War and Incomplete Information*

Bahar Leventoğlu
Assistant Professor
Department of Political Science
Stony Brook University
Stony Brook, NY 11794
e-mail: bahar.leventoglu@stonybrook.edu

Ahmer Tarar
Assistant Professor
Department of Political Science
Texas A&M University
4348 TAMU
College Station, TX 77843-4348
e-mail: ahmertarar@polisci.tamu.edu

March 10, 2006

Abstract

It has long been argued that incomplete information can lead to war between more-or-less rationally led states. Fearon (1995) and Powell (1999) formalize this idea and show that a "risk-return tradeoff" emerges under private information, whereby a state may trade off a positive risk of war to try to get a better deal at the negotiating table. We generalize Powell's canonical model of crisis bargaining in the shadow of power, and find that although risk-return equilibria exist, so do "non-risk" equilibria in which the probability of war is zero. The results suggest that private information by itself often merely leads to delay in reaching a preferred-to-war negotiated settlement, and if inefficient war occurs before then, it is only because the dissatisfied state is sufficiently impatient. The results significantly modify our understanding of one of the canonical "rationalist explanations for war."

*We gratefully acknowledge the support of the National Science Foundation (SES-0518185 and SES-0518945).

1 Introduction

The idea that inefficient war could result between two more-or-less rationally led states due to some kind of incomplete information, uncertainty, or misperception between them has a long history among students of international relations (e.g., Blainey 1988; Jervis 1976; Van Evera 1999). For example, Blainey (1988) argues that wars often occur when both sides are very optimistic about their chances of victory.

Fearon (1995) formalizes the war-due-to-uncertainty idea in expected utility and game theoretic terms, and shows that even in a scenario where there exist negotiated settlements that both sides strictly prefer to war, war could be a rational outcome between two states when there is private information about military capabilities or resolve and incentives to misrepresent it. Fearon uses a take-it-or-leave-it (henceforth TILI) offer crisis bargaining game to establish his argument. Powell (1996a, 1996b, 1999) generalizes Fearon's model to allow for a potentially endless numbers of offers and counteroffers by both sides. Powell, too, shows that war can result under incomplete information even though there exist negotiated settlements that both sides strictly prefer to war. Today, the idea that private information and incentives to misrepresent it is a sound explanation for war between two rationally led states is probably one of the most commonly accepted ideas in international relations, and indeed is the most prominent of Fearon's (1995) three "rationalist explanations for war."

But is private information as complete an explanation for war as previous studies have implied? In Fearon and Powell's models, war emerges under incomplete information due to what Powell calls the "risk-return tradeoff." Suppose one state, call it a "satisfied" state, can make an offer of a division of a disputed territory to the "dissatisfied" state. Under complete information, the satisfied state optimally offers just enough to make the dissatisfied state indifferent between accepting this offer and going to war. The dissatisfied state accepts the offer, and war is avoided. Suppose, however, that the satisfied state is uncertain of the dissatisfied state's utility from war. Then, the satisfied state does not know exactly how large an offer will be needed to get the dissatisfied state to accept it, thereby avoiding war.

If the satisfied state is sufficiently confident at the beginning of the game that it faces a low-resolve type of dissatisfied state (i.e., one whose cost of war is relatively high), then the satisfied state makes a relatively low offer which is rejected by highly resolved types. In all equilibria of Powell's (1996a, 1996b) model, the highly resolved types reject this offer and go to war rather than make a counteroffer. Thus, war occurs if the dissatisfied state ends up being highly resolved, and hence these are called "risk-return tradeoff" equilibria: the satisfied state trades off a positive risk of war for a greater return at the negotiating table if war does *not* occur.

Previous formal work on war-due-to-uncertainty thus suggests that under conditions of uncertainty, an uncertain satisfied state faces a hair-trigger decision: either its initial offer is accepted, or it is rejected and war occurs. This work suggests that pre-war bargaining is completely futile in the presence of incentives to misrepresent one's private information in order to get a better deal (Fearon 1995). However, from an intuitive point of view, this seems somewhat odd: if the satisfied state's initial offer is too small, why would the dissatisfied state go to war rather than make a counteroffer, when it is commonly known there exist negotiated settlements that both sides *strictly* prefer to war? Or, to put it another way, why is the risk-return tradeoff so risky?

In this paper, we show that the risk-return tradeoff is the *unique* equilibrium outcome in Powell's model under incomplete information only because that model imposes the restriction that a state can initiate a war only in periods where the other side makes an offer, and not in periods where it itself makes an offer. This restriction means that in equilibrium, the satisfied state has all of the bargaining leverage, and because of this, the dissatisfied state is always better off going to war if it gets a low initial offer, rather than making a counteroffer. Thus, the satisfied state is always in the hair-trigger risk-return tradeoff scenario: either its initial offer is accepted, or it is rejected and war occurs.

We then generalize Powell's canonical model of crisis bargaining in the shadow of power to allow states to go to war in *any* period. In the generalized model, unlike in Powell's model, both sides have bargaining leverage in equilibrium, as we would intuitively expect from a model in which both sides can make proposals. Because both sides have bargaining leverage, it turns out that, unlike in Powell's model, the dissatisfied state sometimes finds it optimal to make a counteroffer rather than go to war if the other side's initial offer is too small. In particular, when the dissatisfied state's discount factor is sufficiently high, there exist perfect Bayesian equilibria (PBE) under incomplete information in which the dissatisfied state makes a counteroffer (that is accepted) rather than goes to war, if the satisfied state's initial offer is too low. Thus, contrary to the results of Fearon and Powell's more restrictive models, uncertainty does not always lead to a hair-trigger risk-return tradeoff decision for the satisfied state. There also exist equilibria in which, if it were to make a low offer, this offer is simply rejected and agreement is reached in the next period, and in which the probability of war is therefore zero.

This result has a number of important implications for research on the causes of war. By only identifying risk-return equilibria, previous work has overstated the extent to which uncertainty leads to war between rationally led states, especially when it is commonly known that there exist negotiated settlements that both sides strictly prefer to war. Our results suggest that the only effect of private information *per se* is to delay the reaching of a preferred-to-war negotiated settlement, and if inefficient war occurs before then, it is only because the dissatisfied state is too impatient to wait until then. Thus, it appears that private information *plus* impatience are jointly sufficient to produce war, and not private information alone. This is a significant modification of our understanding of how private information leads to war.

In addition, to the best of our knowledge, ours is the only formal model of crisis bargaining in which an agreement can be reached in equilibrium after some delay, without conflict occurring in the meantime. Of course, this is quite common empirically. For example, in the Agadir Crisis of 1911 (also commonly called the Second Moroccan Crisis) sparked by the deployment of the German gunboat *Panther* off the coast of French-controlled Morocco, a number of offers and counteroffers were made before an agreement was finally reached. The German foreign minister initially demanded the entire French Congo as compensation for

German acceptance of a French protectorate over Morocco — in the final agreement reached, however, Germany accepted considerably less, and no conflict occurred in the meantime (e.g., Barraclough 1982; Lowe 1994, 174-83). Previous formal models of crisis bargaining cannot account for cases like this.

The rest of the paper is organized as follows. In the next section, we briefly present the bargaining approach to war. We then show how Powell generalizes Fearon's TILI offer bargaining model to allow a potentially endless number of offers and counteroffers, but still imposes a restriction that leads to the risk-return tradeoff being the *unique* PBE outcome under incomplete information. We then generalize Powell's model and establish the main results, under complete and incomplete information. We conclude with a discussion of the implications of these results for research on the causes of war.

2 Fearon's Model

Figure 1, drawn from Fearon (1995), graphically illustrates the bargaining approach to war pioneered by Fearon and Powell (1996a and 1996b). Two countries (labeled D and S) are involved in a dispute over a divisible good (e.g., territory) whose value to both sides is normalized to 1. The two sides can either peacefully reach an agreement on a division of the good, or they can go to war, in which case the side that wins obtains the entire good and the side that loses receives none of it. Moreover, war is costly, with side D and S's cost of war being $c_D, c_S > 0$, respectively. Assume that if war occurs, side D wins with probability $1 > p > 0$ and side S wins with probability $1 - p$. Then, country D's expected utility from war is $EU_D(war) = (p)(1) + (1-p)(0) - c_D = p - c_D$. Similarly, country S's expected utility from war is $EU_S(war) = (p)(0) + (1-p)(1) - c_S = 1 - p - c_S = 1 - (p + c_S)$. Thus, as seen in Figure 1, there is a bargaining range of agreements $[p - c_D, p + c_S]$ such that for all agreements in this range, both sides prefer the agreement to war (and both sides *strictly* prefer any agreement in the interior of this range).[1]

[1]As Powell (2002) points out, the interpretation that the war is total and the victorious side wins everything while the losing side gains nothing is not necessary for this argument. Simply interpret p to be the expected division of the good resulting from war. War is inefficient ex post, because the two sides could have

In Fearon's framework, a rationalist explanation for war has to explain why costly war may occur between two rational unitary states despite the existence of this bargaining range. Fearon presents three such explanations: (i) private information about military capabilities or resolve and incentives to misrepresent it, (ii) issue indivisibilities, and (iii) dynamic commitment problems, such as when one side is growing in power relative to the other. In this paper, we focus on the private information explanation. Powell (2004b, 2006) points out that issue indivisibilities leading to war is really due to a commitment problem, and shows how a range of seemingly different commitment problems can all be described by a single, general inefficiency condition.

Fearon game-theoretically analyzes this framework using a TILI offer bargaining model in which one side, say S, can propose some division of the disputed good, say $(x, 1-x)$, where $1 \geq x \geq 0$ is D's share and $1-x$ is S's share. D can either accept this offer, in which case each side's utility is simply its proposed share (we assume risk neutrality throughout this paper), or it can reject it, in which case each side gets its payoff from war. Under complete information, this TILI bargaining protocol gives a unique subgame-perfect equilibrium (SPE), in which D only accepts all offers that give it at least its utility from war (i.e., it accepts all offers such that $x \geq p - c_D$), and S offers D exactly its utility from war. That is, agreement is reached on $[p - c_D, 1 - (p - c_D)]$, and war is avoided. Because S can make a TILI offer, it gets all of the gains from avoiding war, i.e., it gets its most preferred outcome in the bargaining range (e.g., Romer and Rosenthal 1978).[2]

Fearon then supposes that S is uncertain about D's cost of war, c_D. Suppose that c_D lies in the range $[c_{D_l}, c_{D_h}]$. That is, c_{D_l} is the lowest cost "type" of D that S might be facing (or most resolved, because its expected payoff from war is the highest; see Figure 2), and c_{D_h} is the highest cost type (or least resolved). In equilibrium, each type of D only accepts all agreements that give it at least its expected payoff from war. Thus, S faces a tradeoff: it can either make the big offer $p - c_{D_l}$ which *all* types of D accept and thus avoid war with

peacefully divided the good in the same ratio and avoided the costs of war (Fearon 1995).

[2]We prefer to call this the gains from avoiding war rather than the usual term "gains from cooperation," because the negotiated settlement and the process that leads to it may resemble coercive diplomacy much more so than cooperation.

certainty, or S can make a lower offer which only less resolved types of D accept, but which leaves S with a bigger share of the pie *if* it is accepted. If S's initial belief about D's type puts sufficient weight on D being a less resolved type, then S's optimal offer in equilibrium is to offer less than $p - c_{D_l}$, and thus war occurs if D ends up being a highly resolved type. Powell (1999) calls this the "risk-return tradeoff," because S trades off a positive risk of war for a greater return at the negotiating table if war does *not* occur.

Fearon then points out that previous explanations of war due to incomplete information do not explain why D does not use diplomacy or some other means of communication to convey its true cost of war (or resolve).[3] In equilibrium, war occurs if D ends up being a highly resolved type. Why, then, do not highly resolved types of D simply reveal their true cost of war in order to avoid an inefficient war?[4] Fearon then generalizes the model to allow D to send a costless message to S once D knows its cost of war. He shows that in any equilibrium of this model, S makes the same offer regardless of the message it receives, and the probability of war is the same as in the model without communication. The reason is that, if there was an equilibrium in which different messages cause S to make different offers, then all types of D would have an incentive to send the message that causes S to make the most favorable offer. But then no information would be conveyed by the message, and hence S could not rationally be conditioning its offer on the message. In short, the incentive by less resolved types to mimic the behavior of highly resolved types in order to get a better deal prevents information from being meaningfully conveyed in equilibrium.[5] Therefore, Fearon points out, it is not private information *per se* that leads to war, but private information *plus incentives to misrepresent it in order to get a better deal.*

[3] Earlier formalizations of the war-due-to-uncertainty idea include, e.g., Bueno de Mesquita and Lalman (1992) and Morrow (1989).

[4] With a TILI bargaining protocol, there is a trivial answer to this question. But the logic of Fearon's answer is very general.

[5] Some information could be conveyed in equilibrium; however, not enough to cause S to actually condition its offer on the message received. In other words, the different types of D cannot be separating "too much" in equilibrium, in terms of the messages they send.

3 Powell's Model

The major limitation of Fearon's model is that only one side can make an offer, and rejection of that offer automatically results in war. That is, the other side cannot make a counteroffer. However, in most actual bargaining situations, there is no reason why the other side cannot make a counteroffer and why rejection of the initial offer must automatically result in war.[6] Thus, Fearon's model is not ideal for studying the actual decision to deliberately launch a war, because in his model that decision is not separated from the decision to reject an offer.

Powell (1996a, 1996b, 1999) generalizes Fearon's model to allow each side to make a counteroffer if it rejects an offer. More specifically, Powell's model is shown in Figure 3 (only three periods are shown in the figure, but this is actually an infinite horizon game). One side (which in the figure is D, but this is not necessary) begins by making an initial offer, say $(x_0, 1-x_0)$, of a division of the good. The other side can accept that offer, in which case each side's payoff is simply the present discounted value of getting that share of the pie forever after, $x_0/(1-\delta)$ for D and $(1-x_0)/(1-\delta)$ for S, where $1 > \delta > 0$ is the players' common discount factor. Or it can reject an offer and go to war, in which case each side's payoff is the present discounted value of getting its payoff from war forever after, $(p-c_D)/(1-\delta)$ for D and $(1-p-c_S)/(1-\delta)$ for S. So far, this is exactly like Fearon's model. However, the third choice that the player has is to reject an offer but make a counteroffer (in the next period, in which payoffs are discounted by δ) rather than going to war. There is some status quo division of the good, $(q, 1-q)$, from which the players derive utility each period until they reach an agreement or go to war (this seems to be a reasonable assumption regarding most issues over which states have disputes, e.g., territory, which in most cases is already divided in some ratio). Thus, if an agreement is reached on some division of the pie $(z, 1-z)$ in period t $(t = 0, 1, 2, \ldots)$, then D's payoff is $\sum_{i=0}^{t-1} \delta^i q + \sum_{i=t}^{\infty} \delta^i z$ and S's payoff is $\sum_{i=0}^{t-1} \delta^i (1-q) + \sum_{i=t}^{\infty} \delta^i (1-z)$. If they go to war in some period t $(t = 0, 1, 2, \ldots)$, then D's payoff is $\sum_{i=0}^{t-1} \delta^i q + \sum_{i=t}^{\infty} \delta^i (p - c_D)$ and S's payoff is $\sum_{i=0}^{t-1} \delta^i (1-q) + \sum_{i=t}^{\infty} \delta^i (1-p-c_S)$.

[6] Alternatively, instead of a bargaining model, Fearon's model can be interpreted as one side using force to present the other with a military fait accompli, which the other side can either accept or go to war to try to overturn.

Powell's model is well suited for analyzing the decision to rationally launch a war, because in this model war only occurs when an actor decides that this is preferred to bargaining further. That is, in this model bargaining further is always a possibility, and hence war will only be chosen if an actor decides that war is better. This is in contrast to Fearon, in which in responding to a TILI offer, an actor's only choice is to accept it or go to war.

Introducing a status quo division of the good into the model necessitates making a distinction between satisfied and dissatisfied states. Powell calls a state satisfied if the current (i.e., status quo) division of the good provides it with at least as much utility as its payoff from going to war. In contrast, a state is dissatisfied if it strictly prefers to go to war rather than live with the status quo. D is satisfied if $q \geq p - c_D$ and dissatisfied if $q < p - c_D$ (this can be easily visualized in Figure 1). S is satisfied if $1 - q \geq 1 - p - c_S$, or $q \leq p + c_S$. S is dissatisfied if $q > p + c_S$. Both sides are satisfied if $p + c_S \geq q \geq p - c_D$ (i.e., the status quo lies within the preferred-to-war bargaining range). Only D is dissatisfied if $q < p - c_D$, and only S is dissatisfied if $q > p + c_S$. If the two sides agree on the probability that each prevails in war, then at most one state can be dissatisfied.

A dissatisfied state can credibly threaten to use force to try to change the status quo, because by definition a dissatisfied state is one that prefers war to living with the status quo. A satisfied state cannot credibly threaten to use force to try to change the status quo, because by definition it prefers the status quo to war. When both sides are satisfied, then neither can credibly threaten to use force to try to change the status quo, and hence war does not occur in equilibrium, nor does any revision of the status quo take place. Because the status quo division of the good is efficient (they are on the Pareto frontier, i.e., there are no mutual gains to be had), no revision will occur when neither can credibly threaten to use force to push for a revision.[7]

[7] In terms of bargaining models in economics, the way to think of this is a model in which the players have inside options (the status quo division of the good) as well as outside options (the decision to go to war rather than bargain further) (Muthoo 1999). One way in which this model differs from most economic models (e.g., Muthoo 1999) is that the inside options are jointly efficient, i.e., there is no surplus to be divided, and hence no mutual gains to be had. But the reason there are incentives to bargain is that if one side is dissatisfied, i.e., its outside option exceeds the value of its inside option, then it can credibly threaten to exercise its outside option, but the exercising of the outside option is jointly inefficient, i.e., $(p - c_D) + (1 - p - c_S) = 1 - c_D - c_S < 1$.

Now suppose one side is dissatisfied, and (without loss of generality), in particular, suppose D is dissatisfied (hence the labels D and S), i.e., suppose $q < p - c_D$ (we will henceforth refer to D as a "he" and S as a "she"). Then, as Powell points out, in the subgame perfect equilibria of this game, war is avoided.[8] Whenever S makes an offer, it offers D just its utility from war, i.e., it offers $p - c_D$, and keeps the rest of the pie for itself. This is the same outcome in Fearon's model, when S gets to make a TILI offer. Note that S gets all of the gains from avoiding war whenever it makes an offer. However, a rather odd agreement is reached when D makes an offer. In the SPE that Powell identifies, in any period in which D makes an offer, it proposes for itself the share of the pie $q(1 - \delta) + \delta(p - c_D)$, and the rest of the pie for S, and S accepts this offer. The strange thing is that D's offer for itself is *less than its utility from war*: $q(1 - \delta) + \delta(p - c_D) < p - c_D$ for $\delta < 1$. This can be seen from Figure 4, which graphically illustrates the equilibrium shares of the pie that D and S propose for D, as δ changes. Notice that D's offer for itself is a weighted average (convex combination) of its status quo payoff and its utility from war. Since its status quo payoff is strictly less than its utility from war, this weighted average is strictly less than its payoff from war.[9]

As seen in Figure 4, there are three somewhat odd findings in Powell's complete information results. First, there is a first mover *disadvantage*, which is atypical for a Rubinstein-type (1982) bargaining model with complete information. In particular, each player would rather have the other make the first offer (which is accepted). Second, the satisfied state has all of the bargaining leverage, in that it gets *all* of the gains from avoiding war (and gains even more than that when D makes the first offer). In equilibrium, S has the same bargaining power that it would if it could make a TILI offer (even more, when D makes the first offer), even though *both sides have proposal power.* And third, which is perhaps the oddest of all

[8]Powell (1996a, 263) identifies one (stationary) SPE and states that it is the unique SPE. However, because of an indifference condition, there are in fact an infinite number of SPE. However, the average per-period payoffs to the players are the same in all of these equilibria, which is what is important. We characterize the subgame perfect equilibria in the appendix, while generalizing the model to allow the players to have different discount factors.

[9]Figure 4 is drawn for $p = 0.5$, $c_D = c_S = 0.2$, and $q = 0.1$. It has the same general shape for all values of these parameters.

from a substantive viewpoint, D actually proposes for itself less than its payoff from war, fully knowing that this proposal will be accepted. (Note from Figure 4 that as $\delta \rightarrow 1$, the first and third problems disappear in the limit. In the limit, the outcome approaches one in which S can make a TILI offer, regardless of who gets to make the first proposal, and so the second problem remains, and we will argue, this is what drives the incomplete information results.)

Under complete information, war does not occur in equilibrium, just as in Fearon's model. Instead, the status quo is peacefully revised in D's favor, with S getting (at least, depending on who makes the first offer) all of the gains from avoiding war.

Powell then analyzes the case where the two sides are uncertain about each other's cost of war. He shows that in the *unique* perfect Bayesian equilibrium (PBE) outcome, Fearon's risk-return tradeoff emerges. In particular, a dissatisfied type of player D never rejects S's initial offer to make a counteroffer. Instead, it goes to war if it gets a proposal that gives it less than its utility from war. Therefore, S is essentially in the position of making its optimal TILI offer (as in Fearon's model), knowing that if it is rejected, war will result rather than a counteroffer. If S's initial belief puts sufficient weight on D being a less resolved type, then S's optimal offer in equilibrium is low enough that war occurs with positive probability. Moreover, this risk-return tradeoff is the *unique* perfect Bayesian equilibrium outcome.

The main result that Powell uses to establish the uniqueness of this PBE outcome is that a dissatisfied type of D would never reject S's initial offer in order to make a counteroffer. The logic behind the proof of this result is as follows (this is straight from Powell 1999, 248). Suppose that at least one dissatisfied type would do so in an equilibrium. Consider the lowest cost (i.e., most resolved) type of dissatisfied D that does this in equilibrium (this is also the lowest cost type of D among the satisfied as well as dissatisfied types who make a counteroffer). In the next period, S knows (through Bayesian updating on the equilibrium path) that this is the lowest-cost (i.e., most resolved) type that it could be facing. Thus, *the most favorable* (for D) behavior that S could be displaying in this subgame is if it believes with *certainty* that D is this lowest-cost (most resolved) type, and hence plays as though it is in a complete information game with this lowest-cost type. We already know from the complete information results that the best negotiated settlement that D could get in this subgame is the most resolved type's payoff-equivalent from war (if S's offer in the next period in this best-case scenario is accepted), or strictly less (if D's offer in the current period is accepted). Or, war might result in this second period or later. In any of these cases, the lowest-cost type of dissatisfied D is strictly better off going to war in the first period rather than rejecting an offer to make a counteroffer. Since this is true for any conceivable lowest-cost dissatisfied type that makes a counteroffer, this establishes that no dissatisfied type of D would reject an offer to make a counteroffer. It either accepts the initial offer (if it is as least as good as its expected utility from war) or goes to war. Because of this, S is essentially in the position of making its optimal TILI offer, because if it is rejected, war results, rather than a counteroffer. Thus, we are exactly back in Fearon's scenario.

To summarize, because Powell's model gives all of the bargaining leverage to S in equilibrium, D never rejects an offer to make a counteroffer, because in the next period, the best he can get is his utility from war (since S has all the bargaining leverage, i.e., gets all of the gains from avoiding war). And it is better to get the war payoff now rather than live with the (worse) status quo for the present and get (at best) the war payoff in the next period. Therefore, *every* perfect Bayesian equilibrium of Powell's model is of the risk-return tradeoff type, where D goes to war if it gets too low of an initial offer.

4 A Generalization of Powell's Model

We have thus argued that the reason that every incomplete information PBE equilibrium in Powell's model is of the risk-return tradeoff type is because that model gives all of the bargaining leverage to S. What is it in Powell's model that leads to this skewed distribution of bargaining leverage? The driving factor is that D cannot initiate a war in periods in which it makes an offer. Consider a period in which S makes an offer (we are now discussing the complete information case again). All it needs to offer D is its expected utility from war. D cannot credibly be demanding anything greater than that (unless it expects to get more in the next period, which in equilibrium it does not). So in any period in which S makes an offer, it gets all the gains from avoiding war.

Now consider a period in which D makes an offer. In the next period, S expects to get all of the gains from avoiding war, i.e., it just needs to offer D the amount $p - c_D$, and keep the rest of the pie for itself. Thus, in the current period S is demanding a share of the pie that makes it at least as well off as rejecting it, thereby getting the big payoff of $1 - q$ in the current period (if it rejects an offer) and the slightly lower (but still big) payoff of $1-(p-c_D)$ in the next period. Thus, S is demanding *even more* than all of the gains from avoiding war, in periods in which D makes an offer. If his offer is rejected, D has to live with the (awful) status quo in the current period and just his utility-equivalent from war in the next period. Thus, D chooses to offer the huge amount that leaves D with less than its utility from war.[10]

Thus, the driving factor here is that D cannot initiate a war in a period in which it makes an offer. If it could (and there is no reason why it should not be able to do so in an anarchic international environment; Waltz 1979), it would never choose to make an offer that leaves it with less than its payoff from war, because it would rather go to war instead.

To capture this, we generalize Powell's model by allowing an actor to initiate a war if its offer has been rejected. We also allow the actors to have different discount factors, δ_S and δ_D. The model is shown in Figure 5. Because D can now threaten to initiate a war if its offer is rejected, S cannot afford to be greedy and demand more than all of the gains from avoiding war in periods in which D makes an offer. In fact, S can no longer credibly demand more than its *own* payoff from war (thereby not getting *any* of the gains from avoiding war), if D is choosing to go to war if S rejects its offer. The following propositions describe SPE of this model, for different values of δ_D and δ_S, when D is dissatisfied.[11] (These propositions

[10]Technically, when the discount factors are the same (as in Powell's model), D is indifferent between making this offer and making a lower offer that results in agreement being reached in the next period, which is why there are an infinite number of SPE, which, however, all result in the same average per-period payoffs to the actors. This is all discussed in more detail in the appendix.

[11]These SPE are stationary up to the point that, at certain decision nodes, an actor may be indifferent among different courses of actions, and hence may be choosing different actions at histories which are different but which lead to structurally identical subgames. When δ_D is not too large, as in Propositions 1 and 2, we believe, but have not been able to prove, that these are all of the SPE. When δ_D is large, as in Proposition 3, then there exist genuinely non-stationary SPE, which are described in Proposition 4.

are stated in rather technical form, and the uninterested reader can skip to the discussion that follows.)

Proposition 1 *If* $\delta_D \leq \frac{(p-c_D)-q}{(p+c_S)-q}$, *then the following are SPE:*
(a) D always proposes $(x^*, 1-x^*)$, *where* $x^* = p+c_S$. *He always accepts any offer* $(y, 1-y)$ *such that* $y \geq p - c_D$. *In any period in which he gets a lower offer than this, he fights (does not say no). In any period in which S rejects his offer, he fights rather than passes.*
(b) S always proposes $(y^*, 1-y^*)$, *where* $y^* = p-c_D$. *She always accepts any offer* $(x, 1-x)$ *such that* $x \leq p+c_S$. *In any period in which she gets a worse offer, she is indifferent between fighting and saying no (since in the latter case D fights anyway), and hence can be choosing either (or mixing). In any period in which D rejects her offer, she passes rather than fights.*

Proposition 2 *If* $\frac{(p-c_D)-q}{(p+c_S)-q} \leq \delta_D \leq \frac{(p-c_D)-q}{\delta_D[(p+c_S)-q]}$, *then the following are SPE:*
(a) D always proposes $(x^*, 1-x^*)$, *where* $x^* = p+c_S$. *He always accepts any offer* $(y, 1-y)$ *such that* $y \geq q(1-\delta_D) + \delta_D(p+c_S)$. *In any period in which he gets a lower offer than this, he says no (does not fight). In any period in which S rejects his offer, he fights rather than passes.*
(b) In any period in which S makes a proposal, she (i) proposes $y^* = q(1-\delta_D) + \delta_D(p+c_S)$ *if* $\delta_S > \delta_D$, *(ii) proposes some* $y < y^*$ *if* $\delta_S < \delta_D$ *(this proposal is rejected and agreement is reached on* x^* *in the next period), and (iii) is indifferent between proposing* y^* *and proposing some* $y < y^*$ *if* $\delta_S = \delta_D$ *(this latter proposal is rejected and agreement is reached on* x^* *in the next period), and hence can be choosing either (or mixing). She always accepts any offer* $(x, 1-x)$ *such that* $x \leq p + c_S$. *In any period in which she gets a worse offer, she is indifferent between fighting and saying no (since in the latter case D fights anyway), and hence can be choosing either (or mixing). In any period in which D rejects her offer, she passes rather than fights.*

Proposition 3 *If* $\frac{(p-c_D)-q}{\delta_D[(p+c_S)-q]} \leq \delta_D$, *then the following are SPE:*
(a) D always proposes $(x^*, 1-x^*)$, *where* $x^* = \frac{(p-c_D)-q(1-\delta_D^2)}{\delta_D^2}$. *He always accepts any offer* $(y, 1-y)$ *such that* $y \geq \frac{(p-c_D)-q(1-\delta_D)}{\delta_D}$. *In any period in which he gets a lower offer than this,*

he says no (does not fight). In any period in which S rejects his offer, he fights with probability $\beta \in (0,1)$ *and passes with probability* $1-\beta$*, where (i)* $\beta = \frac{(1-\delta_S\delta_D)[(p-c_D)-q]}{\delta_D\{\delta_D[(p+c_S)-q]-\delta_S[(p-c_D)-q]\}}$ *if* $\delta_S > \delta_D$*, (ii)* $\beta = \frac{(1-\delta_S^2)[(p-c_D)-q]}{\delta_D^2[(p+c_S)-q]-\delta_S^2[(p-c_D)-q]}$ *if* $\delta_S < \delta_D$*, and (iii)* $\beta = \frac{(1-\delta^2)[(p-c_D)-q]}{\delta^2(c_D+c_S)}$ *if* $\delta_D = \delta_S = \delta$*. (Note that* $\beta \to 1$ *from below as* $\delta_D \to \frac{(p-c_D)-q}{\delta_D[(p+c_S)-q]}$ *from above, and that* $\beta \to 0$ *from above as* $\delta_D, \delta_S \to 1$ *from below.)*
(b) In any period in which S makes a proposal, she (i) proposes $y^* = \frac{(p-c_D)-q(1-\delta_D)}{\delta_D}$ *if* $\delta_S > \delta_D$*, (ii) proposes some* $y < y^*$ *if* $\delta_S < \delta_D$ *(this proposal is rejected and agreement is reached on* x^* *in the next period), and (iii) is indifferent between proposing* y^* *and proposing some* $y < y^*$ *if* $\delta_S = \delta_D$ *(this latter proposal is rejected and agreement is reached on* x^* *in the next period), and hence can be choosing either (or mixing). She always accepts any offer* $(x, 1-x)$ *such that* $x \leq \frac{(p-c_D)-q(1-\delta_D^2)}{\delta_D^2}$*. In any period in which she gets a worse offer, she says no (does not fight). In any period in which D rejects her offer, she passes rather than fights.*

These results are illustrated graphically in Figure 6 for the case where $\delta_D = \delta_S = \delta$. The figure shows the equilibrium proposals for D, x^* and y^*, when D and S make proposals, respectively.[12] These proposals are accepted in equilibrium, and so these are D's actual average per-period payoffs in the model, depending on who gets to make the first proposal.

When δ is low, each side's proposal just offers the other side its payoff from war, i.e., whoever gets to make the first proposal gets *all* of the gains from avoiding war. Note that in this range, D finds it optimal to fight rather than move to the next period, if his proposal is rejected or if S makes a small offer. Because D optimally chooses to fight if given a low offer, he cannot credibly demand more than his payoff from war when S is making a proposal, and hence S just needs to offer D his utility-equivalent from war. Similarly, because D chooses to fight if S rejects his offer, S cannot credibly be demanding more than her payoff from war when D is making a proposal (this is in contrast to Powell's model, in which D cannot choose to fight at this stage and so S can be greedy and demand more), and so D just offers S her utility-equivalent from war.

[12]Figure 6 is drawn for $p = 0.5$, $c_D = c_S = 0.2$, and $q = 0.1$. It has the same general shape for all values of these parameters.

When δ gets high enough, however, then D is sufficiently patient that he prefers to move to the next period (and get all of the gains from avoiding war therein) if S makes a low offer, rather than go to war. Therefore, in periods where S makes an offer, D is now demanding his average payoff for getting the status quo in the current period and all of the gains from avoiding war in the next period. Since this average payoff is now (i.e., with δ high enough) greater than his payoff from war, D is demanding more than his payoff-equivalent from war, and so S now has to compromise when she makes a proposal. Thus, as seen in the figure, y^* (S's proposal for D) is increasing in this range. However, S still gets some of the gains from avoiding war when she makes a proposal, i.e., $y^* < p + c_S$. In this range, δ is still low enough that D prefers to go to war if S rejects his offer (because D gets the relatively small payoff of y^* in the next period, and with sufficient discounting war is preferred to living with the status quo in the current period and not getting some of the gains from avoiding war y^* until the next period), and so S cannot credibly demand more than her utility-equivalent from war. Therefore, D is still getting all of the gains from avoiding war when he makes a proposal (i.e., $x^* = p + c_S$).

When δ gets even larger, however, then the credibility of D's threat to go to war if S rejects D's offer starts diminishing (the offer is never rejected, and hence this is an off-the-equilibrium path threat). This is because D will get at least some of the gains from avoiding war in the next period (recall that $y^* > p - c_D$), and hence for δ large enough, D prefers to live with the status quo in the current period and get some of the gains from avoiding war in the next period rather than going to war in the current period, if S rejects his offer. However, this diminished credibility means that S can be demanding more than just her payoff-equivalent from war, and hence D starts compromising when he makes a proposal (i.e., x^* starts decreasing). And because D gets a smaller payoff in periods in which he makes a proposal, S can offer less to D in periods where she makes a proposal (and, hence, y^* also starts decreasing). In this range, D is probabilistically choosing between fighting and moving to the next period if S rejects his proposal, and as δ increases, the probability with which D choose to fight (β in Proposition 3) decreases.

Thus, as seen in Figure 6, the three odd results that emerge under complete information in Powell's model do not emerge in the generalized model. First, there is a first-mover advantage, which is typical in Rubinstein-type (1982) bargaining models with complete information. Each side strictly prefers to make the first proposal. Second, each side has genuine bargaining leverage in equilibrium, which we would expect in a model in which both sides can make proposals. When δ is low, each side gets all of the gains from avoiding war when it makes a proposal. When δ is in the medium range, D gets all of the gains from avoiding war when it makes a proposal, and both sides get some of the gains when S makes a proposal. When δ is large, both sides get some of the gains from avoiding war, regardless of who gets to make the first proposal. If the two sides are uncertain of who will get to make the first proposal (e.g., if they randomly decide), then each side expects to do strictly better from negotiations than from war, regardless of the size of the discount factor. Finally, and perhaps most importantly from a substantive viewpoint, the agreement reached gives each side at least its utility from war, which we would intuitively expect given that under anarchy, a state can launch a war at any time.

So far, we have been discussing the case where the two sides have the same discount factor, i.e., where $\delta_D = \delta_S = \delta$. However, now suppose that the discount factors can differ. Figure 6 then shows D's equilibrium proposal for itself, x^* (which is accepted), and y^* is now the minimal demand made by D in periods where S makes an offer, but not necessarily the *actual* offer that S makes. δ on the horizontal axis should now read δ_D. When δ_D is low, then D goes to war if S makes a low offer, and hence S always offers y^* (i.e., regardless of the value of δ_S). However, once δ_D gets in the medium or high range, then D says no if S makes a low offer, rather than going to war, for the reasons discussed above. In the medium or high range, if $\delta_S > \delta_D$, then S offers y^* (which is accepted) whenever she makes a proposal. However, if $\delta_S < \delta_D$, i.e., if S is sufficiently impatient, then S proposes some $y < y^*$, which is rejected, and agreement is reached on x^* in the next period.[13] What is the intuition behind this delaying behavior? S prefers the agreement y^* to x^*. However, S prefers the status

[13]If $\delta_S = \delta_D$, then S is indifferent, and hence can be choosing either, or mixing.

quo even more. If S is sufficiently impatient, then she cares primarily about her payoff in the current period, and hence she would rather get the status quo for one more period, even though that means that in the long run, she lives with a less favorable agreement (x^*) than the one that she could have gotten (y^*). In crisis bargaining in the shadow of power, an impatient satisfied state may delay reaching an agreement and ultimately accept a less favorable agreement in order to enjoy the benefits of the very favorable status quo for a little longer. However, this can only occur if the dissatisfied state is sufficiently patient that it accepts that delay rather than going to war immediately if S makes a low offer.[14] To the best of our knowledge, this is the only formal model of crisis bargaining in which an agreement can be reached in equilibrium after some delay, without conflict occurring in the meantime.[15]

So far, we have been discussing the stationary SPE (which we believe are the unique ones). However, before moving on to incomplete information, note that, for δ_D and δ_S sufficiently high, a continuum of agreements can be reached in non-stationary SPE. This is established in the following folk-theorem type result.

Proposition 4 *If* $\frac{(p-c_D)-q}{\delta_D[(p+c_S)-q]} \leq \delta_D$ *(the same condition for Proposition 3) and* $\delta_S = \delta_D$*, then, in the model where* D *makes the first offer, any agreement that gives* D *between* $p - c_D$ *and* $p + c_S$ *can be reached in the first period of a non-stationary SPE, and in the model where* S *makes the first offer, any agreement that gives* D *between* $p - c_D$ *and* $q(1 - \delta_D) + \delta_D(p + c_S)$ *(this is the linearly increasing portion of the dotted line in Figure 6) can be reached in the first period of a non-stationary SPE.*

[14]This points to another interesting feature of crisis bargaining over an already-divided good, namely that delay is not inherently costly. In the Rubinstein (1982) model, because the actors get no utility until an agreement is reached, delay is inherently costly in the sense that, for any given agreement that gives each side at least some of the pie, each actor strictly prefers to reach that agreement sooner rather than later. In this model, because there is already a status quo division of the disputed good, delay is not necessarily costly. For any given agreement that is strictly better than the status quo for one actor (which is of course strictly worse for the other actor), that actor strictly prefer to reach that agreement sooner rather than later, but strictly prefers the opposite for any agreement that is worse than the status quo.

[15]In recent models of intra-war bargaining with incomplete information, an agreement can be reached in equilibrium after some delay; however, conflict occurs in the meantime. In the models of Filson and Werner (2002), Slantchev (2003), and Smith and Stam (2004), conflict *must* occur each time an offer is rejected. In Powell's (2004a) model, an offer can be rejected without conflict occurring, but this never happens in equilibrium. In Morrow's (1989) model, an agreement can be reached in the second period in equilibrium, without conflict occurring before then; however, in that model, there is no *opportunity* for conflict to occur before then, as there is in our model.

To understand how these non-stationary SPE are generated, note that, when δ_D is high, then in the stationary SPE of Proposition 3, D is mixing (randomly choosing) between fighting and passing, at decision nodes at which S has said no to D's offer. D can be mixing because he is indifferent (given the equilibrium strategies for the rest of the game), and hence can be choosing either option. In the stationary SPE of Proposition 3, he is randomizing with the same probability (fight with probability β, pass with probability $1-\beta$) at *each* such decision node. However, suppose that at the *first* such decision node, he chooses to fight with certainty, and then uses the stationary strategy (of Proposition 3) in subsequent periods. Because D chooses to fight with certainty if his first offer is rejected, S can credibly demand no more than her payoff from war in that period, and hence D gets all of the gains from avoiding war. This is the upper bound of what D can achieve in a non-stationary SPE when δ_D is high. On the other hand, suppose that D chooses to *pass* with certainty at the first such decision node, and then uses the stationary strategy in subsequent periods. This allows S to demand a lot more in that period, and hence D's share of the pie is lower. This establishes the lower bound of what D can achieve in a non-stationary SPE when δ_D is high. Any payoff between these two extreme values can also be achieved in a non-stationary SPE, by supposing that at the first such decision node, D chooses to fight with the appropriate probability.

The above method of generating the non-stationary SPE also suggests an equilibrium selection argument for which equilibrium is most likely to be played. D's payoff is maximized if he chooses to fight with certainty at the first such decision node. Both players recognize this, and hence it seems reasonable that this is where their expectations will converge. Thus, it can reasonably be argued that the equilibrium most likely to be played when δ_D is high is the one in which the trend in Figure 6 when δ_D is medium continues even when δ_D becomes large, i.e., D continues to get all of the gains from avoiding war when he makes the first proposal, and S compromises more and more (the linearly increasing dotted line) as δ_D becomes larger and larger when she makes the first proposal. This equilibrium selection argument will also be referred to in the incomplete information results.

5 Incomplete Information

We now turn to crisis bargaining under incomplete information. Recall that, in Powell's (1996a, 1996b, 1999) model, there is a *unique* perfect Bayesian equilibrium outcome in which, if S's initial offer is too small, D goes to war rather than makes a counteroffer. Thus, if war occurs due to incomplete information, it is because of the risk-return tradeoff: S trades off a positive risk of war for a greater return at the negotiating table if war does *not* occur. In this section, we want to examine whether this is the same causal mechanism that leads to war in the generalized model.

To examine this, we consider a case of one-sided uncertainty, in which S is uncertain about D's cost of war c_D. We assume that D's cost of war takes on one of two values. S believes that D's cost is c_{D_l} with probability $1 > s > 0$ and c_{D_h} with probability $1 - s$, with $c_{D_l} < c_{D_h}$, i.e., c_{D_l} is the more resolved (low-cost) type, because its expected utility from war is higher (see Figure 2). We assume that both types of D are dissatisfied, i.e., suppose that $q < p - c_{D_h}$. Finally, consider the model in which S makes the first offer.

We first show that, as in Fearon and Powell's models, there exist risk-return tradeoff equilibria, in which war occurs if S makes a small initial offer and D ends up being the highly resolved type. These equilibria exist when δ_D is low, medium, or high. We discuss the intuition behind them, and point out how they differ from Powell's equilibria. We then show that, unlike in Powell's model, when δ_D is sufficiently high, there also exist peaceful, "non-risk" equilibria in which D makes a counteroffer (which is accepted) if S makes too small an initial offer. We then discuss the implications of these findings.

5.1 Risk-Return Tradeoff Equilibria

Proposition 5 *If $\delta_D \leq \frac{(p-c_{D_h})-q}{(p+c_S)-q}$, there is a perfect Bayesian equilibrium (PBE) in which, in the first period, type c_{D_l} accepts all offers $(y, 1-y)$ such that $y \geq p - c_{D_l}$ and goes to war for any lower offer, and type c_{D_h} accepts all offers $(y, 1-y)$ such that $y \geq p - c_{D_h}$ and goes to war for any lower offer. If $s \geq s_{critical}$, where $s_{critical} = \frac{c_{D_h}-c_{D_l}}{c_{D_h}+c_S}$, then S makes the large initial offer $y^* = p - c_{D_l}$, which both types accept, and war is avoided. If $s \leq s_{critical}$,* *then S makes the small initial offer $y^* = p - c_{D_h}$, which only type c_{D_h} accepts. Type c_{D_l} rejects it and goes to war instead. If the second period is reached in this equilibrium (this is off-the-equilibrium path behavior), agreement would be reached on $x^* = p + c_S$.*

Proposition 6 *If $\frac{(p-c_{D_h})-q}{(p+c_S)-q} \leq \delta_D \leq min\{\frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}, \frac{(p-c_{D_l})-q}{(p+c_S)-q}\}$, there is a perfect Bayesian equilibrium (PBE) in which, in the first period, type c_{D_l} accepts all offers $(y, 1-y)$ such that $y \geq p - c_{D_l}$, and goes to war for any lower offer, and type c_{D_h} accepts all offers $(y, 1-y)$ such that $y \geq q(1-\delta_D) + \delta_D(p + c_S)$ and says no (rather than fight) for any lower offer. If s is sufficiently high, S makes the large initial offer $y^* = p - c_{D_l}$, which both types accept, and war is avoided. If s is sufficiently low, then S makes a low initial offer of $y^* = q(1-\delta_D)+\delta_D(p+c_S)$ if $\delta_S \geq \delta_D$ and some even lower offer if $\delta_S \leq \delta_D$. Type c_{D_l} rejects these low offers and goes to war. If the second period is reached, agreement is reached on $x^* = p + c_S$.*

Proposition 7 *If $\delta_D \geq max\{\frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}, \frac{(p-c_{D_h})-q}{(p-c_{D_l})-q}\}$, there is a perfect Bayesian equilibrium (PBE) in which, in the first period, type c_{D_l} accepts all offers $(y, 1-y)$ such that $y \geq p - c_{D_l}$, and goes to war for any lower offer, and type c_{D_h} accepts all offers $(y, 1-y)$ such that $y \geq \frac{(p-c_{D_h})-q(1-\delta_D)}{\delta_D}$ and says no (rather than fight) for any lower offer. If s is sufficiently high, S makes the large initial offer $y^* = p - c_{D_l}$, which both types accept, and war is avoided. If s is sufficiently low, then S makes a low initial offer of $y^* = \frac{(p-c_{D_h})-q(1-\delta_D)}{\delta_D}$ if $\delta_S \geq \delta_D$ and some even lower offer if $\delta_S \leq \delta_D$. Type c_{D_l} rejects these low offers and goes to war. If the second period is reached (by type c_{D_h}), agreement is reached on $x^* = \frac{(p-c_{D_h})-q(1-\delta_D^2)}{\delta_D^2}$.*

We have thus constructed PBE in which the risk-return tradeoff emerges, whether δ_D is low, medium, or high. What is interesting is that the substantive dynamics behind the risk-return tradeoff equilibria are quite different from Powell's, and they differ depending on whether δ_D is low, medium, or high. Recall that in Powell's model, *no* dissatisfied type of D rejects an offer in order to make a counteroffer. It either accepts the initial offer (if it is as least as great as its utility from war), or goes to war. When δ_D is low, this is what happens in our model as well. However, the reason is quite different. In Powell's model, because S has all the bargaining leverage in equilibrium, the best the lowest-cost type of D can get in the second period is its utility from war, and this is the reason why it never makes a counteroffer. In our equilibrium, however, agreement would be reached on $x^* = p + c_S$ in the second period, i.e., D would get *all* of the gains from avoiding war. This is why δ_D has to be sufficiently low in this equilibrium. For D to go to war instead of waiting until then, he has to be so impatient that he would rather get the war payoff in the current period rather than the (worse) status quo, even though that means that he forsakes getting all of the gains from avoiding war (the agreement x^*) from the next period onward. When δ_D is low, if war occurs due to incomplete information (because of the risk-return tradeoff), it is because the dissatisfied state is too impatient to wait one period to get all of the gains from avoiding war.

Now consider when δ_D is in the medium range. In this equilibrium, only the highly-resolved (low cost) type of D goes to war if he gets a low initial offer. The less-resolved (higher cost) type makes a counteroffer in the next period (which is accepted) if he gets a low offer. This equilibrium shows that in the generalized model, and unlike in Powell's model, it is possible for a dissatisfied type of D to reject an offer in order to make a counteroffer, rather than go to war.[16]

What is the intuition behind δ_D having to be medium? Again, notice that in this equilibrium, the agreement $x^* = p + c_S$ would be reached in the second period, i.e., D would get *all* of the gains from avoiding war. Thus, D would only go to war rather than get all of the gains from avoiding war (in the next period, if he gets a low initial offer) if he is so impatient that he would rather go to war than live with the (worse) status quo in the current period, even though that means forsaking getting all of the gains from avoiding war from the next period onward. Because the two types have different payoffs from going to war, they have different thresholds for δ_D below which they would rather go to war than wait and get all of the gains from avoiding war from the next period onward. In this equilibrium, δ_D is

[16]Powell's proof does not depend on there being a continuum of types, as is the case in his analysis. That is, Powell's result would hold in his model even if there are only two types of D, as in our analysis. That is to say, our finding that a dissatisfied type of D can reject an offer in order to make a counteroffer is not because we only consider two types, but because we consider a more general model that gives both sides bargaining leverage in equilibrium.

below the highly-resolved type's threshold (who thus would rather go to war), but above the less-resolved type's threshold (who thus would rather move to next period if he gets a low offer).

Finally, consider the equilibrium when δ_D is high. This is similar to the previous one in that only the highly-resolved type goes to war if he gets a low initial offer — the less-resolved type makes a counteroffer (which is accepted) if he gets a low offer. However, unlike the previous equilibria, D does not get all of the gains from avoiding war in the second period. As seen in Figure 6 and from our discussion of the complete information results, when δ_D gets high, in the stationary SPE of Proposition 3, the credibility of D's threat to use force diminishes, because he would rather move to the next period and get at least some of the gains from avoiding war therein rather than going to war. Because the credibility of his threat to use force is diminishing, both x^* and y^* (under complete information) start decreasing. In the incomplete information equilibrium when δ_D is high, what is happening is that δ_D is so high that the most favorable (for D) agreement that could be reached in the second period is actually less than the highly-resolved type's payoff from war, but exceeds the low-resolve type's payoff from war. Thus, even though δ_D is high, the highly-resolved type prefers to go to war rather than move to the next period (in which he could at best get a worse agreement than war), whereas the low-resolve type prefers to move to the next period and get some of the gains from avoiding war therein.

5.2 Peaceful ("Non-Risk") Equilibria

Proposition 8 *If* $\frac{(p-c_{D_l})-q}{(p+c_S)-q} \leq \delta_D \leq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$, *then for any value of s, there is a PBE in which, in the first period, both types of D accept all offers* $(y, 1-y)$ *such that* $y \geq q(1-\delta_D) + \delta_D(p+c_S)$ *and say no (rather than fight) for any lower offer. If* $\delta_S \geq \delta_D$, *then S offers* $y^* = q(1-\delta_D) + \delta_D(p+c_S)$ *in the first period, which both types accept. If* $\delta_S \leq \delta_D$, *then S offers some* $y < y^*$, *and agreement is reached on* $x^* = p + c_S$ *in the second period.*

Proposition 9 *If* $\delta_D \geq max\{\frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}, \frac{(p-c_{D_l})-q}{(p+c_S)-q}\}$, *then for any value of s, there is a PBE in which, in the first period, both types of D accept all offers* $(y, 1-y)$ *such that*

$y \geq q(1-\delta_D) + \delta_D(p + c_S)$ and say no (rather than fight) for any lower offer. If $\delta_S \geq \delta_D$, then S offers $y^* = q(1-\delta_D) + \delta_D(p+c_S)$ in the first period, which both types accept. If $\delta_S \leq \delta_D$, then S offers some $y < y^*$, and agreement is reached on $x^* = p + c_S$ in the second period.

These are equilibria under incomplete information, when δ_D is sufficiently large, in which the risk-return tradeoff does *not* emerge. That is, these are equilibria in which even the highly-resolved (low cost) type makes a counteroffer (which is accepted) if he gets a low initial offer, rather than going to war (again, this is behavior that cannot emerge in Powell's model). What is the intuition behind these results? We have constructed equilibria in which agreement would be reached on $x^* = p + c_S$ in the second period, i.e., D would get *all* of the gains from avoiding war. Thus, as long as D is sufficiently patient, he prefers to make a counteroffer rather than go to war if he gets a low initial offer. The risk-return tradeoff is the *unique* PBE outcome in Powell's model only because that model gives all of the bargaining leverage to S. When D also has bargaining leverage, as in the generalized model, there exist equilibria when δ_D is sufficiently high in which D finds it worthwhile to make a counteroffer rather than go to war, and in which the probability of war is therefore zero.

5.3 Implications

What are the implications of these results? First, the risk-return tradeoff, and hence the risk of war under incomplete information, is not as omnipresent as Fearon and Powell's models suggest. Under incomplete information, there also exist non-risk equilibria in which the probability of war is zero even when S's initial belief puts sufficient weight on D being a low-resolve type that she is tempted to make a small initial offer. Under incomplete information, S does not necessarily face a hair-trigger decision in the initial period, as Fearon and Powell's models suggest.

Perhaps more importantly, these results suggest that the causal mechanism by which incomplete information leads to war is different than that implied by previous work. First consider the case where D is very impatient. When δ_D is low, then *only* risk-return tradeoff equilibria exist. This is because, even if D gets *all* of the gains from avoiding war in the second period (the best he can possibly do), he would rather go to war in the first period than wait until then (and live with the awful status quo in the current period), for δ_D sufficiently low. So, if S's initial belief puts sufficient weight on D being the low-resolve type that she makes a low initial offer, war occurs if D ends up being the high-resolve type. However, it is not quite correct to say that private information is *the* cause of war in this scenario. The only effect of private information *per se* is to delay the reaching of an agreement until the second period. The only reason why this leads to war is because D is too impatient to wait until then. Thus, when δ_D is low, it is private information *plus* impatience that is the cause of war.

Now consider when δ_D is medium or large. In this case, there are risk-return as well as non-risk equilibria. War can only occur if the actors are using the strategies of the risk-return equilibria. Thus, it is not quite correct to say that incomplete information is *the* cause of war, because empirically explaining the outbreak of war in a particular case (when δ_D is medium or high) would also require us to explain why the actors were using the strategies of the risk-return equilibrium rather than the non-risk equilibrium. A complete explanation of war arising due to incomplete information must account for either impatience (when δ_D is low) or an equilibrium selection issue (when δ_D is medium or high).

Regarding the equilibrium selection issue, a natural question to ask is whether the risk-return equilibria Pareto-dominate the non-risk equilibria, or vice-versa, when they both exist. If so, there would be good reasons for expecting the Pareto-dominant equilibrium to be played. However, it turns out that neither equilibrium Pareto-dominates the other. When δ_D is sufficiently high, the equilibria of Propositions 7 as well as 9 exist. In Proposition 7, when s is low, S makes a low initial offer that is only accepted by the less-resolved type — the high-resolve type rejects it and goes to war. Thus, war occurs with probability s, and an agreement is reached with probability $1 - s$. As δ_D converges to 1, this agreement that is reached converges to $p - c_{D_h}$, i.e., S gets all of the gains from avoiding war against the less-resolved type. In Proposition 9, S makes an offer that both types accept. As δ_D

converges to 1, this agreement converges to $p + c_S$, i.e., D (both types) gets all of the gains from avoiding war. Thus, as δ_D converges to 1, both types of D strictly prefer the non-risk equilibrium of Proposition 9, whereas for s sufficiently low, S prefers the risk-return equilibrium of Proposition 7, in which she is increasingly likely to get all of the gains from avoiding war, and less likely to have to go to war. Thus, for δ_D high and s low, both types of D strictly prefer the non-risk equilibrium, in which S compromises a lot, whereas S prefers the risk-return equilibrium, in which she just has to offer the low-resolve type of D its payoff-equivalent from war, and war is very unlikely. Neither equilibrium Pareto-dominates the other.

However, there are other reasons for expecting the non-risk equilibrium to be played when δ_D is high, rather than the risk-return equilibrium. We saw from Proposition 4 that when δ_D is high, a broad continuum of agreements can be reached in a non-stationary SPE. In the *stationary* SPE of Proposition 3, as δ_D approaches 1, D's share of the pie approaches $p - c_D$ (i.e., D does not get *any* of the gains from avoiding war), regardless of who makes the first proposal. In constructing the risk-return equilibrium when δ_D is high (Proposition 7), we have to use this stationary SPE in the second period. That is, when δ_D is high, the only way we can get the highly-resolved type of D to go to war rather than make a counteroffer if S's initial offer is too small, is by constructing an equilibrium in which, if the second period is reached, the highly-resolved type of D does not get *any* of the gains from avoiding war. As long as D gets *some* of the gains from avoiding war in the second period, the risk-return tradeoff cannot exist for δ_D high enough, because D would rather move to the second period and get those gains (no matter how small) than go to war in the first period. To put it another way, constructing a risk-return equilibrium when δ_D is very high requires us to use a SPE in which D does not get *any* of the gains from avoiding war in the second period, and there are many SPE in which he *does* get some of the gains from avoiding war. Moreover, we argued earlier that the SPE most likely to be played when δ_D is high is one in which D in fact gets a substantial amount of the gains from avoiding war. If this SPE is indeed the one expected to be played if the second period is reached, then the risk-return tradeoff cannot exist for δ_D high enough.

6 Conclusion

It is probably a fair statement to say that most students of international relations agree that private information and incentives to misrepresent it comprise a sound rationalist explanation for war, even when it is commonly known that there exist negotiated settlements that both sides strictly prefer to war. Indeed, virtually all of the formal work being done today on the informational problems associated with war focus on how information conveyed on the battlefield and through bargaining behavior *during* the course of a war can help resolve the informational problems that can lead to the rational *outbreak* of war (e.g., Filson and Werner 2002; Morrow 1989; Powell 2004a; Slantchev 2003; Smith and Stam 2004; Wagner 2000; Wittman 1979).

In this paper, we take a step back to reexamine the soundness of the informational rationale for the *outbreak* of war, which essentially suggests that due to incentives to misrepresent one's private information, pre-war bargaining can do little in the presence of informational assymetries. In formal terms, the informational rationale is captured by the risk-return tradeoff: when the satisfied state is uncertain about the dissatisfied side's resolve for war, if S is sufficiently confident that D is a less-resolved type, then S's optimal offer entails some risk of rejection, and hence, war. However, it is not clear from an intuitive point of view why D would necessarily go to war if he receives a low initial offer, rather than make a counteroffer, especially when it is commonly known that there exist negotiated settlements that both sides strictly prefer to war. In Fearon's (1995) model, rejection *by assumption* means war, and in Powell's (1996a, 1996b, 1999) model, in equilibrium rejection *effectively* means war rather than a counteroffer. Thus, the risk-return tradeoff appears to be very robust, and the complete explanation for how war arises under incomplete information. However, we have shown here that this is simply because Powell's model gives all of the bargaining leverage to S, effectively ruling out any incentive for D to make a counteroffer.[17]

[17]The risk-return tradeoff is also built in to recent models of intra-war bargaining. For example, in the

We generalized Powell's canonical model of crisis bargaining in the shadow of power and showed that, in the generalized model, both sides have bargaining leverage in equilibrium. And because of that, the risk-return tradeoff is the unique incomplete information PBE outcome only when the dissatisfied state is so impatient that even if he would get *all* of the gains from avoiding war in the next period, he prefers to go to war now rather than live with the (worse) status quo in the current period and have to wait until the next period to start getting those gains. When the dissatisfied state is sufficiently patient, there also exist equilibria in which he makes a counteroffer and thereby gets some of the gains from avoiding war, if he gets a low initial offer, and in which the probability of war is therefore zero. Moreover, when the dissatisfied state is sufficiently patient, there are good reasons to expect such a "non-risk" equilibrium to be played rather than a risk-return equilibrium.

This analysis suggests that empirically explaining the outbreak of war due to private information becomes more complicated than previous work has implied. When δ_D is low, if war breaks out, it is because (i) there is private information and incentives to misrepresent it, (ii) S's initial belief about D's type causes her to make a small initial offer, (iii) D turns out to be a highly-resolved type, and (iv) D is too impatient to wait until the next period to get some of the gains from avoiding war. When δ_D is high, if war breaks out, it is because conditions (i)-(iii) above are satisfied, and (iv) for some reason, the actors are using the strategies of the risk-return rather than the non-risk equilibrium. The causal mechanisms that lead to war under incomplete information are more complex than previous work has implied, and the very idea that private information (plus incentives to misrepresent it) *by itself* comprises a sound rationalist explanation for war is not quite correct.

---

models of Filson and Werner (2002), Slantchev (2003), and Smith and Stam (2004), conflict *must* occur every time an offer is rejected, and hence in the first period, state 1 faces the risk-return tradeoff (in Filson and Werner's model, state 1 has the additional option of quitting the crisis if its first offer is rejected, but conflict must occur before another offer can be made). In Powell's (2004a) model, a state can reject an offer and peacefully move to the next period. Indeed, this model also allows each state to go to war in any period, as in the model we analyze here. However, only S is allowed to make offers, which gives all the bargaining leverage to S, and hence D never peacefully moves to the next period in equilibrium. Thus, the risk-return tradeoff is built in by only allowing S to make offers. This suggests that while these models are well-suited for examining how events on the battlefield and bargaining behavior during the course of war can help reveal information *if* war occurs, they are not ideally suited for examining the rationalist *outbreak* of war. A next-generation intra-war bargaining model would relax the war-as-a-costly-lottery assumption (which we maintain), but not *build in* the risk-return tradeoff.

To put it another way, when incomplete information leads to war, it is either because (a) the uncertainty leads to delay in reaching a preferred-to-war negotiated settlement and the dissatisfied state is too impatient to wait until then, or (b) the dissatisfied state is not too impatient, but for some reason the states are using the strategies of the risk-return rather than the non-risk equilibria (in particular, they are using strategies in which, if the second period is reached, the highly-resolved type of dissatisfied state does not get *any* of the gains from avoiding war). This is a significant modification of our understanding of how incomplete information leads to war. Finally, to the best of our knowledge, this is the only formal model of crisis bargaining in which an agreement can be reached in equilibrium after some delay, without conflict occurring in the meantime. In short, our results suggest that the consensus that has seemingly developed on the soundness of the informational rationale for war is perhaps premature, and that more valuable research can be done on this topic.

In addition to its theoretical implications for the causal mechanisms by which incomplete information leads to war, this analysis also has potential policy implications for conflict resolution/prevention strategies. There is a growing literature that examines how third party actors, such as international organizations or foreign mediators, can help resolve or prevent bilateral conflict. Since Fearon's (1995) seminal article on rationalist explanations for war, a lot of this work examines how third party actors can help disputants overcome informational asymmetries or make credible commitments, Fearon's two primary rationalist explanations (e.g., Kydd 2003; Walter 2002). Our analysis suggests additional methods, if perfect information revelation cannot be achieved. First of all, third party actors should help the disputants recognize the possibility of making counteroffers, if previous offers are not acceptable. In a TILI offer bargaining setting, the risk-return tradeoff is the unique equilibrium outcome under incomplete information. As we have shown here, however, in an offer-counteroffer bargaining setting, non-risk equilibria also exist. Second, third party actors can help the disputants to have a longer time horizon (i.e., be more patient). Perhaps the most practical way of doing this is to structure the bargaining so as to reduce the time between counteroffers, or between negotiating rounds (which effectively increases the actors' discount factors). Finally, third party actors should help emphasize agreements that give both sides some of the gains from avoiding war. In an interesting article, Kydd (2003) uses a game-theoretic model to show that if conflict occurs because of private information and incentives to misrepresent it, then biased mediators reduce the probability of conflict more than unbiased mediators. Kydd builds on Fearon's (1995) TILI offer bargaining model to establish his argument. We have shown here that in a more general bargaining setting, when impatience is not a factor, the only way incomplete information can lead to war is if the dissatisfied side does not expect to get *any* of the gains from avoiding war in the future. By placing emphasis on agreements that give both sides some of the gains from avoiding war, the likelihood of conflict is reduced, and it seems that unbiased mediators would be better willing to play this role.

7 Appendix

7.1 Powell's Model

Powell (1996a, 263) identifies one (stationary) SPE of his model and states that it is the unique SPE. However, because of an indifference condition (when the discount factors are equal, i.e., when $\delta_S = \delta_D$, as in Powell's model), there are in fact an infinite number of SPE (however, the average per-period payoffs to the players are the same in all of these equilibria). Powell refers the reader to Osborne and Rubinstein's (1990) analysis of the Rubinstein (1982) model with outside options for a proof. However, the models are not quite equivalent, since there are no inside options in Osborne and Rubinstein's (1990) model. Below, we characterize SPE in Powell's model, while generalizing the model to allow the players to have different discount factors — Powell's SPE is the one in which $\delta_S = \delta_D$ and D chooses to always satisfy S's minimal demand when making a proposal.

Proposition 10 *The following are SPE when D is dissatisfied, i.e., when $q < p - c_D$.*
(a) In any period in which D makes a proposal, he (i) proposes $x^* = q(1-\delta_S) + \delta_S(p - c_D)$ if $\delta_D < \delta_S$, (ii) proposes some $x > x^*$ if $\delta_D > \delta_S$ (this proposal is rejected and agreement is reached on y^* in the next period), and (iii) is indifferent between proposing x^* and proposing some $x > x^*$ if $\delta_D = \delta_S$ (this latter proposal is rejected and agreement is reached on y^* in the next period), and hence can be choosing either (or mixing). He always accepts any offer $(y, 1-y)$ such that $y \geq p - c_D$. In any period in which he gets a lower offer than this, he fights (does not continue to the next period).
(b) S always proposes $(y^*, 1-y^*)$ where $y^* = p - c_D$, and always accepts any offer $(x, 1-x)$ such that $x \leq q(1-\delta_S) + \delta_S(p - c_D)$. In any period in which she gets a worse offer than this, she continues to the next period (does not fight).

Proof: First consider S's decisions. Given D's acceptance rule, S is strictly best off proposing $y^* = p - c_D$ whenever she makes a proposal (if S proposes a lower y, D chooses war, which is strictly worse for S). Now consider periods in which D makes an offer. We have just shown that S's optimal continuation value for moving to the next period is $(1-q) + \frac{\delta_S(1-y^*)}{1-\delta_S}$. If she goes to war instead, her payoff is $\frac{1-p-c_S}{1-\delta_S}$. It is easy to show that the former is strictly greater than the latter, and hence S cannot credibly reject any proposal $(x, 1-x)$ such that $\frac{1-x}{1-\delta_S} \geq (1-q) + \frac{\delta_S(1-y^*)}{1-\delta_S}$, or $x \leq q(1-\delta_S) + \delta_S(p-c_D)$, and she must move to the next period rather than fight if she gets a worse offer.

Now consider D's decisions. Given S's acceptance rule, the best (for himself) acceptable proposal that D can make in a period in which he makes a proposal is $x^* = q(1-\delta_S)+\delta_S(p-c_D)$. (Because this is strictly greater than q, if that agreement will eventually be reached, it is strictly better to reach it now rather than in a later period.) D's other option is to propose some $x > x^*$, which is rejected, and in the next period, S proposes $y^* = p - c_D$, which D can (i) accept, (ii) reject it and fight, or (iii) reject it and make a counteroffer. In case (iii), D is back in the same scenario that he is in the current period, except that he has received q for two periods. Since $x^* > q$, offering x^* in the current period is strictly better than ending up in case (iii). Cases (i) and (ii) give D the same total payoff of $q + \frac{\delta_D(p-c_D)}{1-\delta_D}$. Therefore, D is strictly best off offering x^* in the current period rather than some $x > x^*$ if and only if $\frac{x^*}{1-\delta_D} > q + \frac{\delta_D(p-c_D)}{1-\delta_D}$, which can be shown to hold if and only if $\delta_D < \delta_S$. If $\delta_D > \delta_S$, then D strictly prefer to offer some $x > x^*$ (which is rejected and leads to agreement being reached on $y^* = p - c_D$ in the next period), and if $\delta_D = \delta_S$, then D is indifferent between proposing x^* and some $x > x^*$, and hence can be choosing either, or mixing.

Now consider a period in which S makes an offer. If D chooses to fight upon receiving a low offer, his payoff is $\frac{p-c_D}{1-\delta_D}$. If he chooses to move to the next period instead, he either (depending on the relative values of δ_S and δ_D) (i) finds it optimal to offer x^* (if $\delta_D \leq \delta_S$), which is accepted, or (ii) finds it optimal to propose some $x > x^*$ (if $\delta_D \geq \delta_S$), which is rejected. In case (ii), D is back in the same position that he is in the current period (in which S makes an offer), except that he has received q for two periods, and he can get at best $y^* = p - c_D$ in the period that he now finds himself in. Since $p - c_D > q$ and S's strategy does not change, if case (ii) holds, D is strictly better fighting rather than saying no if gets a low offer in the current period, and therefore he cannot credibly reject any offer $(y, 1-y)$ such that $y \geq p - c_D$. If case (i) holds, then D's optimal continuation value for moving to the next period is $q + \frac{\delta_D x^*}{1-\delta_D}$. It is easy to show that $\frac{p-c_D}{1-\delta_D}$ is strictly greater than this, and hence D cannot credibly reject any proposal $(y, 1-y)$ such that $y \geq p - c_D$, and must choose to fight if he gets a lower offer. Q.E.D.

7.1.1 Discussion

As discussed in the main body of the paper, D's proposal for himself, x^*, is strictly less than S's proposal for D, $y^* = p - c_D$, which is also D's payoff from war (as $\delta_S \to 1$, $x^* \to y^*$ from below). Thus, when $\delta_D \leq \delta_S$, D proposes an agreement that is worse for him than war, and this agreement is accepted (when $\delta_D = \delta_S$, he does not have to be doing this, but can be doing this, as in the SPE that Powell describes). Thus, the odd results that emerge are that (i) there is a first-mover disadvantage (which disappears in the limit as $\delta_S \to 1$), (ii) D proposes an agreement that he prefers less than war (again, this disappears in the limit as $\delta_S \to 1$), and (iii) S gets all of the gains from avoiding war (even more, when D makes the first offer, except in the limit as $\delta_S \to 1$). When δ_S and δ_D are allowed to differ, then in Figure 4, the horizontal axis should be labeled δ_S instead of δ, and x^* now represents S's minimal demand when D makes an offer, but not necessarily the offer that D actually *makes.*

When $\delta_D \leq \delta_S$, then D proposes the agreement x^*, which is accepted, whereas if $\delta_D \geq \delta_S$, then D proposes an unacceptable agreement so that agreement is reached on y^* in the next period. What is the intuition behind this delaying behavior? Recall that $q < x^* < y^*$. When δ_D is low (i.e., when $\delta_D \leq \delta_S$), D primarily cares about his payoff in the current period, and so he would rather get x^* now and forever after rather than get q in the current period (which is worse than x^*) and only get the better agreement of y^* beginning in the next period. On the hand, when δ_D is high, then D cares primarily about his long-run payoffs, and so he would rather live with the status quo in the current period and get the favorable payoff of y^* beginning in the next period rather than get the worse payoff of x^* now and forever after. But regardless of whether he chooses to offer x^* or delay reaching an agreement until the next period, D's average per-period payoff (which is the proper way of evaluating a player's welfare in a game like this) is strictly less than his payoff from war, which means that the results are dependent on the assumption that he cannot launch a war in a period in which he makes an offer.

7.2 Proof of Proposition 1

First consider D's decisions. Given S's acceptance rule, D is strictly best off proposing $x^* = p+c_S$ whenever he makes a proposal, as this is the best possible payoff he can effectively get in the model (given that S can choose to fight instead of accepting a worse offer, and if S chooses to fight, D's payoff is $p - c_D$, which is strictly worse than $p + c_S$). Now consider a period in which S makes an offer. If she makes a low offer and D chooses to fight, his payoff is $\frac{p-c_D}{1-\delta_D}$. If he chooses to say no instead, we have just shown that his optimal continuation value is $q + \frac{\delta_D x^*}{1-\delta_D}$ (note that if D says no, S chooses to pass rather than fight). For the upper bound on δ_D in this equilibrium, the former is greater than the latter, and hence D cannot credibly reject any offer $(y, 1-y)$ such that $y \geq p - c_D$, and must be choosing to fight rather than say no if he gets a worse offer. Now suppose D has made an offer and S has said no. If D chooses to fight, his payoff is $\frac{p-c_D}{1-\delta_D}$. If he chooses to pass instead, we have just shown that his payoff is $q + \frac{\delta_D y^*}{1-\delta_D}$. The former is strictly greater than the latter, and hence D must be choosing to fight.

Now consider S's decisions. Given D's acceptance rule, S is strictly best off proposing $y^* = p - c_D$ (if S makes a lower offer, D chooses to fight, in which case S is strictly worse off). Now consider a period in which D makes an offer. Since D is choosing to fight if S says no to his offer, S cannot credibly reject any offer than gives her at least her utility from war, i.e., she cannot credibly reject any offer $(x, 1-x)$ such that $x \leq p + c_S$. If she gets a worse offer, she is indifferent between fighting and saying no, since in the latter case, D chooses to fight anyway. Therefore, she can be doing either, or she can be mixing. Now suppose S has made an offer and D has said no. If S chooses to fight, her payoff is $\frac{1-p-c_S}{1-\delta_S}$. If she chooses to pass instead, we have just shown that her payoff is $(1-q) + \frac{\delta_S(1-x^*)}{1-\delta_S}$. The latter is strictly greater than the former, and hence S must be choosing to pass. Q.E.D.

7.3 Proof of Proposition 2

First consider D's decisions. The same argument as above shows that D is strictly best off proposing $x^* = p+c_S$ whenever he makes a proposal, given S's acceptance rule. Now consider a period in which S makes an offer. If she makes a low offer and D chooses to fight, his payoff is $\frac{p-c_D}{1-\delta_D}$. If he chooses to say no instead, we have just shown that his optimal continuation value is $q + \frac{\delta_D x^*}{1-\delta_D}$ (note that if D says no, S chooses to pass rather than fight). For the lower bound on δ_D in this equilibrium, the latter is greater than the former, and hence D cannot credibly reject any offer $(y, 1-y)$ such that $\frac{y}{1-\delta_D} \geq q + \frac{\delta_D x^*}{1-\delta_D}$, or $y \geq q(1-\delta_D) + \delta_D(p+c_S)$, and must choose to say no rather than fight if he gets a worse offer. Now suppose D has made an offer and S has said no. If D chooses to fight, his payoff is $\frac{p-c_D}{1-\delta_D}$. If he chooses to pass instead, his payoff is $q + \frac{\delta_D y^*}{1-\delta_D}$. (This is based on the assumption that S proposes y^* in the next period. As we show below, depending on the size of δ_S relative to δ_D, S may or may not find it optimal to offer y^* in the next period — however, whatever S chooses to do, D's average per-period payoff in the subgame beginning in the next period is y^* (this is because y^* is the offer that makes him just indifferent between accepting it and moving to the next period and reaching agreement on x^* therein), and hence this argument is fine.) For the upper bound on δ_D in this equilibrium, the former is greater than the latter, and hence D chooses to fight rather than pass.

Now consider S's decisions. S cannot credibly reject any offer that gives her at least her utility from war, i.e., any offer $(x, 1-x)$ such that $x \leq p + c_S$. This is because D chooses to fight if S says no to his offer. If S gets a worse offer, she is indifferent between fighting and saying no (since in the latter case D fights anyway), and hence can be doing either, or mixing. Now suppose D has said no to S's offer. We have just shown that S's continuation value for passing is $(1-q) + \frac{\delta_S(1-x^*)}{1-\delta_S}$. This is strictly greater than her payoff $\frac{1-p-c_S}{1-\delta_S}$ for fighting, and hence S must be passing rather than fighting. Now consider periods in which S makes an offer. Given D's acceptance rule, the best possible (for herself) acceptable agreement that S can propose in the current period is $y^* = q(1-\delta_D) + \delta_D(p+c_S)$, for a total payoff of $\frac{1-y^*}{1-\delta_S}$. Setting this greater than her payoff for proposing a lower y that is rejected and leads to agreement being reached in the next period, $(1-q)+\frac{\delta_S(1-x^*)}{1-\delta_S}$, and simplifying, we obtain $\delta_S > \delta_D$. Hence, if $\delta_S > \delta_D$, S is strictly best off proposing y^*. If $\delta_S < \delta_D$, S is strictly best off proposing some $y < y^*$ which is rejected and leads to agreement being reached on x^* in the next period. If $\delta_S = \delta_D$, then S is indifferent between proposing y^* and some $y < y^*$, and hence can be choosing either, or mixing. Q.E.D.

7.4 Proof of Proposition 3

Note that, in this proof, we use the "one-stage-deviation principle," henceforth OSDP, for infinite horizon games with discounting of future payoffs (Fudenberg and Tirole 1991, 108-110). This principle states that, to verify that a profile of strategies comprises a SPE, one just has to verify that, given the other players' strategies, no player can improve her payoff at any history at which it is her turn to move by deviating from her equilibrium strategy at that history and then reverting to her equilibrium strategy afterwards.

We want to look for a SPE in which D is mixing between passing and fighting, at any decision node at which S has said no to D's offer.[18] Suppose that in this (supposed) SPE, D's average per-period payoff for the subgame beginning in the next period (in which S makes an offer) is y'. Then, for mixing to be okay, it must be the case that D is indifferent between fighting and passing, i.e., $\frac{p-c_D}{1-\delta_D} = q + \frac{\delta_D y'}{1-\delta_D}$, or $y' = \frac{(p-c_D)-q(1-\delta_D)}{\delta_D}$. It is easy to verify that, given S's strategy and D's strategy for the rest of the game, D's average per-period payoff for the subgame beginning in the next period is indeed y', and hence D's strategy of mixing at this stage satisfies the OSDP. Therefore, suppose that D is choosing to fight with some probability $\beta \in (0,1)$ and pass with probability $1-\beta$.

Now consider when D has to make a proposal. Given S's acceptance rule, the best D

[18] A natural way to establish continuity with Propositions 1 and 2 would be if, when δ_D is high, as in Proposition 3, D chooses to pass with certainty even when S rejects his offer. In Proposition 2, δ_D is high enough that D chooses to say no (rather than fight) if S makes too low an offer, but still low enough that he prefers to fight if S rejects D's offer (his continuation value for moving to the next period is higher in the former case than in the latter, since in the former case he gets to make the proposal in the next period). So, it would be natural to expect that when δ_D is even higher, as in Proposition 3, D would choose to pass with certainty even when S rejects his offer. However, it turns out that this behavior cannot be supported as part of a stationary SPE, and instead D starts passing with positive probability, and this probability begins from zero and approaches one as the players' discount factors approach one. So, Proposition 3 *is* a natural continuation of Propositions 1 and 2, but uses a mixed strategy.

can do if he wants an agreement to be reached in the current period is to propose $x^* = \frac{(p-c_D)-q(1-\delta_D^2)}{\delta_D^2}$. If he proposes some bigger x, S says no and D's expected payoff (if he then uses his equilibrium strategy for the rest of the game) is $\beta(\frac{p-c_D}{1-\delta_D}) + (1-\beta)[q + \frac{\delta_D y'}{1-\delta_D}]$. Setting $\frac{x^*}{1-\delta_D}$ strictly greater than the latter and simplifying, we obtain $q < p - c_D$, which is true (note that β drops out of the simplification, so this is true for any value of β). Therefore, D cannot profitably deviate from proposing x^*, and then revert to his equilibrium strategy, and hence D's strategy satisfies the OSDP at histories at which D makes a proposal.

Now suppose S has just made an offer to D. If D fights, his payoff is $\frac{p-c_D}{1-\delta_D}$. If he says no instead, then, according to his equilibrium strategy for the rest of the game, his overall payoff will be $q + \frac{\delta_D x^*}{1-\delta_D}$ (note that S's strategy is to pass if D says no, and so the next period will be reached). Setting the latter strictly greater than the former and simplifying, we obtain $q < p - c_D$, which is true. Therefore, D is strictly better off saying no rather than fighting, if S's offer is too small. Therefore, he cannot do any better (assuming that he uses his equilibrium strategy in the future) than use the acceptance rule of accepting any offer $(y, 1-y)$ such that $\frac{y}{1-\delta_D} \geq q + \frac{\delta_D x^*}{1-\delta_D}$, or $y \geq \frac{(p-c_D)-q(1-\delta_D)}{\delta_D}$, and say no (rather than fight) if he gets a lower offer.[19] We have thus verified that D's strategy satisfies the OSDP, i.e., there exists no history at which D can profitably deviate from his equilibrium strategy at that stage and then revert back to his equilibrium strategy. Now we have to verify that the same is true for S.

Suppose D has just said no to S's offer. If S fights, her payoff is $\frac{1-p-c_S}{1-\delta_S}$. If she passes instead and follows her equilibrium strategy in the future, her payoff is $(1-q) + \frac{\delta_S(1-x^*)}{1-\delta_S}$. Setting the latter strictly greater than the former and simplifying, we obtain $\delta_D^2 > \frac{\delta_S[(p-c_D)-q]}{(p+c_S)-q}$, which is implied by our restriction in this proposition that $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$, and hence S's strategy satisfies the OSDP at this stage.

Now consider periods in which S makes a proposal. Given D's acceptance rule, the most favorable (for herself) acceptable agreement that S can propose is $y^* = \frac{(p-c_D)-q(1-\delta_D)}{\delta_D}$,

[19] Note that the only acceptance rule which is as good as this one for *any* offer by S, i.e., at *any* history at which S has just made an offer to D, is to accept any proposal $(y, 1-y)$ such that $y > \frac{(p-c_D)-q(1-\delta_D)}{\delta_D}$, and say no (rather than fight) if he gets a lower offer.

leaving her with an overall payoff of $\frac{1-y^*}{1-\delta_S}$. If she instead proposes some $y < y^*$, D rejects it and agreement is reached on x^* in the next period (assuming that S follows her equilibrium strategy), giving S an overall payoff of $(1-q) + \frac{\delta_S(1-x^*)}{1-\delta_S}$. Setting the former strictly greater than the latter and simplifying, we obtain $\delta_S > \delta_D$. Hence, (i) when $\delta_S > \delta_D$, S is strictly better off proposing y^* rather than doing something else and then reverting to her equilibrium strategy, (ii) when $\delta_S < \delta_D$, she is strictly better off proposing some $y < y^*$ rather than doing something else and then reverting to her equilibrium strategy, and (iii) when $\delta_S = \delta_D$, she is indifferent between proposing y^* and some $y < y^*$, and hence can be choosing either, or mixing. (And, this is strictly preferred to doing something else, namely proposing some $y > y^*$, and then reverting to her equilibrium strategy, although the latter point is moot because the agreement will be accepted and the game will end.) Therefore, we have verified that S's strategy satisfies the OSDP at histories at which she makes a proposal.

Finally, we need to verify that S's acceptance rule satisfies the OSDP. We consider the three cases in turn.

Case (i): $\delta_S > \delta_D$

Consider a period in which D makes an offer. According to S and D's equilibrium strategies, in the next period, agreement would be reached on y^* (since $\delta_S > \delta_D$). Therefore, in the current period, S's continuation value for saying no (if she uses her equilibrium strategy in the future) is $\beta[\frac{1-p-c_S}{1-\delta_S}]+(1-\beta)[(1-q)+\frac{\delta_S(1-y^*)}{1-\delta_S}]$. If she fights instead, her payoff is $\frac{1-p-c_S}{1-\delta_S}$. Setting the former strictly greater than the latter and simplifying, we obtain $\delta_D > \frac{\delta_S[(p-c_D)-q]}{(p+c_S)-q}$, which is implied by our restriction in this proposition that $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$. Therefore, S is strictly better off saying no rather than fighting, if she gets a low offer. Therefore, she cannot do any better (assuming she uses her equilibrium strategy in the future) than use the acceptance rule of accepting any offer $(x, 1-x)$ such that $\frac{1-x}{1-\delta_S} \geq \beta[\frac{1-p-c_S}{1-\delta_S}]+(1-\beta)[(1-q)+\frac{\delta_S(1-y^*)}{1-\delta_S}]$, and say no (rather than fight) if she gets a worse offer. Setting this equivalent to the acceptance rule described in the statement of the proposition (namely, $x \leq \frac{(p-c_D)-q(1-\delta_D^2)}{\delta_D^2}$) and solving for β, we obtain $\beta = \frac{(1-\delta_S\delta_D)[(p-c_D)-q]}{\delta_D\{\delta_D[(p+c_S)-q]-\delta_S[(p-c_D)-q]\}} \in (0,1)$. That is, when β takes on this value, S's acceptance rule as described in the proposition satisfies the OSDP at any history at which D has just made an offer to S. Note that $\beta \to 1$ (from below) as $\delta_D \to \frac{(p-c_D)-q}{\delta_D[(p+c_S)-q]}$ (from above). That is, this equilibrium converges to that of Proposition 2. Also note that our requirement in this proposition that $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$ means that $\beta > 0$ always. As $\delta_S, \delta_D \to 1$ (from below), $\beta \to 0$ (from above).

Case (ii): $\delta_S < \delta_D$

Consider a period in which D makes an offer. According to S and D's equilibrium strategies, in the next period, S will propose some $y < y^*$, which D rejects, and agreement will be reached on x^* in the following period. Therefore, in the current period, S's continuation value for saying no (if she uses her equilibrium strategy in the future) is $\beta[\frac{1-p-c_S}{1-\delta_S}]+(1-\beta)[(1-q)+\delta_S(1-q)+\frac{\delta_S^2(1-x^*)}{1-\delta_S}]$. If she fights instead, her payoff is $\frac{1-p-c_S}{1-\delta_S}$. Setting the former strictly greater than the latter and simplifying, we obtain $\delta_D^2 > \frac{\delta_S^2[(p-c_D)-q]}{(p+c_S)-q}$, which is implied by our restriction in this proposition that $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$. Therefore, S is strictly better off saying no rather than fighting, if she gets a low offer. Therefore, she cannot do any better (assuming she uses her equilibrium strategy in the future) than use the acceptance rule of accepting any offer $(x, 1-x)$ such that $\frac{1-x}{1-\delta_S} \geq \beta[\frac{1-p-c_S}{1-\delta_S}] + (1-\beta)[(1-q) + \delta_S(1-q) + \frac{\delta_S^2(1-x^*)}{1-\delta_S}]$, and say no (rather than fight) if she gets a worse offer. Setting this equivalent to the acceptance rule described in the statement of the proposition and solving for β, we obtain $\beta = \frac{(1-\delta_S^2)[(p-c_D)-q]}{\delta_D^2[(p+c_S)-q]-\delta_S^2[(p-c_D)-q]} \in (0,1)$. Note that $\beta \to 1$ (from below) as $\delta_D \to \frac{(p-c_D)-q}{\delta_D[(p+c_S)-q]}$ (from above). That is, this equilibrium converges to that of Proposition 2. Also note that $\beta > 0$ always, since $\delta_D > \delta_S$. As $\delta_S, \delta_D \to 1$ (from below), $\beta \to 0$ (from above).

Case (iii): $\delta_S = \delta_D = \delta$

Consider a period in which D makes an offer. According to S and D's equilibrium strategies, S's continuation value for saying no in the current period (regardless of whether she chooses to propose y^* or some $y < y^*$, or mix, in the next period) is $\beta[\frac{1-p-c_S}{1-\delta}]+(1-\beta)[(1-q)+\frac{\delta(1-y^*)}{1-\delta}]$. If she fights instead, her payoff is $\frac{1-p-c_S}{1-\delta}$. Setting the former strictly greater than the latter and simplifying, we obtain $c_S + c_D > 0$, which is true. Therefore, S is strictly better off saying no rather than fighting, if she gets a low offer. Therefore, she cannot do any better (assuming she uses her equilibrium strategy in the future) than use the acceptance rule of accepting any offer $(x, 1-x)$ such that $\frac{1-x}{1-\delta} \geq \beta[\frac{1-p-c_S}{1-\delta}] + (1-\beta)[(1-q) + \frac{\delta(1-y^*)}{1-\delta}]$, or $x \leq \beta(p+c_S) + (1-\beta)(p-c_D)$, and say no (rather than fight) if she gets a worse offer. Setting this equivalent to the acceptance rule described in the statement of the proposition and solving for β, we obtain $\beta = \frac{(1-\delta^2)[(p-c_D)-q]}{\delta^2(c_D+c_S)} \in (0,1)$. Note than $\beta \to 1$ (from below) as $\delta \to \frac{(p-c_D)-q}{\delta[(p+c_S)-q]}$ (from above), and hence this equilibrium converges to that of Proposition 2. Also note that $\beta \to 0$ (from above) as $\delta \to 1$ (from below).

Therefore, we have verified that D and S's strategies satisfy the OSDP at any history at which it is their turn to move. Q.E.D.

7.5 Proof of Proposition 4

The SPE characterized in Proposition 3 are stationary, except that when $\delta_S \leq \delta_D$, S can be choosing different actions (among which she is indifferent) at different histories (but that lead to structurally identical subgames) at which it is her turn to make an offer, and this allows for non-stationarity (but D and S's payoffs are the same in all of these SPE).

It turns out that when δ_D is high, there are also SPE that are non-stationary in a more genuine sense. Suppose that $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$, so that Proposition 3 holds. Consider the model in which D makes the first offer. Suppose that, in the subgame beginning in the second period, D and S use the strategies of Proposition 3, which we already know are best responses to each other. Then, in the last decision node of the first period, D is indifferent between passing and fighting, and so suppose D is choosing to fight with certainty (as opposed to fighting with probability β, as in Proposition 3). Then, in the first period, S's acceptance rule must be to accept any proposal $(x, 1-x)$ such that $x \leq p + c_S$, and so in the first period, D optimally proposes $x^* = p + c_S$. Hence, in Figure 6, there exist non-stationary SPE when δ_D is high in which x^* remains at $p + c_S$, rather than gradually decreasing to $p - c_D$, as in the stationary SPE.

In fact, we can suppose that in the last decision node of the first period, D fights with some probability λ and passes with probability $1-\lambda$. When $\lambda = 1$, we are in the SPE described above, and when $\lambda = \beta$, we are in the stationary SPE of Proposition 3. As λ

decreases, D's proposal for himself in the first period, x^*, decreases. When $\delta_S \geq \delta_D$ (so that agreement will be reached on $y^* = \frac{(p-c_D)-q(1-\delta_D)}{\delta_D}$ in the second period — see Proposition 3), then in the first period, S accepts all agreements $(x, 1-x)$ such that $\frac{1-x}{1-\delta_S} \geq \lambda[\frac{1-p-c_S}{1-\delta_S}] + (1-\lambda)[(1-q) + \frac{\delta_S(1-y^*)}{1-\delta_S}]$, and says no (rather than fight) for any worse offer.

This can be simplified to obtain that in the first period, S accepts all offers $(x, 1-x)$ such that $x \leq x^*$, where $x^* = \lambda[(p+c_S) - q] + \delta_S y^*(1-\lambda) + q(1-\delta_S) + \lambda q \delta_S$. When $\lambda = 1$, $x^* = p + c_S$, and when $\lambda = 0$, $x^* = \delta_S y^* + q(1-\delta_S)$. (And, since x^* is a continuous function of λ, any value of x^* in between these two extreme values can be obtained, for the right value of λ.) $\frac{\partial x^*}{\partial \lambda} > 0$ can be simplified to obtain $\delta_D > \frac{\delta_S[(p-c_D)-q]}{(p+c_S)-q}$, which is implied by our stipulation that $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$. Therefore, what S allows D to keep for himself in the first period is increasing in λ, which makes intuitive sense. (Also note that $\frac{\partial x^*}{\partial \delta_S} = 0$ when $\lambda = 1$ and $\frac{\partial x^*}{\partial \delta_S} > 0$ when $\lambda < 1$.)

Therefore, in the model in which D makes the first offer, there exist non-stationary SPE when $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$ (the binding condition of Proposition 3) and $\delta_S \geq \delta_D$ in which D's offer to himself (which S accepts) in the first period ranges from a minimum of $x^* = \delta_S y^* + q(1-\delta_S)$ to a maximum of $x^* = p + c_S$. Note that when $\lambda = 0$ and $\delta_S = \delta_D$, then $x^* = p - c_D$, i.e., D offers himself just his payoff from war. Of course, x^* can never be lower than $p - c_D$ in a SPE, unlike in Powell's model.

Now consider the model in which S makes the first offer. Suppose that, beginning in the third period and forever afterwards, both players use the strategies of Proposition 3. Then, in the last stage of the second period, D can be choosing to fight with any probability $\lambda \in [0, 1]$, and hence his payoff in the second period ranges from a minimum of $x^* = \delta_S y^* + q(1-\delta_S)$ to a maximum of $x^* = p + c_S$. Now consider the first period. If $x^* = p + c_S$ in the second period (i.e., $\lambda = 1$), then our restriction on δ_D means that, in the first period, D must be accepting any offer $(y, 1-y)$ such that $y \geq q(1-\delta_D) + \delta_D(p + c_S)$, and saying no (rather than fight) for any lower y. Since we have been assuming that $\delta_S \geq \delta_D$, S chooses to offer $y^* = q(1-\delta_D) + \delta_D(p + c_S)$ in the first period. Thus, there exist non-stationary SPE (in which $\lambda = 1$) when $\delta_S \geq \delta_D$ in which, in Figure 6, the trend continues as δ_D moves from medium to large, i.e., $x^* = p + c_S$ and $y^* = q(1 - \delta_D) + \delta_D(p + c_S)$ even when δ_D becomes very large. Note that, in these equilibria, as $\delta_D \to 1$, the agreement reached approaches $p + c_S$, i.e., D gets *all* of the gains from avoiding war, regardless of who gets to make the first proposal.

Now suppose that $\lambda = 0$ and $\delta_S = \delta_D$ (the worse possible case for D). Then, in the second period, D will propose $x^* = p - c_D$, and hence, in the first period, D's acceptance rule must be to accept any offer $(y, 1 - y)$ such that $y \geq p - c_D$, and fight (rather than say no) for any lower y. Thus, S proposes $y^* = p - c_D$, and hence S gets *all* of the gains from avoiding war.

Thus, when $\delta_S \geq \delta_D$ and S makes the first offer, the agreement reached in a non-stationary SPE can range from a minimum (for D) of $y^* = p - c_D$ (when $\lambda = 0$ and $\delta_S = \delta_D$) to a maximum of $y^* = q(1 - \delta_D) + \delta_D(p + c_S)$ (when $\lambda = 1$; as $\delta_D \to 1$, this converges to $p + c_S$).

We can generate additional non-stationary SPE. Suppose that in the model in which D makes the first proposal, the two players begin using the strategies of Proposition 3 beginning in the fourth period. Then, in the last stage of the third period, D can be choosing to fight with any probability $\lambda \in [0, 1]$. We can continue to build non-stationary SPE like this. The key is to stipulate that the players eventually start using the strategies of Proposition 3 in some period in which S makes an offer, and forever afterwards. In the last stage of the period just before then, D can be choosing to fight with any probability $\lambda \in [0, 1]$, and the choice of λ uniquely determines what has to occur previous to that stage (i.e., everything before then can be solved using backwards induction). Thus, we can create non-stationary SPE for any value of λ and any number of non-stationary initial periods. Thus, we have a folk-theorem type result when $\delta_D^2 \geq \frac{(p-c_D)-q}{(p+c_S)-q}$ (the binding condition of Proposition 3), in which a whole lot of payoff combinations can be supported as SPE (these payoffs lie between the upper and lower bounds identified earlier, since those bounds are determined by the most and least favorable agreements that D can possibly get in the first period in which he makes a proposal). The crucial thing is that, unlike in Powell's model, it cannot be the case that an agreement is reached in a SPE that gives one player less than his or her utility from war, because that player would be strictly better off instigating war in that period, which is always an option in our model, but not in Powell's.

7.6 Proof of Proposition 5

We want to construct a PBE in which neither type of D rejects S's initial offer in order to make a counteroffer. Each type accepts all initial offers $(y, 1-y)$ such that y is at least as great as its expected utility from war, and fights (rather than says no) if it gets a lower offer. We also want that if the second period is reached (this is off-the-equilibrium path behavior), the strategies of the players are such that agreement is reached on $x^* = p + c_S$, i.e., D gets all of the gains from avoiding war. First note that if such an agreement were to be reached, then S would be strictly best off passing rather than fighting if D says no to S's initial offer, i.e., $\frac{1-p-c_S}{1-\delta_S} < (1-q) + \frac{\delta_S(1-x^*)}{1-\delta_S}$. Then, for type c_{D_l} to be fighting rather than saying no if he gets a low initial offer, it must be that $\frac{p-c_{D_l}}{1-\delta_D} \geq q + \frac{\delta_D x^*}{1-\delta_D}$, or $\delta_D \leq \frac{(p-c_{D_l})-q}{(p+c_S)-q}$. This also ensures that type c_{D_l} cannot credibly reject any initial offer $(y, 1-y)$ such that $y \geq p - c_{D_l}$. Similarly, for type c_{D_h}'s acceptance rule to be to accept any initial offer $(y, 1-y)$ such that $y \geq p - c_{D_h}$ and go to war (rather than say no) for a lower y, it must be that $\frac{p-c_{D_h}}{1-\delta_D} \geq q + \frac{\delta_D x^*}{1-\delta_D}$, or $\delta_D \leq \frac{(p-c_{D_h})-q}{(p+c_S)-q}$. Since $c_{D_h} > c_{D_l}$, the latter is the binding restriction on δ_D.

Now we need to construct a PBE of the subgame beginning in the second period, which is in never reached in equilibrium, in which agreement is reached on $x^* = p + c_S$. The simplest way to do this is to stipulate that if this subgame is reached, S believes that it is facing type c_{D_l} (the low-cost, or highly resolved, type) with certainty (and that this belief never changes later on), and that S and type c_{D_l} therefore use the complete information strategies of Proposition 1, which are best responses to each other (note that we could also stipulate that S believes she is facing type c_{D_h} with certainty, and this belief never changes; the argument below would require only minor modifications). This requires the binding condition of Proposition 1 to hold (when D's cost of war is c_{D_l}), namely that $\delta_D \leq \frac{(p-c_{D_l})-q}{(p+c_S)-q}$, which is already implied by our binding condition in the previous paragraph.

Now we need to construct a strategy (in this subgame) for type c_{D_h} that is a best response to S's strategy (which is given by Proposition 1). Given S's acceptance rule, type c_{D_h} is strictly best off proposing $x^* = p + c_S$ whenever he makes a proposal. Now suppose S has just made a low offer to type c_{D_h}. We have just shown that if type c_{D_h} says no, his optimal continuation value is $q + \frac{\delta_D x^*}{1-\delta_D}$ (note that S's strategy is to pass rather than fight if D says no, and hence the next period will be reached). Given the upper bound on δ_D that we have derived earlier, namely that $\delta_D \leq \frac{(p-c_{D_h})-q}{(p+c_S)-q}$, type c_{D_h}'s payoff from war, $\frac{p-c_{D_h}}{1-\delta_D}$, is greater than this, and hence type c_{D_h}'s acceptance rule must be to always accept any offer $(y, 1-y)$ such that $y \geq p - c_{D_h}$, and fight if he gets a lower offer. Finally, suppose S has said no to type c_{D_h}'s offer. Given S's proposal and the acceptance rule we have just derived for type c_{D_h}, the latter's continuation value for passing is $q + \frac{\delta_D(p-c_{D_l})}{1-\delta_D}$. Given the upper bound on δ_D that we have derived earlier, namely that $\delta_D \leq \frac{(p-c_{D_h})-q}{(p+c_S)-q}$, type c_{D_h}'s payoff from war, $\frac{p-c_{D_h}}{1-\delta_D}$, is strictly greater than this, and hence type c_{D_h} must always be choosing to fight rather than pass. This completes the description of type c_{D_h}'s best response to S's strategy.

All that remains is to specify the optimal offer that S makes in the first period of the game. Given the acceptance rules of types c_{D_l} and c_{D_h}, S's best response is either to make the big offer $y^* = p - c_{D_l}$, which both types accept (and so war is avoided with certainty), or to make the lower offer $y^* = p - c_{D_h}$, which only type c_{D_h} accepts. Type c_{D_l} rejects it and goes to war. It is easy to see that no other proposal can be a best response. If $0 < s < 1$ is the prior probability that D is of type c_{D_l}, then making the big offer is a best response if and only if $\frac{1-(p-c_{D_l})}{1-\delta_S} \geq s[\frac{1-p-c_S}{1-\delta_S}] + (1-s)[\frac{1-(p-c_{D_h})}{1-\delta_S}]$, or $s \geq \frac{c_{D_h}-c_{D_l}}{c_{D_h}+c_S} \in (0,1)$. Q.E.D.

7.7 Proof of Proposition 6

This equilibrium is similar to the previous one in that, if the second period is reached, agreement is reached on $x^* = p+c_S$. However, because we have stipulated in this proposition that $\delta_D \geq \frac{(p-c_{D_h})-q}{(p+c_S)-q}$, in the first period, if type c_{D_h} gets a low initial offer, he prefers to move to the second period and get x^* rather than fight. Thus, his acceptance rule in the first period must be to accept any offer $(y, 1-y)$ such that $\frac{y}{1-\delta_D} \geq q + \frac{\delta_D x^*}{1-\delta_D}$, or $y \geq q(1-\delta_D)+\delta_D(p+c_S)$, and say no (rather than fight) for any lower y. Because we have stipulated in this proposition that $\delta_D \leq \frac{(p-c_{D_l})-q}{(p+c_S)-q}$, type c_{D_l} prefers to go to war if he gets a low initial offer rather than get x^* in the next period, and so his acceptance must be to accept all initial offers $(y, 1-y)$ such that $y \geq p - c_{D_l}$, and go to war for any lower y. Because agreement will be reached on x^* in the next period, if D says no to S's initial offer, S is strictly better off passing rather than fighting.

If the second period is reached, we stipulate that S believes with certainty that she is facing type c_{D_h}, and this belief never changes. (If the second period is reached on-the-equilibrium path, this belief follows from Bayes' rule, and if it is reached off-the-equilibrium path, we as the analyst stipulate that this is S's belief, since Bayes' rule does not apply. This off-the-equilibrium path belief is quite reasonable, because type c_{D_h}'s payoff from war is lower and hence he is more likely to say no rather than go to war than type c_{D_l}.) Since we have stipulated in this proposition that $\frac{(p-c_{D_h})-q}{(p+c_S)-q} \leq \delta_D \leq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$, the conditions for Proposition 2 are satisfied (when D is of type c_{D_h}), and hence we stipulate that, beginning in the second period, S and type c_{D_h} play the strategies of Proposition 2, which are best responses to each other. Because S is using the strategy of Proposition 2, agreement will indeed be reached on $x^* = p + c_S$, which we have been assuming. It is easy to construct type c_{D_l}'s best response to S's strategy, in the subgame beginning in the second period.

All that remains is to determine S's optimal offer in the first period. She can either make the large offer $y^* = p - c_{D_l}$, which both types of D accept, or make a smaller offer $y \leq q(1-\delta_D) + \delta_D(p + c_S)$, which type c_{D_l} rejects and goes to war. It is easy to see that no other offer can be a best response. We know from our proof of Proposition 2 that if $\delta_S \geq \delta_D$, then if S prefers to make the small offer, she prefers to offer exactly $y^* = q(1-\delta_D)+\delta_D(p+c_S)$, whereas if $\delta_S \leq \delta_D$, then if S prefers to make the small offer, she prefers to offer some $y < y^*$, so that agreement is reached on x^* in the next period (if D turns out to be type c_{D_h}). We consider the two cases in turn.

If $\delta_S \geq \delta_D$, then making the large offer $y^* = p - c_{D_l}$ is a best response if and only if $\frac{1-(p-c_{D_l})}{1-\delta_S} \geq s[\frac{1-p-c_S}{1-\delta_S}] + (1-s)[\frac{1-q(1-\delta_D)-\delta_D(p+c_S)}{1-\delta_S}]$, or $s \geq \frac{(p-c_{D_l})-[q(1-\delta_D)+\delta_D(p+c_S)]}{(p+c_S)-[q(1-\delta_D)+\delta_D(p+c_S)]} \in [0, 1)$.

If $\delta_S \leq \delta_D$, then making the large offer $y^* = p - c_{D_l}$ is a best response if and only if

$\frac{1-(p-c_{D_l})}{1-\delta_S} \geq s[\frac{1-p-c_S}{1-\delta_S}] + (1-s)[(1-q) + \frac{\delta_S(1-x^*)}{1-\delta_S}]$, or $s \geq \frac{(p-c_{D_l})-[q(1-\delta_S)+\delta_S(p+c_S)]}{(p+c_S)-[q(1-\delta_S)+\delta_S(p+c_S)]} \in [0,1)$. Q.E.D.

7.8 Proof of Proposition 7

We want to construct a risk-return tradeoff equilibrium even when δ_D is high. That is, type c_{D_l} goes to war rather than saying no if he gets a low initial offer. We stipulate that if the second period is reached, S believes that it is facing type c_{D_h} with certainty (on-the-equilibrium path, this will follow from Bayes' rule, and off-the-equilibrium path, this is a reasonable stipulation, for the same reason as in the previous proof), and that this belief never changes. Therefore, since we want to allow δ_D to be high, we stipulate that in the subgame beginning in the second period, S and type c_{D_h} use the strategies of Proposition 3, which are best responses to each other if $\delta_D \geq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$, which we therefore stipulate to hold in this proposition. Therefore, we know from the proof of Proposition 3 that type c_{D_h}'s acceptance rule in the first period is optimal, given that he expects agreement to be reached on $x^* = \frac{(p-c_{D_h})-q(1-\delta_D^2)}{\delta_D^2}$ if the second period is reached.

Now consider type c_{D_l}'s behavior in the first period. He knows that if the second period is reached, S adopts the strategy of Proposition 3, treating D as if he is of type c_{D_h} (since S believes this with certainty). What is type c_{D_l}'s best response to this (in the subgame beginning in the second period)? It is either to propose $x^* = \frac{(p-c_{D_h})-q(1-\delta_D^2)}{\delta_D^2}$, or to propose some $x > x^*$ which S rejects, and then go to war. (Since type c_{D_h} is indifferent between passing and fighting if S rejects D's offer in the second period, type c_{D_l}, whose payoff from war is strictly higher, strictly prefers to fight rather than pass.) He prefers to offer x^* if $\frac{x^*}{1-\delta_D} \geq \frac{p-c_{D_l}}{1-\delta_D}$, or $\delta_D \leq \frac{(p-c_{D_h})-q}{\delta_D[(p-c_{D_l})-q]}$. On the other hand, if $\delta_D \geq \frac{(p-c_{D_h})-q}{\delta_D[(p-c_{D_l})-q]}$, then c_{D_l} prefers to instigate war in the second period rather than offer x^*.

Note that $\frac{(p-c_{D_h})-q}{\delta_D[(p-c_{D_l})-q]} > \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$. Therefore, suppose that $\delta_D \geq \frac{(p-c_{D_h})-q}{\delta_D[(p-c_{D_l})-q]}$, so that c_{D_l} is best off instigating war in the subgame beginning in the second period. He is strictly better off going to war in the first period than in the second period, and therefore his acceptance rule in the first period is fine. Now suppose that $\frac{(p-c_{D_h})-q}{\delta_D[(p-c_{D_l})-q]} \geq \delta_D \geq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$, so that c_{D_l} is best off proposing x^* in the subgame beginning in the second period. Then, c_{D_l}'s acceptance rule in the first period is fine as long as $\frac{p-c_{D_l}}{1-\delta_D} \geq q + \frac{\delta_D x^*}{1-\delta_D}$, or $\delta_D \geq \frac{(p-c_{D_h})-q}{(p-c_{D_l})-q}$. Therefore, as long as $\delta_D \geq max\{\frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}, \frac{(p-c_{D_h})-q}{(p-c_{D_l})-q}\}$, type c_{D_l}'s acceptance rule in the first period is fine, and therefore we stipulate this to hold in this proposition.

Note that, if D says no to S's initial offer, the proof of Proposition 3 shows that S is strictly better off passing rather than fighting, since she expects agreement to be reached on x^* in the next period.

All that remains is to determine S's optimal offer in the first period. She can either make the large offer $y^* = p - c_{D_l}$, which both types of D accept, or make a smaller offer $y \leq \frac{(p-c_{D_h})-q(1-\delta_D)}{\delta_D}$, which type c_{D_l} rejects and goes to war. It is easy to see that no other offer can be a best response. We know from the proof of Proposition 3 that if $\delta_S \geq \delta_D$, then if S prefers to make the small offer, she prefers to offer exactly $y^* = \frac{(p-c_{D_h})-q(1-\delta_D)}{\delta_D}$, whereas if $\delta_S \leq \delta_D$, then if S prefers to make the small offer, she prefers to offer some $y < y^*$, so that agreement is reached on x^* in the next period (if D turns out to be type c_{D_h}). We consider the two cases in turn.

If $\delta_S \geq \delta_D$, then making the large offer $y^* = p - c_{D_l}$ is a best response if and only if $\frac{1-(p-c_{D_l})}{1-\delta_S} \geq s[\frac{1-p-c_S}{1-\delta_S}] + (1-s)[\frac{1-\frac{(p-c_{D_h})-q(1-\delta_D)}{\delta_D}}{1-\delta_S}]$, or $s \geq \frac{[q(1-\delta_D)+\delta_D(p-c_{D_l})]-(p-c_{D_h})}{[q(1-\delta_D)+\delta_D(p+c_S)]-(p-c_{D_h})} \in [0,1)$.

If $\delta_S \leq \delta_D$, then making the large offer $y^* = p - c_{D_l}$ is a best response if and only if $\frac{1-(p-c_{D_l})}{1-\delta_S} \geq s[\frac{1-p-c_S}{1-\delta_S}] + (1-s)[(1-q) + \frac{\delta_S(1-x^*)}{1-\delta_S}]$, or $s \geq \frac{\delta_D^2[(p-c_{D_l})-q]-\delta_S[(p-c_{D_h})-q]}{\delta_D^2[(p+c_S)-q]-\delta_S[(p-c_{D_h})-q]} \in [0,1)$. Q.E.D.

7.9 Proof of Proposition 8

We want to construct a PBE in which agreement is reached on $x^* = p + c_S$ if the second period is reached, and in which both types of D make counteroffers rather than go to war if S's initial offer is too small. The natural way to do this is to use the results of Proposition 2, in which agreement is reached on $x^* = p + c_S$ whenever D makes a proposal, and D says no rather than fights, if S makes a small offer. The stipulation in this proposition that $\frac{(p-c_{D_l})-q}{(p+c_S)-q} \leq \delta_D \leq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$ means that the conditions of Proposition 2 hold for both types of D, i.e., it means that $\frac{(p-c_{D_l})-q}{(p+c_S)-q} \leq \delta_D \leq \frac{(p-c_{D_l})-q}{\delta_D[(p+c_S)-q]}$ and $\frac{(p-c_{D_h})-q}{(p+c_S)-q} \leq \delta_D \leq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$.

Thus, we simply specify that, beginning from the very first period and continuing forever after, both types of D use the strategy of Proposition 2. Note that this strategy does not depend on D's cost of war in any way, and hence both types are adopting identical strategies. Since they are adopting identical strategies, S's best response is to adopt the strategy of Proposition 2, regardless of the value of s. And, given that S is adopting the strategy of Proposition 2, the best response of both types of D is to use the strategy of Proposition 2.

If the second period is reached on-the-equilibrium path, S's belief will remain at s, by Bayes' rule. If it is reached off-the-equilibrium path, then we specify that S's belief can be anything, and that S as well as both types of D continue to use the strategies of Proposition 2. If the third period is reached (this can only happen off-the-equilibrium path, since both types of D fight if S rejects D's offer in the second period), then the belief can be anything, and everyone continues to use the strategies of Proposition 2, and so on. Q.E.D.

7.10 Proof of Proposition 9

We want to construct a PBE in which both type of D make counteroffers rather than fight, if S's initial offer is too small, and in which δ_D can be very high. Because we want to allow δ_D to be high, we use the results of Proposition 3. In particular, we stipulate that if the third period is reached (it will turn out that this can only happen off-the-equilibrium path), then S believes with certainty that she is facing type c_{D_h} (which, as we have been discussing earlier, is a sensible belief), and this belief never changes. Thus, we stipulate that, beginning in the third period, S and c_{D_h} use the strategies of Proposition 3 (with S treating D as though it is type c_{D_h} with certainty), which are best responses to each other if $\delta_D \geq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$, which we therefore stipulate to hold in this proposition. Type c_{D_l}'s best response to S's strategy is easy to construct.

Now consider the last decision node of the second period, in which D has to decide whether to pass or fight. We know from Proposition 3 that type c_{D_h} is indifferent between fighting and passing, since he expects his payoff from Proposition 3 to be obtained in the next period. Since he is indifferent, we stipulate that he chooses to fight with certainty (as opposed to fighting with probability β, as in Proposition 3). Since c_{D_h} is indifferent, type c_{D_l}, whose payoff from war is strictly higher, strictly prefers to fight. Therefore, since both types of D are choosing to fight if S rejects D's offer in the second period, S's acceptance rule in the second period must be to accept any offer $(x, 1-x)$ such that $x \leq p + c_S$, regardless of what her belief about D's type is at that point. Therefore, both types of D are strictly best off proposing $x^* = p + c_S$ in the second period.

Now consider the first period. Our previously derived stipulation that $\delta_D \geq \frac{(p-c_{D_h})-q}{\delta_D[(p+c_S)-q]}$ ensures that type c_{D_h} strictly prefers to say no (and get x^* in the next period) rather than fight if S's initial offer is too small, and therefore c_{D_h}'s acceptance rule in the first period is fine. Type c_{D_l} prefers to say no and obtain x^* in the next period rather than fight if S's initial offer is too low as long as $\delta_D \geq \frac{(p-c_{D_l})-q}{(p+c_S)-q}$, which we therefore stipulate to hold in this proposition, and hence c_{D_l}'s acceptance rule in the first period is fine. The last thing to note is that if D says no to S's initial offer, S is strictly better of passing rather than fighting, since she expects agreement to be reached on x^* in the next period. Q.E.D.

References

[1] Barraclough, Geoffrey. 1982. *From Agadir to Armageddon: Anatomy of a Crisis*. London: Weidenfeld and Nicolson.

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

[3] Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason*. New Haven: Yale University Press.

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

[5] Filson, Darren, and Suzanne Werner. 2002. "A Bargaining Model of War and Peace." *American Journal of Political Science* 46(October):819-38.

[6] Fudenberg, Drew, and Jean Tirole. 1991. *Game Theory*. Cambridge: MIT Press.

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

[8] Kydd, Andrew. 2003. "Which Side Are You On? Bias, Credibility, and Mediation." *American Journal of Political Science* 47:597-611.

[9] Lowe, John. 1994. *The Great Powers, Imperialism and the German Problem, 1865-1925*. London and New York: Routledge.

[10] Morrow, James D. 1989. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining." *American Journal of Political Science* 33(November):941-972.

[11] Muthoo, Abhinay. 1999. *Bargaining Theory with Applications*. Cambridge: Cambridge University Press.

[12] Osborne, Martin, and Ariel Rubinstein. 1990. *Bargaining and Markets*. New York: Academic Press.

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

[14] Powell, Robert. 1996b. "Stability and the Distribution of Power." *World Politics* 48(January):239-67.

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

[16] Powell, Robert. 2002. "Bargaining Theory and International Conflict." *Annual Review of Political Science* 5(June):1-30.

[17] Powell, Robert. 2004a. "Bargaining and Learning While Fighting." *American Journal of Political Science* 48(2):344-361.

[18] Powell, Robert. 2004b. "The Inefficient Use of Power." *American Political Science Review* 98(May):231-41.

[19] Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 60:169-203.

[20] Romer, Thomas, and Howard Rosenthal. 1978. "Political Resource Allocation, Controlled Agendas, and the Status-quo." *Public Choice* 33:27-43.

[21] Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50(January):79-109.

[22] Slantchev, Branislav. 2003. "The Principle of Convergence in Wartime Negotiations." *American Political Science Review* 97(4):621-632.

[23] Smith, Alastair, and Allan C. Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48(6):783-813.

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

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

[26] Walter, Barbara F. 2002. *Committing to Peace: The Successful Settlement of Civil Wars*. Princeton: Princeton University Press.

[27] Waltz, Kenneth. 1979. *Theory of International Politics*. New York: McGraw-Hill.

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

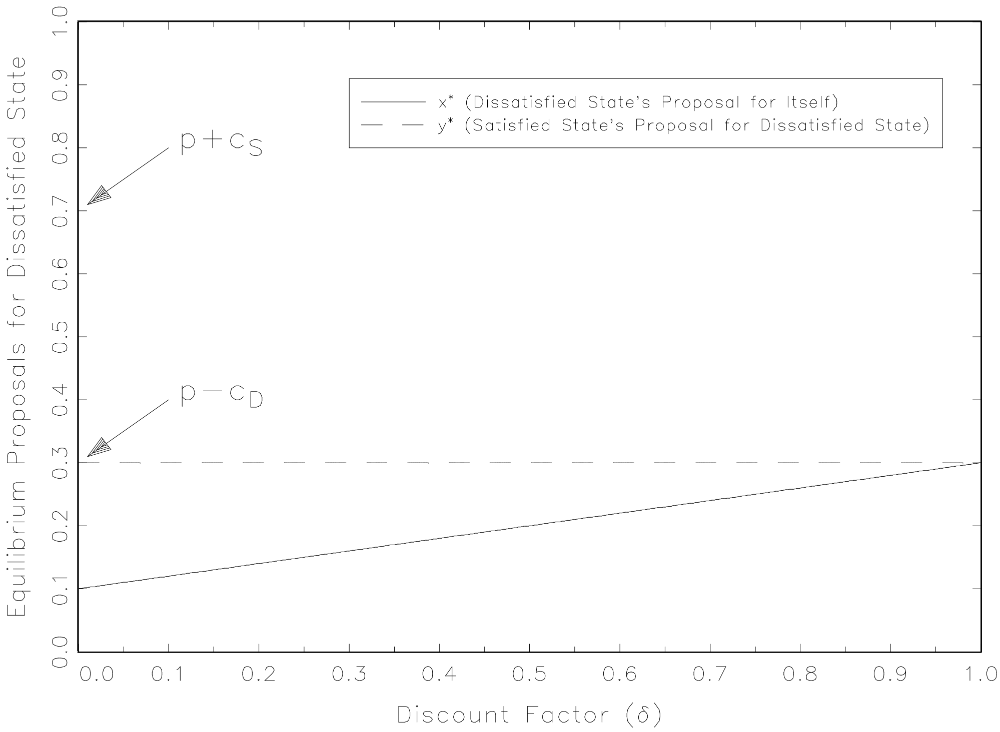

Figure 4: Equilibrium Proposals in Powell's Model

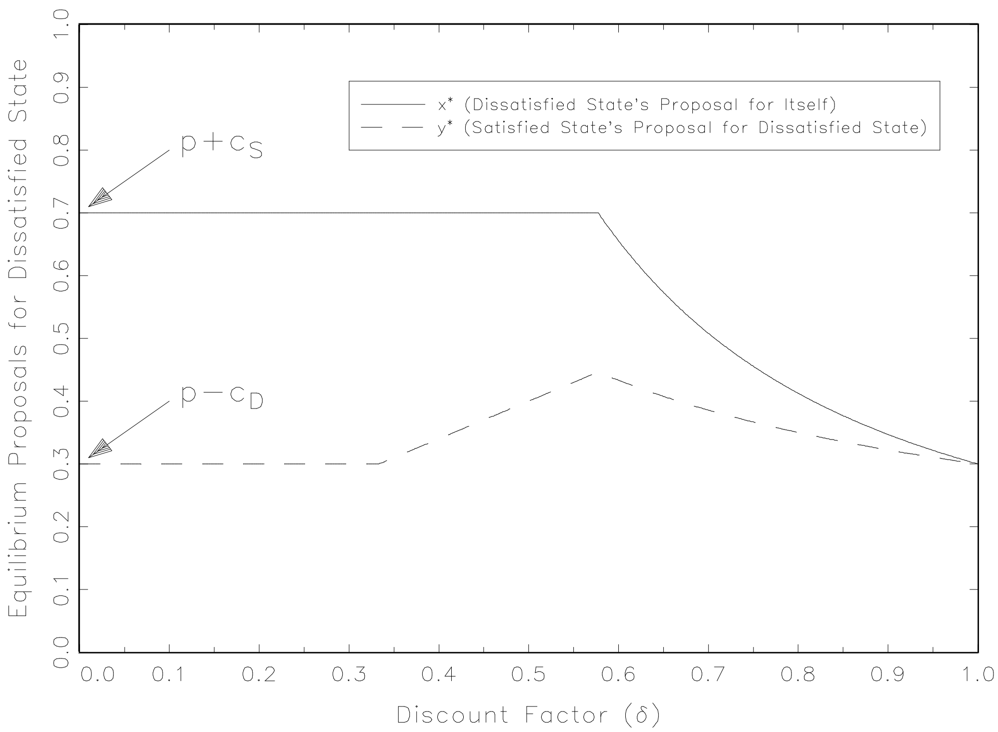

Figure 6: Equilibrium Proposals in the Generalized Model

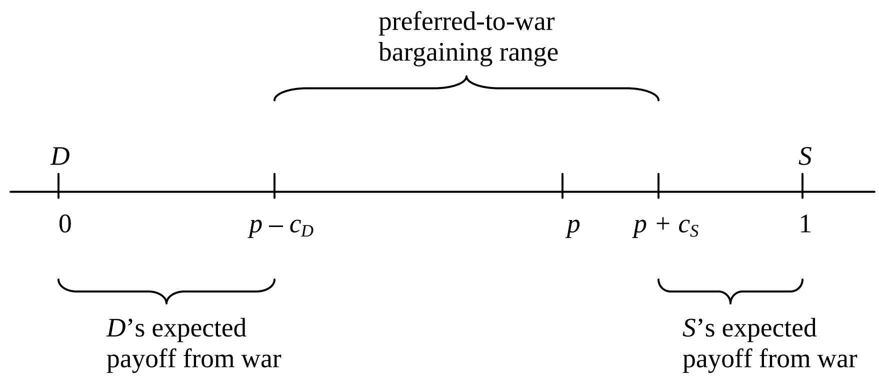

Figure 1: Fearon's (1995) Bargaining Model

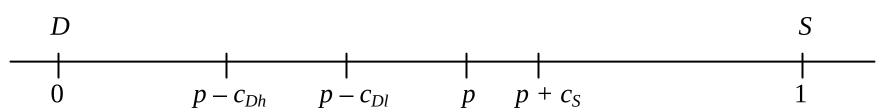

Figure 2: Uncertainty About *D*'s Cost of War

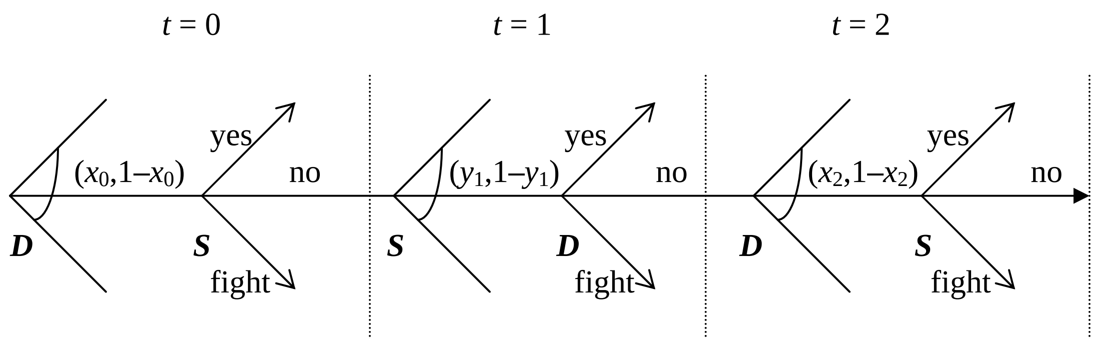

Figure 3: Powell's Model

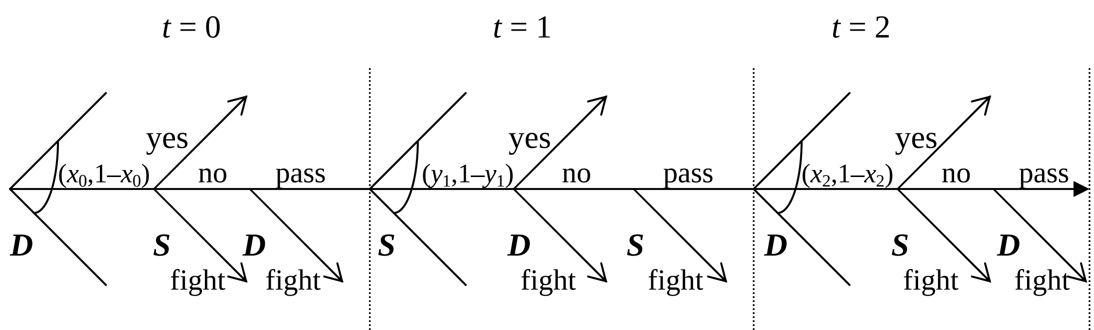

Figure 5: A Generalization