claim to present an all-inclusive model for war. Rather, we have attempted to provide game form free results for an important class of strategic situations discussed in the literature on private information and war.

One basic assumption that we have used to define this class of games is that any peaceful agreement must be acceptable to both sides of the conflict. This "voluntary agreements" assumption is intended to be general and substantively motivated. We require that there be an action that allows decision makers to reject settlements that give them less than they might expect from fighting, but we do not require that this action leads to a war lottery or even immediate fighting. This flexibility implies that our results hold in many different kinds of games. For example, the fact that our results hold in game forms that include capitulation as a possible outcome distinct from war and peace also aids in clarifying the application of our model to several historical cases. While there are many clear examples of the direct, unilateral use of force such as the Japanese attack on Pearl Harbor and the American invasion of Iraq in 2003, there are other historical cases that could be better categorized as capitulation, such as the invasions of Denmark and Luxembourg in 1940. Thus, our results apply to both kinds of circumstances, including those that involve armed aggression but that ultimately do not involve significant armed conflict. Moreover, as noted above, we can see our results materialize in a wide variety of extensive forms, from alternating-offers bargaining and wars of attrition to crisis bargaining models and models of extended deterrence (Fearon 1994; Leventoğlu and Tarar 2008; Powell 1996; Schultz 1999; Smith 1998).

While our view is that our "voluntary agreements" assumption is a natural consequence of the anarchic nature of the international arena, there may remain some question as to how best to describe the process leading countries to war. One alternative assumption is that war occurs only if both countries agree to fight and otherwise the status quo or an imposed settlement prevails. This assumption has been used to consider the idea of mutual optimism as a cause of war. Central to this literature (Blainey 1988; Wittman 1979) is the claim that *both* countries must want to fight for a war to occur, hence the name *mutual* optimism. Not surprisingly, this alternative assumption yields different conclusions about the possibility of war. Thus our results for the case of voluntary agreements highlight the role of private information and the incentives to misrepresent and can be interpreted as the complement to the mutual optimism argument.[29]

The other basic assumption we use to define the class of crisis bargaining games is that whatever form the interaction of the two sides takes, the end result is always either a negotiated settlement or war. Dividing outcomes into these two categories seems to us to be quite natural. At the same time, because we make the common assumption that the war outcome is a game-ending costly lottery, we should consider how our results speak to the recent

[29] For more on the issue of mutual optimism and assumptions about how countries end up in war, see Fey and Ramsay (2007, 2008).

formal conflict literature on "war as a bargaining process" (Filson and Werner 2002; Powell 2004; Slantchev 2003b). In a manner similar to Powell (2004), we can relax the assumption that fighting is a game-ending move by supposing that fighting creates a chance of a militarily decisive outcome but can also be indecisive and lead to further negotiations or further fighting. We can fit such a model into our theoretical framework by identifying all terminal nodes with a decisive military outcome as "war" outcomes and all other terminal nodes as "peaceful settlement" outcomes. However, as a number of rounds of fighting can precede any such outcome, we must make changes to two of the assumptions we make about our class of crisis bargaining games. First, we must modify the assumption that peaceful settlements are efficient because if the countries fight for some time before reaching a peaceful settlement, the payoffs to the two countries for such a strategy profile will sum to less than one. Second, for a similar reason we must modify our assumption that a country's war payoff does not depend on the actions taken in the game. The longer the countries choose to fight before a decisive military outcome results, the lower is the expected value of this outcome.

How do our results fare if we make these two changes to our framework in order to accommodate models of war as a bargaining process? Reassuringly, it is possible to show that all of our main results continue to hold. In the case of the first change, we show in a companion paper (Fey and Ramsay 2009) that permitting settlement outcomes to be inefficient does not change any of the qualitative features of our results. This is also true with the second change, permitting the war payoff to vary with the actions taken in the game. Rather than provide formal proofs here, instead we briefly explain why our results continue to hold.[30] Our reasoning rests on the fact that even if a country's war payoff depends on its actions, as its equilibrium actions depend on its type, we can use the revelation principle to find a direct mechanism in which a country's war payoff depends directly on its type. In this way, Lemma 1 continues to hold. As this lemma is a key ingredient in our proofs, this result, coupled with the fact that the actions of a country can only change its payoff by incurring the costs of multiple rounds of fighting, yields the same characterizations as found in our propositions.

Conclusion

This article has set out to establish general results about the fundamental incentives inherent in crisis bargaining. Rather than fix a particular model and derive results that are limited to this single extensive form, we develop a method to identify properties shared by all equilibria of a large class of crisis bargaining models. The value of this approach is twofold. First, such an approach allows us to cut through the clutter of the endless variety of modeling assumptions, in particular assumptions about sequencing of moves, and to show how uncertainty and the incentive to misrepresent affect the probability of war in a variety of crisis bargaining settings. Second, our "game form free" approach allows us to compare the effect that differences in the kinds of private information possessed by actors have on strategic choices of the actors. We have developed these results across a range of information environments in order to identify conditions under which the positive probability of war is an unavoidable consequence of private information and conditions under which peaceful resolution of conflicts is possible. We have also characterized the nature of such peaceful resolutions. In short, we have found that while it is possible to give general results that support the rationalist explanation of war as a consequence of private information and the incentive to misrepresent this information, the link between private information and war depends in important ways on the types of uncertainty that states face.

Throughout this article, we have emphasized the generality of our approach and the necessity of utilizing such a general approach to justify broad claims about the causes of war. Although we are able to give results that identify specific relationships between information and outcomes, there is an obvious trade-off between general results and specific predictions. The relative value of the two must depend on the question being asked. In particular, questions regarding particular behaviors in a specific institutional setting call for an approach very different from the one we take here. On the other hand, as a theoretical matter, we may be interested in results that are not institution specific, or that apply to a wide class of different institutions and strategic settings. To the extent that answers to general questions about the relationship between uncertainty and war, independent of institutional specifics, are the object of interest, no iteration of the process of analyzing specific protocols will reveal these general equilibrium phenomena. For these questions, the game-free approach is a useful tool. As a guiding principle, we endorse the "two-step" approach outlined in Banks (1990) in which general results are established first using the approach developed here and then additional behavioral predictions are generated from a particular game form that captures specific institutional features of interest.

In the end, while this article has laid out a framework for analyzing some general questions about war

[30] Details are presented in a supplemental appendix available online.

and peace, several additional aspects of this approach remain to be explored. First, while we have focused on how various theoretical sources of uncertainty may influence the possibility for "well-designed" institutions to eliminate unwanted, inefficient conflict, it is also true that institutions do not just arrive from nowhere; they are endogenous to the negotiation process between states. It is natural, then, to consider the process by which countries bargain over possible game forms that then govern their interaction. It should be possible to use our approach to deal with such a situation by viewing the institution selection process as itself part of a more broadly defined game and applying the revelation principle.[31] Second, although each of the two kinds of uncertainty that we have considered in this article have involved one-dimensional types, it should also be possible to generalize our approach to allow for multiple-dimensional types. One example of this kind of uncertainty would be a case in which countries possessed private information about both their own cost of conflict as well as the relative balance of power. We leave the analysis of this case as a topic for further research.

References

Baliga, Sandeep, and Tomas Sjöström. 2008. "Mechanism Design: New Developments." In *The New Palgrave Dictionary of Economics* (2nd ed.), ed. Steven N. Durlauf and Lawrence E. Blume. New York: Palgrave Macmillan, 542–48.

Banks, Jeffrey S. 1990. "Equilibrium Behavior in Crisis Bargaining Games." *American Journal of Political Science* 34(3): 599–614.

Baron, David P. 2000. "Legislative Organization with Informational Committees." *American Journal of Political Science* 44(3): 485–505.

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

Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason: Domestic and International Imperative*. New Haven, CT: Yale University Press.

Chwe, Michael Suk-Young. 1999. "Minority Voting Rights Can Maximize Majority Welfare." *American Political Science Review* 93(1): 85–97.

Compte, Olivier, and Philippe Jehiel. 2009. "Veto Constraint in Mechanism Design: Inefficiency with Correlated Types." *American Economic Journal: Microeconomics* 1(1): 182–206.

Cramton, Peter C., and Thomas R. Palfrey. 1995. "Ratifiable Mechanisms: Learning from Disagreement." *Games and Economic Behavior* 10(2): 255–83.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3): 577–92.

[31] See Holmstrom and Myerson (1983) for more on this approach.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3): 379–414.

Fey, Mark, and Kristopher W. Ramsay. 2007. "Mutual Optimism and War." *American Journal of Political Science* 51(4): 738–54.

Fey, Mark, and Kristopher W. Ramsay. 2009. "Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types." *Review of Economic Design* 13(3): 233–50.

Fey, Mark, and Kristopher W. Ramsay. 2010. "Mutual Optimism and Unilateral War." Typescript, Princeton University.

Filson, Darren, and Suzanne Werner. 2002. "A Bargaining Model of War and Peace." *American Journal of Political Science* 46(4): 819–38.

Forges, Francoise. 1999. "Ex Post Individually Rational Trading Mechanisms." In *Current Trends in Economics: Theory and Applications*, ed. Charalambos Aliprantis, Nicholas Yannelis, and Ahmet Alkan. New York: Springer, 157–76.

Gilpin, Robert. 1981. *War and Change in World Politics*. New York: Cambridge University Press.

Holmstrom, Bengt, and Roger B. Myerson. 1983. "Efficient and Durable Decision Rules with Incomplete Information." *Econometrica* 51(6): 1799–1819.

Jackson, Matthew O. 2003. "Mechanism Theory." In *Encyclopedia of Life Support Systems*, ed. Ulrich Derigs. Oxford: EOLSS.

Keohane, Robert. 1984. *After Hegemony*. Princeton, NJ: Princeton University Press.

Kindleberger, Charles P. 1973. *The World in Depression: 1929–1939*. Berkeley: University of California Press.

Kydd, Andrew. 2003. "Which Side Are You On? Bias, Credibility, and Mediation." *American Journal of Political Science* 47(4): 597–611.

Leventoğlu, Bahar, and Ahmer Tarar. 2008. "Does Private Information Lead to Delay or War in Crisis Bargaining?" *International Studies Quarterly* 52: 533–53.

Matthews, Steven A., and Andrew Postlewaite. 1989. "Pre-Play Communication in Two-Person Sealed-Bid Double Auctions." *Journal of Economic Theory* 48(1): 238–63.

Milgrom, Paul, and Ilya Segal. 2002. "Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70(2): 583–601.

Morrow, James D. 1985. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining." *American Journal of Political Science* 33(4): 941–72.

Muthoo, Abhinay. 1999. *Bargaining Theory with Applications*. New York: Cambridge University Press.

Myerson, Roger B. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47(1): 61–73.

Myerson, Roger B. 2008. "Mechanism Design." In *The New Palgrave Dictionary of Economics* (2nd ed.), ed. Steven N. Durlauf and Lawrence E. Blume. New York: Palgrave Macmillan, 533–42.

Myerson, Roger B., and Mark A. Satterthwaite. 1983. "Efficient Mechansims for Bilateral Trading." *Journal of Economic Theory* 29(2): 265–81.

Organski, A. F. K., and Jacek Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.

Pillar, Paul R. 1983. *Negotiating Peace: War Termination as a Bargaining Process*. Princeton, NJ: Princeton University Press.

Powell, Robert. 1996. "Bargaining in the Shadow of Power." *Games and Economic Behavior* 15(2): 255–89.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton, NJ: Princeton University Press.

Powell, Robert. 2004. "Bargaining and Learning While Fighting." *American Journal of Political Science* 48(2): 344–61.

Ramsay, Kristopher W. 2004. "Politics at the Water's Edge: Crisis Bargaining and Electoral Competition." *Jounral of Conflict Resolution* 48(4): 459–86.

Reed, William. 2003. "Information, Power, and War." *American Political Science Review* 97(4): 633–41.

Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50: 97–109.

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

Schultz, Kenneth. 1998. "Domestic Opposition and Signaling in International Crisis." *American Political Science Review* 94(4): 829–44.

Schultz, Kenneth. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53(2): 233–66.

Schultz, Kenneth. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.

Slantchev, Branislav L. 2003a. "The Power to Hurt: Costly Conflict with Completely Informed States." *American Political Science Review* 47(1): 123–33.

Slantchev, Branislav L. 2003b. "The Principle of Convergence in Wartime Negotiation." *American Journal of Political Science* 97(4): 621–32.

Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92(3): 623–38.

Smith, Alastair, and Allan C. Stam. 2006. "Divergent Beliefs in 'Bargaining and the Nature of War'." *Journal of Conflict Resolution* 50(4): 614–18.

Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44(3): 469–84.

Waltz, Kenneth. 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley.

Waltz, Kenneth. 2001. *Man, the State and War: A Theoretical Analysis*. New York: Columbia University Press.

Wittman, Donald. 1979. "How Wars End." *Journal of Conflict Resolution* 23(4): 743–63.

Wittman, Donald. 2009. "Bargaining in the Shadow of War: When Is a Peaceful Resolution More Likely?" *American Journal of Political Science* 53(3): 588–602.

Yarhi-Milo, Keren. 2009. "Knowing Thy Adversary: Assessment of Intentions in International Politics." PhD dissertation, University of Pennsylvania.