The Power to Hurt: Costly Conflict with Completely Informed States

Branislav L. Slantchev*
Department of Political Science
University of Rochester
Rochester, NY 14627

bsla@troi.cc.rochester.edu
Phone: 716-737-8260

February 16, 2002

Abstract

According to common explanations of war, since war is costly and risky, rational states have incentives to strike a bargain and avoid conflict. War is then the result of private information about capabilities or resolve and incentives to misrepresent this information, or inability to make credible commitments (Fearon 1995).

This article presents another rationalist explanation, which does not depend on either of these mechanisms. By modeling warfare as a costly bargaining process, the paper demonstrates that inefficient fighting can occur in equilibrium under very general assumptions favoring peace. Specifically, it is assumed that peace can be supported in equilibrium, and that fighting brings no benefits to either state, only costs. Despite the fact that there exist many agreements that Pareto-dominate the final settlement, states are unable to agree to them and fighting can occur in equilibrium.

The equilibrium is driven by the ability of states to impose costs on their opponent, and bear costs. The power to hurt the opponent establishes a range of acceptable settlements, and the existence of such a range makes inefficient equilibria possible. The article suggests that (1) states may fight if they believe that seeking peace means accepting an unpalatable settlement; and (2) a diminished ability to hurt the enemy is a major reason for war termination. Historical examples are provided to illustrate these results.

*I would like to thank Randall Stone, Curtis Signorino, John Duggan, Robert Powell, Joanne Gowa, David Lake, and two anonymous reviewers for thoughtful comments. I am especially indebted to Robert Powell for pointing out that the original version of Proposition 3.1 assumed implicitly that Γ was a simultaneous-move game. I also gratefully acknowledge financial support from the Watson Center for Conflict and Cooperation, University of Rochester. A previous version was presented at the Peace Science Society (International) Conference, October 2001, Atlanta, GA. Needless to say, all errors remain strictly the responsibility of the author.

I. Introduction

One of the central puzzles in international relations theory is why wars occur. Since fighting is always inefficient *ex post*, it should be possible to design an agreement that both states would accept in lieu of going to war and negotiating thereafter. Fearon (1995) conceptualized war as bargaining failure and identified rationalist explanations of this puzzle: Possession of private information with incentives to misrepresent it, and inability to commit not to use one's advantage to extract further concessions in the future.[1]

Two crucial assumptions underlie these results: Crisis negotiation is modeled as a single-shot bargaining event, and war is modeled a costly lottery over exogenously fixed outcomes. These assumptions, however, are quite problematic as approximations of war as usually conceptualized by practitioners and historians. One of the classic statements in this regard belongs to Clausewitz, who argued that "war [...] always lasts long enough for influence to be exerted on the goal and for its own course to be changed in one way or another" (Clausewitz 1832, p. 87).[2] Modeling conflict using the two common assumptions misses the point that war is a process, in which players can condition their strategies depending on past play. Recent formal theoretical advances have begun to address these issues, bringing them closer to the non-formal treatments.[3]

Posing the question in terms of a bargaining process that occurs in the shadow of fighting is useful in addressing the related puzzle of war termination: If fighting is costly, why do the sides delay in reaching a settlement to end it?[4] Most wars do not end in

---

[1]There is also a third mechanism that relies on issue indivisibilities but Fearon finds it less compelling. For an explanation why such indivisibilities may arise and become important in civil wars, see Walter (2001). Fearon's argument has largely channeled recent inquiry into the causes of war and related inefficient phenomena. Although there are several works that explain conflict as a dynamic commitment problem (e.g. Powell 1997), it is the asymmetric information argument that seems to have received most attention.

[2]Fuller's (1961, pp. 63-5) highest praise for Clausewitz is reserved for latter's recognition that war is an instrument of policy and that interstate bargaining does not cease when hostilities begin. Schelling (1966, p. 7), of course, was blunt: "War [is] a bargaining process—dirty, extortionate, and often quite reluctant as a bargaining process on one side or both—nevertheless a bargaining process." More recently, Wagner (2000) argues that war should be conceptualized as a bargaining process, and provides an early formal treatment.

[3]See, for example, Filson and Werner (2001), Kim (2001), Powell (2001), Smith and Stam (2001), and Slantchev (2002).

[4]Considering the importance of this question, it may be surprising to find that relatively few works are concerned with it. General works include Kecskemeti (1958), Iklé (1971), Dunnigan and Martel (1987),

complete military victory that eliminates the opponent's ability to continue fighting but in some negotiated settlement (Pillar 1983). Focusing on victory and defeat, as the costly lottery approach does, can be therefore misleading in this respect as well.

In this article, I model war as a process instead of outcome, and allow outcomes to be endogenous. One of the goals is to examine the strategic incentives for inefficient behavior in the absence of asymmetric information, especially as it relates to ending costly conflict. To this end, I stack the model against war by making several assumptions: (i) peace can be supported in equilibrium; (ii) peace is the most preferred outcome; (iii) fighting produces costs and no direct benefits; and (iv) there is complete information. Therefore, in this model war is not profitable by assumption. I find that even under these strict assumptions there exist inefficient equilibria, in which fighting occurs.

Thus, this article provides another plausible explanation of why states may fight despite the inefficiency. This explanation, which turns on the ability of states to impose costs on their opponents, and the ability to bear them, does not require asymmetric information, does not derive from credibility problems, and does not depend on issue indivisibilities. Also, several useful additional insights emerge from the analysis.

Once war is disaggregated from a lottery over exogenous outcomes into a process where war aims arise endogenously, it is possible to make a subtle distinction between two types of costs that are associated with conflict: The ability to bear costs, and the ability to impose costs. To my knowledge, there exists no theory in international relations, either formal or non-formal, that alludes to such difference, or that is even capable of conceptualizing it within its framework.[5] This distinction is relevant because it induces bargaining strategies that produce outcomes, which cannot be predicted from the usual models.

These results draw attention to two underappreciated areas in research on war: The endogeneity of war aims, and the ability of states to inflict pain on their adversaries.

Smith (1995), and Kegely and Raymond (1999). The paucity of empirical research is also staggering, but there are two recent interesting works that take very different approaches: See Goemans (2000) for a detailed case study of World War I, and Stam (1999) for a quantitative analysis.

[5]It is worth pointing out that these abilities are related, but not simply isomorphic. That is, it is not the case that state A's ability to impose costs on state B is the same as state B's ability to bear costs. This is because state A must in turn incur costs when it imposes costs on B, which determines the extent to which A can hurt B. These costs are not entirely dependent on B's defenses because they depend on other factors, such as A's domestic politics, or international situation. During the Korean War, the U.S. had the option of using nuclear weapons against North Korea and China. The costs this would impose on the opponent would have been substantial. On the other hand, concern with Soviet reaction, and perhaps domestic public opinion hampered U.S. ability to impose these costs.

Although coercion has been recognized as perhaps the premier instrument of wartime politics, theories of war have generally ignored the possibility that states might condition negotiation strategies on their performance in the war, and that such strategies might depend on their ability to impose costs in unforeseen ways.

It is interesting that inefficiency can occur under complete information. There is little doubt that uncertainty is a pervasive feature of international politics, and almost all models, in which war occurs in equilibrium, rely on asymmetric information to explain it (Powell 2001). As Goemans (2000, p. 24) notes, "if both sides knew how the pie would be divided after the war, both would be better off if they divided accordingly before the war."[6] However, simply knowing how the pie would be divided does not necessarily mean that the division is the only one that can be supported in equilibrium *ex ante* even if such outcome is preferable because it is Pareto-improving.

As this article shows, there exists a range of negotiated agreements that can be supported as outcomes of efficient equilibria, and some of these leave players distinctly worse off. The threat to switch to one of these different equilibria depending on the path of play can support a variety of inefficient equilibria that involve fighting.[7] In these inefficient equilibria, almost any agreement prior to the one that eventually obtains would leave both players better off, and yet the strategies are such that neither player wants to deviate and stop fighting.

I should emphasize that the result of inefficiency under complete information has much in common with several formal models of wage bargaining in economics. This literature has also normally invoked asymmetric information as the cause of delay and suboptimal outcomes, but there have been several recent development. Fernandez and Glazer (1991) study a complete information model, in which a firm and a union bargain over wages, and in which the union can go on strike. They are able to demonstrate that there exist inefficient equilibria, in which the union strikes for some time.[8] Haller and Holden (1990) analyze essentially the same model with similar results. Unlike these models, in which only one player can impose costs on the other, I allow both players to

[6]This is almost identical to Fearon's (1995, pp. 380-8) conclusion.

[7]This is the standard "equilibrium switching" method commonly used to establish Folk Theorems for repeated games. See, for example Fudenberg and Maskin (1986). It is well known that it is relatively easy to obtain inefficient outcomes in repeated games. The fear of retaliation (which comes from players being able to condition their actions on past history of the game) can produce outcomes, including inefficient ones, that would otherwise not occur. The model of war as a bargaining process, however, is not a repeated game because agreement ends it, and the standard Folk Theorems therefore do not apply.

[8]Their paper has an error in the proof, which is corrected by Bolt (1995).

engage in a conflict game after any offer is rejected.

This approach is similar to Busch and Wen (1995), who study a more general model, where a disagreement game is embedded in a Rubinstein (1982) bargaining game. They also demonstrate that inefficient behavior can occur when the disagreement game, which is played after an offer is rejected, meets certain conditions. Muthoo (1999), on which my analysis heavily relies, provides a slightly different version of some proofs. The present model extends these models by allowing players to have different discount rates, which turns out to be a non-trivial assumption. In addition to this extension, which is of interest for theoretical reasons, the discussion of the substantive implications points to areas that are not developed (or even recognized) in much of the literature in international relations.

The rest of this paper is organized as follows. Section 2 presents the model. Section 3 analyzes the set of efficient stationary and nonstationary equilibria, and finds the minimum payoffs that can be supported in equilibrium. Section 4 establishes that main result, namely that inefficient equilibria, in which fighting occurs before an agreement is reached, exist. Subsection 4.1 extends the model to allow for repeated renegotiation of agreements. Section 5 provides several illustrative historical cases that highlight different features of the model and the conclusions. Finally, Section 6 concludes.

II. The Model

Consider two states, $i \in \{1, 2\}$, that are bargaining over a two-way partition of a flow of benefits with size π. An agreement is a pair (x_1, x_2), where x_i is state i's share. The set of possible pairs is

$$\mathcal{X} = \left\{(x_1, x_2) \in \mathcal{R}^2 : x_1 + x_2 = \pi \text{ and } 0 \leq x_i \leq \pi \text{ for } i = 1, 2\right\}.$$

States have strictly opposed preferences and each is concerned only with the share of benefits it obtains from the agreement. There is a status quo distribution of benefits, (s_1, s_2) with $s_1 + s_2 = \pi$ that exists prior to any agreement. Because a share x_i in a proposal identifies a distribution uniquely, I shall write x_1, which is equivalent to the pair $(x_1, \pi - x_1)$, and x_2, which is equivalent to the pair $(\pi - x_2, x_2)$.

The two players bargain according to the alternating-offers protocol. Players act in discrete time with potentially infinite horizon, with periods indexed by t $(t = 0, 1, 2, \ldots)$. Players have time preferences with constant interest rates, and so the discount factors are $\delta_i \in (0, 1)$. In even-numbered periods, player 1 proposes a division $x_1 \in \mathcal{X}$ to player

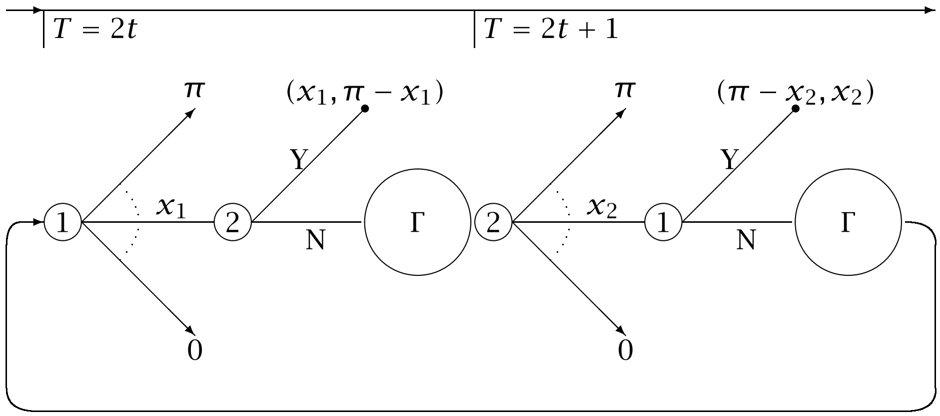

Figure 1: Diagram of two periods of the negotiation game.

2. If player 2 accepts that proposal, an agreement is reached, and players obtain their share in $(x_1, \pi - x_1)$, respectively, in every following period.

If player 2 rejects the proposal, then both states play the conflict game Γ. The description of this game is abstract. Let Γ be some game in normal form where each player i has an action space A_i. Given a profile of actions (a_1, a_2) with $a_1 \in A_1$ and $a_2 \in A_2$, the player's payoff is $r_i(a_1, a_2)$.

In order to stack the model against fighting, assume (i) peace is one of the outcomes in Γ; (ii) peace is the most preferred outcome for both players—that is, fighting is costly and brings no benefits to either player in Γ; and (iii) peace is a subgame perfect equilibrium (SPE) of Γ. Let the peace payoff to player i be its share of the status quo distribution of benefits. If either player attacks, any outcome results in some payoff strictly less than the peace payoff. The payoffs for the worst outcome are normalized to 0. Thus, for all $a_1 \in A_1, a_2 \in A_2$, $r_i(a_1, a_2) \in [0, s_i]$, where the payoff s_i is only obtainable if neither player fights in Γ. The payoffs in the conflict game constitute per-period payoffs in the bargaining game.

After Γ finishes, time advances to the next period where player 2 makes a counteroffer $x_2 \in \mathcal{X}$. Player 1 can either accept it, in which the players obtain $(\pi - x_2, x_2)$ respectively, or reject it, in which case they play Γ again. The game continues until an agreement is reached. Figure 1 shows the schematic for two periods of the negotiation game. Each state's objective is to maximize its average intertemporal payoff $(1-\delta_i) \sum_{t=0}^{\infty} \delta_i^t z_{it}$, where z_{it} is state i's per-period payoff at time t and equals its share of the benefits if an agreement is reached, or r_{it} otherwise.

Since Nash equilibrium (NE) may rely on non-credible threats, the solution concept I shall use is that of subgame perfect equilibrium (SPE), which is a refinement of NE in that it requires each player's strategy to be optimal in every proper subgame, whether or not this subgame is ever reached when players follow their strategies (Selten 1975). In other words, an equilibrium is subgame perfect if the strategies it induces are a Nash equilibrium in every subgame.

III. Efficient Equilibria: The Threat to Hurt

Two additional properties of some SPE can be sometimes useful and so I define them here. An SPE in the bargaining game is *stationary* if players always make the same proposals regardless of history and time. In addition, an SPE is *no-delay* if a player's equilibrium proposal is immediately accepted by the other player. As customary in bargaining games, assume that if a player is indifferent between accepting a proposal and continuing the game, it accepts the proposal. The following proposition, whose proof is the appendix, establishes the existence of stationary efficient equilibria.[9]

Proposition 3.1. *For each subgame perfect equilibrium (SPE) of* Γ*, the bargaining game has a unique stationary no-delay subgame perfect equilibrium, in which for* $i, j \in \{1, 2\}$ *and* $i \neq j$*, state* i *always proposes* x_i^**, accepts* $x_j \leq x_j^*$*, rejects* $x_j > x_j^*$*, and chooses its SPE strategy* a_i^* *in* Γ*, and where*

$$x_i^* = r_i^* + (1 - \delta_j)w$$

with $r_i^* \equiv r_i(a_1^*, a_2^*)$ *and* $w = (\pi - r_1^* - r_2^*)/(1 - \delta_1\delta_2) \geq 0$*. The outcome in all cases is that agreement is reached immediately on the division* $(x_1^*, \pi - x_1^*)$ *and no fighting occurs.*

Since the peace outcome is a SPE of Γ, this proposition implies that the bargaining game has at least one stationary no-delay SPE (SSPE), which is efficient: agreement is reached immediately at $t = 0$ and no fighting occurs.

Corollary 3.2. *The status quo distribution* (s_1, s_2) *can be supported in SSPE.*

Although this game has efficient peaceful equilibria, they are not the only ones. Clearly, any inefficient SPE will involve some playing in Γ that is not Nash (and there-

[9] The analysis relies heavily on Muthoo (1999), which is in turn based on Busch and Wen (1995). In particular, Proposition 3.1 is an adapted version of Proposition 6.5 in Muthoo's (1999) book.

fore not subgame perfect either). I now turn to the analysis of efficient non-stationary SPE.

The strategy, which is a standard game-theoretic argument, is as follows. First, I show that a minimum payoff that is less than the status quo can be supported in an efficient equilibrium. Then, I demonstrate how the threat to revert to the efficient SPE that yields that payoff to the deviating party can keep states from deviating in inefficient equilibria. That is, equilibria, in which costly fighting occurs for nonzero periods of time.

To establish the result stated in Proposition 3.4, it is necessary to go through some preliminaries. Because payoffs are time-invariant, all subgames that begin with a proposal by player i are structurally identical, which implies that the sets of SPE are the same. Let G_i denote a subgame that begins with a proposal by player i. It is useful to characterize the entire set of SPE payoffs for G_i. The existence of SPE established by Proposition 3.1 allows the following definition by ensuring that we are not defining the empty set.

Definition 3.1. Let $U_i = \left\{u_i : \textit{there exists a SPE of } G_i \textit{ with payoff } u_i\right\}$ be the set of payoffs attainable in SPE, and let $m_i = \min U_i$ and $M_i = \max U_i$ be the minimum and maximum payoffs respectively.

Since any outcome that involves conflict is, by assumption, worse than peace, state 2 can keep state 1 strictly below its status quo payoff by fighting in Γ. We can therefore define the minimax payoffs of Γ as follows

Definition 3.2. Let $\underline{v}_1 = \min_{a_2} \max_{a_1} r_1(a_1, a_2) < s_1$ be player 1's worst payoff that player 2 can impose. Define $\underline{v}_2 < s_2$ analogously.

We are now in a position to state the lower bound on the minimum payoff that players can obtain in any SPE. The lemma, whose proof is the appendix, is stated for player 1's payoff, and a symmetric argument establishes a similar result for player 2.

Lemma 3.3. *$m_1 \geq \underline{v}_1 + (1-\delta_2)w_1$, where $w_1 = (\pi - \underline{v}_1 - \overline{v}_2)/(1-\delta_1\delta_2)$, and*

$$\overline{v}_2 = \max_{a_1,a_2} \left\{ r_2(a_1,a_2) - \frac{(1-\delta_1)\delta_2}{\delta_1(1-\delta_2)} \left(\max_{a_1'} \left\{ r_1(a_1', a_2) \right\} - r_1(a_1,a_2) \right) \right\} \tag{1}$$

It is useful to define the pair of actions that produces player 1's minimax outcome in Γ and the SPE pair of actions that produces the peaceful outcome.

Definition 3.3. Let $\underline{\sigma} \equiv (\underline{a}_1, \underline{a}_2)$ be an action profile in Γ such that $r_1(\underline{\sigma}) = \underline{v}_1$. Let $\overline{\sigma} \equiv (\overline{a}_1, \overline{a}_2)$ be the SPE in Γ, such that $\overline{v}_i = s_i$.

I now turn to the main proposition of this section, which establishes that player 1's worst (respectively player 2's best) payoff can be supported in SPE. The proof is in the appendix.

Proposition 3.4 (Extremal SPE). *With sufficiently little discounting, the bargaining game has subgame perfect equilibria, in which players obtain the smallest payoffs from Lemma 3.3:*

1. *If $\delta_1 \geq \delta_2$, the following strategies support a subgame perfect equilibrium, in which agreement $(x_1^*, \pi - x_1^*)$ is obtained immediately:*

   (A) Player 1 always offers $x_1^* = \underline{v}_1 + (1 - \delta_2)w_1$, accepts $x_2 \leq x_2^*$, and rejects $x_2 > x_2^*$. Player 2 always offers $x_2^* = \overline{v}_2 + (1 - \delta_1)w_1$, accepts $x_1 \leq x_1^*$, and rejects $x_1 > x_1^*$. If player 1 rejects an offer, then play $\underline{\sigma}$ in Γ, and if player 2 rejects an offer, then play $\overline{\sigma}$. If player 2 offers $x_2 > x_2^*$, rejects $x_1 \leq x_1^*$, or plays differently in Γ, then switch to phase (B)*;*

   (B) Play the SSPE from Corollary 3.2.

2. *If $\delta_1 \leq \delta_2$, the following strategies support a subgame perfect equilibrium, in which agreement is obtained only in G_2:*

   (A) Player 1 always makes nonserious offers (e.g. $x_1^* = \pi$), accepts $x_2 \leq x_2^*$, and rejects $x_2 > x_2^*$. Player 2 always offers $x_2^* = (1 + \delta_1)^{-1}(\pi - \underline{v}_1 + \delta_1 \overline{v}_2)$, accepts $x_1 \leq \underline{x}_1 = (1 + \delta_1)^{-1}(\overline{v}_1 + \delta_1 \underline{v}_1)$, and rejects $x_1 > \overline{x}_1$. If player 1 rejects an offer, then play $\underline{\sigma}$ in Γ, and if player 2 rejects an offer, then play $\overline{\sigma}$. If player 2 offers $x_2 > x_2^*$, rejects $x_1 \leq \underline{x}_1$, or plays differently in Γ, then switch to phase (B)*;*

   (B) Play the SSPE from Corollary 3.2.

In Type I equilibrium from Proposition 3.4, it is clear that when δ_2 becomes sufficiently high, the equilibrium payoff converges to $\underline{v}_1$, which by definition is the smallest payoff from Γ. The strategies do not support an equilibrium if $\delta_1 < \delta_2$. However, notice that when $\delta_1 = \delta_2$ both equilibria are payoff-equivalent, which implies that in Type II equilibrium and sufficiently high δ_1, the payoff is arbitrarily close to the payoff from a Type I equilibrium.

Proposition 3.4 highlights two important features of the bargaining model. First, regardless of the relationship of the two discount factors, it is possible to sustain in equilibrium payoffs that are strictly smaller than the peaceful SSPE from Corollary 3.2. This is an essential result because the threat to revert to an extremal SPE (ESPE) with the low payoff is sufficient to support fighting in equilibrium, as the next section demonstrates. Second, this proposition illustrates that the common assumption of equal discount factors is not harmless. Although intuitively one should expect the results to depend on how the two players value the future, and especially on whether their valuations diverge, bargaining models usually assume away the difference. The richness of this model, however, introduces a level of complexity that must somehow be reduced. I now examine the implications of the two types of extremal SPE identified by Proposition 3.4 and how a player's smallest payoff differ between them.

IV. Inefficient Equilibria: The Fear of Settlement

Consider player 1's smallest and largest sustainable payoffs. If $\delta_1 \geq \delta_2$, then (using Type I ESPE), the smallest payoff obtained immediately in a subgame beginning with player 1's offer is[10]

$$\underline{v}_1 + \frac{1-\delta_2}{1-\delta_1\delta_2}(\pi - \underline{v}_1 - \overline{v}_2) = \underline{v}_1 + \frac{1-\delta_2}{1-\delta_1\delta_2}(\overline{v}_1 - \underline{v}_1)$$

Since player 2's smallest payoff, which can be supported by a Type II ESPE, is $(1+\delta_2)^{-1}(\overline{v}_2 + \delta_2\underline{v}_2)$, then the largest payoff that player 1 can get is

$$\pi - \frac{\overline{v}_2 + \delta_2\underline{v}_2}{1+\delta_2} = \overline{v}_1 + \frac{\delta_2(\overline{v}_2 - \underline{v}_2)}{1+\delta_2}$$

If $\delta_1 < \delta_2$, then (using Type II ESPE), the smallest payoff obtained in games beginning with player 1's offer is

$$(1-\delta_1)\overline{v}_1 + \delta_1(\pi - x_2^*) = (1-\delta_1)\overline{v}_1 + \frac{\delta_1(\underline{v}_1 + \delta_1\overline{v}_1)}{1+\delta_1} = \frac{\overline{v}_1 + \delta_1\underline{v}_1}{1+\delta_1}$$

Note that the payoff includes the first term because agreement is obtained only in the subgame that begins with player 2's offer. Since player 2's smallest payoff, which can be supported by a Type I ESPE, is $\underline{v}_2 + (1-\delta_1\delta_2)^{-1}(1-\delta_1)(\overline{v}_2 - \underline{v}_2)$, then the largest payoff that player 1 can get is

$$\pi - \underline{v}_2 - \frac{(1-\delta_1)(\overline{v}_2 - \underline{v}_2)}{1-\delta_1\delta_2} = \overline{v}_1 + \frac{\delta_1(1-\delta_2)}{1-\delta_1\delta_2}(\overline{v}_2 - \underline{v}_2)$$

[10] This is always larger than the minimum payoff that can be obtained in a subgame that begins with player 2's offer. This result, due to discounting, is common to bargaining models of this type.

It is useful to condense these results into manageable definitions. To this end, let player 1's minimum payoff that can be sustained in ESPE be

$$\underline{s} = \begin{cases} \underline{v}_1 + (1-\delta_1\delta_2)^{-1}(1-\delta_2)(\overline{v}_1 - \underline{v}_1) & \text{if } \delta_1 \geq \delta_2 \\ (1+\delta_1)^{-1}(\overline{v}_1 + \delta_1\underline{v}_1) & \text{if } \delta_1 < \delta_2 \end{cases} \tag{2}$$

and let player 1's maximum payoff that can be sustained in ESPE be

$$\overline{s} = \begin{cases} \overline{v}_1 + (1+\delta_2)^{-1}\delta_2(\overline{v}_2 - \underline{v}_2) & \text{if } \delta_1 \geq \delta_2 \\ \overline{v}_1 + (1-\delta_1\delta_2)^{-1}\delta_1(1-\delta_2)(\overline{v}_2 - \underline{v}_2) & \text{if } \delta_1 < \delta_2 \end{cases} \tag{3}$$

Claim 4.1. *$\underline{s} < \overline{s}$ for all $\delta_1, \delta_2 \in (0,1)$.*

Proof. Consider the case $\delta_1 \geq \delta_2$. Since for any $\delta_1 < 1$, $\alpha \equiv (1-\delta_1\delta_2)^{-1}(1-\delta_2) < 1$, we have $\alpha\overline{v}_1 + (1-\alpha)\underline{v}_1 < \overline{v}_1$, where the inequality follows from $\overline{v}_1 > \underline{v}_1$. Thus, $\underline{s} < \overline{v}_1$ for all $\delta_1 < 1$. We also have $\overline{s} > \overline{v}_1$ from (3) because the second term is positive. Therefore, $\underline{s} < \overline{s}$, as required.

Consider now the case $\delta_1 < \delta_2$. Since $\underline{s} < \overline{v}_1$ for any $\delta_1 > 0$, and $\overline{s} > \overline{v}_1$ for any $\delta_2 < 1$, the result is established. Q.E.D.

It is instructive to note that in the limit, $\underline{s}$ converges to $\frac{1}{2}(\overline{v}_1 + \underline{v}_1)$, and $\overline{s}$ converges to $\frac{1}{2}(\pi + \overline{v}_1 - \underline{v}_2)$. With these results, I can now state the principal proposition.

Proposition 4.1 (Inefficient SPE). *For any distribution of benefits $s \in (\underline{s}, \overline{s})$, some period $\tau > 0$, and sufficiently little discounting, there exist equilibria of the following type: in all periods $t = (0, 1, \ldots, \tau - 1)$, both players make nonserious offers, reject all proposals, and fight in Γ. In period τ, they agree on the distribution $(s, \pi - s)$. If either player deviates, then immediately play the ESPE from Proposition 3.4 that supports that player's smallest payoff.*

Proof. Recall that the payoff from each player's worst outcome from fighting is normalized to 0. If player 1 follows the proposed strategy, then its payoff is at least $\delta_1^\tau s$. Suppose it deviates in some $t < \tau$, then its payoff is at most $(1-\delta_1^t)\overline{v}_1 + \delta_1^t\underline{s}$, which in the limit, as $\delta_1 \to 1$, converges to $\underline{s} < s$. Thus, for sufficiently high δ_1, deviation is not profitable. The proof for player 2 is equivalent, *mutatis mutandis*. Q.E.D.

The intuition for the result is straightforward. Although there exist many agreements that would be Pareto-improving (such as s reached in periods prior to T, among others), these potentially better settlements cannot be achieved because of the way players respond to deviations. It is the threat to revert to an efficient SPE, which yields the potential deviator its worst possible payoff, that sustains these inefficient equilibria. The result does not disappear when the time between period becomes arbitrarily small, as is usually the case with bargaining models. This is because shortening the periods between offers not only makes alternation faster, but also decreases the costs that players suffer due to fighting. Players still fight because not doing so means obtaining a worse settlement. There is no incomplete information in the model, and so fighting cannot serve as a signal to separate types (which is what delay usually serves to accomplish in bargaining models with private information). Thus, even with quickly alternating offers, there exist equilibria, in which there is a delay in reaching an agreement. This inefficiency is contrary to the so-called Coase Conjecture in economics, which posits that even with incomplete information the ability to alternate offers quickly should produce agreements without costly delay (Coase 1972).

It is worth noting that there exist other equilibria, in which fighting is sporadic, and they are all supported by similar strategies. As Busch and Wen (1995) note, strategic interaction outside the bargaining process will generally generate multiple equilibria. This indeterminacy makes empirical testing a difficult proposition. However, as the purpose of this article is to demonstrate a logical result and point to a mechanism that is neglected in explanations of conflict, it is no great handicap. In Section 5, I give several plausible historical examples that relate to the conclusions from the analysis of this model.

1. Renegotiation of Agreements

Since agreements, even when reached, do not terminate the relations between states, one may wonder what would happen if we allowed players to continue bargaining after they redistribute the flow of benefits. In this section, I extend the model to allow the players to renegotiate an agreement after some offer is accepted. Thus, every time some distribution is adopted, it becomes the new status quo and states play the bargaining game again.

It is now necessary to be more precise about the conflict game Γ that states play each period until an agreement is struck. If we suppose that the payoffs in Γ remain unchanged, then it is possible to extend the analysis in the previous section to the game with renegotiations. All of the equilibrium outcomes are also equilibrium outcomes in the modified model. Since one of the most persistent themes in international relations is the fear of an opponent gaining an advantage and then exploiting it, this simple solution seems unattractive because it requires the assumption that the conflict game does not change after an agreement. Therefore, it seems reasonable to assume that, following some settlement, the payoffs in Γ also change to reflect the new distribution of benefits.

Suppose then that after an agreement is struck, the new distribution can be supported in a peaceful equilibrium as before. Further suppose that the minimax payoff is strictly increasing in the share of resources controlled by the player: Let s_1 and s_1' be two shares for player 1, and let $\underline{v}_1$ and $\underline{v}_1'$ be the minimax payoffs that player 2 can impose in Γ when the distribution is $(s_1, \pi - s_1)$ and $(s_1', \pi - s_1')$ respectively. The assumption states that if $s_1 < s_1'$, then $\underline{v}_1 < \underline{v}_1'$.

Consider a model where the original bargaining game is played and then, after an agreement is reached on some distribution $(s_1^*, \pi - s_1^*)$, the same game is played again with the new distribution as the status quo and the payoffs in Γ constrained as described. Given the new status quo, all classes of equilibria of the original model are also equilibria classes in the subgame following the first settlement in the new model.

Suppose $s_1^* < s_1$, that is, in the interim settlement, player 1's share is strictly less than the initial status quo. Under the assumptions on Γ, this implies that $\overline{v}_1^* < \overline{v}_1$ and $\underline{v}_1^* < \underline{v}_1$. Since $\underline{s}$ is increasing both in $\overline{v}_1$ and in $\underline{v}_1$, it follows that $\underline{s}^* < \underline{s}$. That is, the smallest payoff that player 1 can obtain in equilibrium is now worse than before. In addition, note that $\overline{v}_1^* < \overline{v}_1 \Rightarrow \overline{v}_2^* > \overline{v}_2$, and by a similar argument it follows that player 2's smallest payoff is larger than before. But this implies, together with the fact that $\overline{s}$ is increasing in $\overline{v}_1$, that player 1's largest payoff sustainable in equilibrium is now smaller ($\overline{s} < \overline{s}'$). In other words, player 1's negotiating position has become more tenuous because its ability to withstand punishment and impose costs has decreased, along with the benefits it enjoys from the peaceful status quo. As one would expect intuitively, accepting some offer that leaves the player worse off has deleterious consequences for the player's ability to bargain in subsequent negotiations.

However, as the following example demonstrates, it is nevertheless possible to have equilibria in which player 1 settles for a lower interim payoff in an expectation that it will obtain a better agreement when it comes time to renegotiate the terms. There are, however, limits on the range of interim agreements that it will accept.

Consider the following equilibrium strategies: player 1 offers $x_1 < \underline{v}_1$, which player 2 accepts, and the new status quo becomes s_1'. Then, τ periods later, player 1 offers $x_1' > \underline{v}_1$, which player 2 accepts. If player 1 deviates, then play the ESPE from Proposition 3.4 that yields it $\underline{s}'$. If player 2 deviates, then play the ESPE that yields player 1 $\overline{s}'$.

Player 1's payoff in this equilibrium is

$$(1 - \delta_1^\tau)x_1 + \delta_1^\tau x_1' = x_1 + \delta_1^\tau(x_1' - x_1) \geq \underline{v}_1$$

where the inequality follows from Lemma 7.1 and is a trivial consequence of the fact that player 1 can guarantee itself the original minimax payoff by always making nonserious offers, rejecting all proposals, and playing the appropriate strategy in Γ. Thus, $x_1' > \underline{v}_1$ is a necessary condition for this equilibrium.

In order to induce player 2 to accept x_1' in period τ, player 1 should be able to threaten it with a worse payoff. By the argument above, player 2's worst payoff sustainable in equilibrium is larger (since $\overline{s} > \overline{s}'$), so the threat cannot be too severe. Thus, $x_1' \leq \overline{s}'$ is another necessary condition for this equilibrium. Thus, we have the following condition:

$$\underline{v}_1 < x_1' \leq \overline{s}'$$

which implies that x_1 cannot be too small. This is because $\overline{s}$ is decreasing in the status quo payoff, which implies that if x_1 is too small, $\overline{s}'$ will be correspondingly small, possibly less than $\underline{v}_1$, thus violating the necessary condition for the equilibrium. Therefore, there exists a lower bound on the range of acceptable interim agreements.

The intuition for this result is clear. A player will never accept an interim agreement which leaves it in such a bad position that it cannot hope to recoup its losses. Thus, although it is possible to have equilibria where players settle for less in the expectation that they will obtain a better bargain tomorrow, there are limits to how much they will agree to. These limits are due to the fear that the opponent will use the temporary advantage for future gain. Hence, this result lends credibility to corresponding arguments in international relations theory.

2. *Costs of Fighting and Mutual Coercion*

It is interesting to see how the range of distributions supportable in inefficient SPE varies with the costs of fighting. From equation (2), it is evident that $\underline{s}$ is increasing in player 1's minimax payoff $\underline{v}_1$. Thus, when the cost of fighting decreases, so that player 1 can guarantee itself a higher payoff during conflict, the lower bound on possible agreements increases, improving player 1's worst-case payoff. This result is hardly surprising, and is consistent with findings supporting the conclusion that decreasing the costs of fighting improves a state's bargaining position.

From equation (3), the upper bound on possible agreements is decreasing in $\underline{v}_2$. In other words, as player 2's minimax payoff increases, player 1's best payoff decreases. Since the minimax payoff is by definition the worst payoff that player 1 can impose on player 2, it follows that the capacity to impose costs on its opponent improves player 1's bargaining position by expanding the upper bound on the range of agreements.

This highlights an important aspect of conflict: the ability to bear costs associated with the opponent's effort to inflict pain influences how much a state can be expected to give up in a bargain. The capacity to inflict pain on the opponent, determined by the ability to bear costs associated with this effort influences how much a state can demand in a bargain. Thus, the two types of costs jointly determine the bargaining range that opens up during fighting.

The capacity to minimax the enemy and the magnitude of costs a state can pay when being minimaxed against emerge as central elements in the explanation for war. The bargaining strategies that condition on fighting behavior and that depend on these costs can result in equilibria that may explain many puzzling cases such as those where states have given up fighting after suffering relatively minor casualties or continued fighting in the face of mounting losses.

For example, in 1940 the French surrendered Paris to Germany after suffering about 90,000 dead and 250,000 wounded. This was the same nation that had defended the fortress of Verdun in 1916 at the cost of about 400,000 casualties against the same enemy (Keegan 1999, esp. pp. 71-137). The crucial difference was that unlike its effect in the First World War, the German onslaught in the Second managed to destroy the organizational capacity of the French high command and damaged beyond repair the French ability to inflict substantial losses on the Germans (May 2000). Facing circumstances in which fighting is hopeless because it cannot induce the other side to accept a settlement makes surrender a rational choice. On the other hand, the continued resistance of North Vietnam despite the heavy toll extracted by the United States can also be explained by the simple fact that the Communists could cause enough damage to induce their opponent to withdraw. Similarly, the Afghani rebels could outlast the Soviet occupational forces despite heavy casualties for essentially the same reasons.

The process of war can be usefully viewed as a contest, in which both sides attempt to reduce the opponent's ability to impose costs on them while simultaneously trying to impose costs on the opponent, thereby improving their own bargaining position. Although destroying personnel and materiél may be conducive to diminishing state capacity to fight, it is not necessarily the optimal way of doing so. If a state is weak and/or cannot gain access to the opponent's homeland, its recourse may be to attempt outlasting the enemy. No country, even the richest, can sustain an indefinite involvement in war. Therefore, if a state can deny the opponent the power to hurt (appropriate military tactics, indoctrination, resettlement of population away from urban centers), and simultaneously inflict enough damage to cause the enemy to expand its economic and troop involvement, then even a weaker state has good chances of success.

V. Discussion

Three central implications emerge from this analysis. First, *states may fight as long as they believe that seeking peace prematurely means accepting an unpalatable settlement.* I should emphasize that in the model presented here, it is not the threat to use force, but the *realization* of the power to hurt, or the *compellent* use of force that is important (Art 1980). As Schelling (1966, p.3) notes, "unhappily, the power to hurt is often communicated by some performance of it." However, it is not that states are unsure about each other's power (after all, this is a model of complete information), but rather it is their commitment to particular strategies that produces the unhappy outcome. As in the familiar Prisoner's Dilemma, the strategies that result in the bad outcome are best responses to each other, so states cannot improve the outcome by deviating and, for example, not fighting. The problem is precisely the existence of efficient, but least preferable, bargains and the fact that if a state does not fight, then it will have to accept a diplomatic solution it likes least.

The rationale for forming such expectations empirically can be justified on the basis of a limited information argument that involves signaling behavior.[11] Standard arguments from signalling models demonstrate that players can use costly actions to signal privately known parameters, such as cost or resolve, to other players. In such an environment, states can conceivably prolong the fight in order to convince their opponents about the value of such parameter. In particular, they might choose to fight when they believe that should they quit to make an offer instead, their opponent would interpret this as a revelation of weakness and exploit it by demanding more. A logic roughly anal-

[11] One way to think about these strategies is to regard the present full information model as the limiting case in a series of incomplete information models, where the amount of uncertainty becomes infinitesimal. Since the equilibria in this model will be present in models with incomplete information, possibly along with many other equilibria, one may consider the result as a sharp characterization of one particular class of inefficient equilibria. Remarkably, these inefficient equilibria do not require incomplete information and thus do not depend on the particular type of uncertainty that one might choose to model.

ogous to this operates in the fill information model, the only difference being that states know exactly what worse bargain they will have to accept if they deviate from the fighting strategy. It is worth noting that this is one possible way to arrive at such strategies. Others, like public opinion and the anticipation of the effect of this opinion, can also be invoked.

Second, *the diminished, or eliminated, capacity to hurt the enemy is a major reason to terminate war and seek a negotiated settlement.* This is a far cry from the conventional notions of victory and defeat because it may not involve the complete destruction of the opponent, only of its ability to retaliate. Ho Chi Minh stated that "You can kill ten of my men for every one I kill of yours. But even at those odds, you will lose and I will win." The key in this statement is not, as usually interpreted, the Vietnamese willingness to suffer, although this undoubtedly played a role (Rosen 1972).[12] Rather, it is their ability to kill one American soldier even if it took ten Vietnamese to do so. The Vietnamese correctly surmised that this power to hurt would eventually compel the US to withdraw. It is doubtful that such a policy would have succeeded had the US been able to limit the number of casualties and expenses.

Schelling (1966) observed that "the power to hurt [...] is a kind of bargaining power." Although as a rule formal models of conflict feature the costs of war as an explanatory variable, and despite the prominent role of that variable in solutions, it is always state ability to *bear* costs that is discussed, but not state capacity to *impose* costs on others. The costs associated with these actions are analytically distinct and not necessarily related in a straightforward manner.

Third, *since the power to hurt is a kind of bargaining power, the denial of such power undermines the bargaining position of the opponent.* It is partially for these reasons why the NATO aerial attacks on Kosovo could succeed. Even the most determined opponent would yield if there were no way to hurt the enemy and thus influence the outcome of negotiations. In some respects, this argument also shows why a ground invasion would have been ill-advised: the mountainous terrain, very much unlike the desert in the Persian Gulf, would not have allowed the easy application of NATO military superiority.[13] There can be no question about the eventual outcome of such an engagement, but it is quite possible that mounting casualties might have tempted the US into a more accom-

[12]Rosen incorrectly cites Ho to mean "American" soldiers. In fact, the original statement referred to the French.

[13]Which is precisely why Saddam Hussein's plan "to draw the Coalition into a premature ground offensive in the hope that heavy casualties would lead Western publics to demand an early cease-fire" could not work (Pape 1996, p. 357).

modating position.[14]

It is readily seen that these three points are interconnected insofar as fighting is a way to influence the final outcome, but as such, it is conditioned on the power to hurt the opponent. It is useful to explore several historical cases in some detail to see how this logic operated in causing the termination of war and in influencing the form of settlements. These cases are only illustrations and serve as checks on the plausibility of the arguments developed here.

First, the Japanese decision to surrender in August 1945 is shown to arise, at least in part, from the realization that the military was not able to inflict sufficient damage on the US forces in a land invasion (they could not minimax the US). Second, the Vietnam War is also examined, showing that since both sides expected to have to accept worse settlements should they stop fighting, they continued the costly conflict until the US was forced to quit due to domestic pressure (the threat of reversion to an efficient but unpalatable agreement). These cases serve to illuminate certain features of the model and are useful examples of how the causal logic might work in practice. Obviously, I do not claim that the model provides a complete explanation, but only that it reveals a mechanism and a relationship that are often neglected.

1. Japan's Surrender in 1945

Japan's decision to surrender in 1945 provides a good example of the complex interaction between the (loss of) power to hurt and the expectations for a better settlement.[15] Significantly, at the time of surrender Japan still had over two million men, armed and ready to fight. Neither the incendiary raids on the major cities nor the nuclear attacks had undermined morale to any perceptible degree.[16] Moreover, despite the Allied suc-

[14]Stam (1999) estimates that the probability of success of a ground invasion was 49%, and argues that this is the reason why it was not attempted. I find the estimate doubtful. According to my analysis, even if the probability of success were 100%, a ground invasion would have entailed significant casualties, and thus given Belgrade bargaining power. Air strikes were ideal precisely because they denied Milosevic the power to hurt NATO forces, and thus maneuver Yugoslavia into a better position at the negotiation table.

[15]The treatment is necessarily cursory and omits many fascinating details of this fateful episode. For an excellent recent authoritative discussion, see Sigal (1988). Two other classic, but somewhat dated, studies are Butow (1954) and Brooks (1968). For a penetrating analysis of the different factors that contributed to the decision to surrender and an evaluation of competing explanations, see Pape (1996).

[16]Although there was some absenteeism from work due to massive flight from the cities, as a whole industrial workers turned up at about the same rate as in the US. Anami's orders of August 10 to the Army exhorted the soldiers "to fight resolutely although that may involve our nibbling grass, eating earth, and sleeping in the fields." There is no reason to think that such orders would not have been followed

cess in “Operation Iceberg,” which captured Okinawa, the major islands still remained in Japanese hands. The unconditional surrender on August 15 is thus a rare instance of a nation capitulating before most of its home territory had been conquered and while there still existed an army that could potentially defend it.

The peace overtures had begun in late spring by Togo, who, through the Japanese Moscow ambassador, had been trying to get the Soviets to mediate with the Allies since May. Although Stalin was not forthcoming with the information (or did not indicate that these attempts were serious), the US was aware of them. However, even after the direct message of the July 12 cable, which stated the desire of the Emperor to end the war, no positive steps were taken in that direction. At any rate, negotiations, had they been undertaken, would most probably have foundered on the firm stand of the Army and the Navy. Ever since the USSR repudiated the Neutrality Pact, the Japanese military were concerned with possible Soviet intervention in the war. They correctly surmised that this would have disastrous consequences and correspondingly tried to prevent such entry. Foreign Minister Togo took a dim view and urged the continuation of the war in an effort to improve the diplomatic situation (Butow 1954, pp. 77-8).

Japan was in a precarious position by the early summer. Although Admiral Suzuki had become Prime Minister on April 6, his wavering attempts to find a way to conclude the war found no support in the hawkish cabinet. It is arguable that it was possible to bring about an earlier surrender through a more assertive action, but the undisputed fact remains that War Minister Anami and the Army Chief of Staff Umezu were formidable opponents very much opposed to cessation of hostilities. On June 6, the document entitled “The Fundamental Policy to Be Followed Henceforth in the Conduct of War” was introduced in the Supreme Council. It flatly called for continuation of fighting on the homeland, stating that this was the only way to preserve the nation. Few believed that Japan could turn the tide of Allied victory, but many thought that an operational success would provide a better basis to negotiate a settlement. In fact, this possibility “became a key point in the military’s demand to continue the war through an Allied invasion of the home islands” (Butow 1954, pp. 93-6).

This policy stemmed from the conviction that despite enormous expected Japanese losses, the Allies would also be badly hurt (Kennedy 1983, p. 193). Thus, the idea was not to defeat US forces militarily (such a thing was “beyond all expectation” as General Miyazaki admitted), but to “inflict tremendous losses, forcing [the enemy] to realize the strong fighting spirit of the Japanese Army and people [and thus] bring about the termi-

(Brooks 1968, p. 11). More on morale can be found in Pape (1996, p. 129-30).

nation of hostilities on comparatively favorable terms."[17] The success of this strategy clearly depended on the ability of the Army to actually inflict the necessary amount of damage.

The illusory nature of this hope was revealed by the Soviet declaration of war on August 9. The militarists, who had refused to convene the Supreme Council after the destruction of Hiroshima, now agreed to discuss whether Japan should accept the Potsdam declaration. Even then Umezu claimed that "we would be able to inflict extremely heavy damage on the enemy," but Togo clearly did not believe it, and it is doubtful that the General believed it himself (Brooks 1968, p. 62-3). Upon hearing the news about the debacle of the Manchurian army, Suzuki remarked, "Is the Kwantung army that weak? Then the game is up" (Brooks 1968, pp. 17). The unexpected but complete collapse of what was supposed to be the elite Japanese fighting force finally convinced even the staunchest militarists (except the fanatical Vice-Admiral Onishi) that their hopes of inflicting sufficient casualties in a ground invasion of the main islands were wishful thinking.[18]

It is worth noting that the timing of this meeting undermines the standard claim that the nuclear bomb was instrumental in the Japanese decision to surrender. The militarists ignored the first bombing and the decision had been made prior to the second one. Significantly, it was the Soviet intervention had precipitated the re-evaluation of the situation.[19]

Despite outlandish claims to the contrary, credible estimates of the number of US casualties for an invasion of Kyushu, one of the home islands, are in the neighborhood of 20,000 (Miles Jr. 1985).[20] There was some apprehension caused by the loss in the

[17]Lieutenant General Arisue. Both his and Miyazaki's statements cited in Sigal (1988, p. 228).

[18]It is true that the divisions in Manchuria were severely weakened by the constant drain of manpower and deterioration of equipment. Still, it is perhaps surprising that the Japanese high command had pinned such hopes on a force that was basically incapable of resisting a determined Soviet attack.

[19]Kecskemeti (1958) reaches essentially the same conclusion about the impact of Stalin's declaration, but for different reasons. His contention is that the Soviet entry in the war dispelled the last illusions the Japanese had of using Stalin as a mediator for extracting better terms from the Americans and thereby forced them to accept whatever was being offered. This seems to require a whole lot of unwarranted hoping by the Japanese high command, especially since Stalin had refused to renew the prewar nonaggression pact earlier in the year, had promised to enter the Pacific War several months after concluding the war in Europe, and could generally be expected to want to take part in the division of the spoils in Asia.

[20]There is a common, but erroneous, belief that the dogged defense of Iwo Jima and Okinawa by the Japanese caused the US to update its beliefs about the costs of invasion, which in turn precipitated the use of the atomic bombs (Giangreco 1997, MacEachin 1998). However, as the military commanders argued, the complete victory was a foregone conclusion.

Okinawa campaign, which took the heaviest toll of the war in the Pacific, and it was used as an early excuse for the atomic bomb (especially by an euphoric Churchill, who called it “a miracle of deliverance”). However, the firm stand that the Allies took with respect to Japanese demands to negotiate the terms reveals that they were prepared to bear this cost in order to secure unconditional surrender.

Thus, although there existed a party that sought the termination of war early on, its efforts were thwarted by the hardliners, who believed that Japan had sufficient capability to defend the home islands. Much of Japanese intransigence in the face of overwhelming odds depended on their belief in their remaining power to hurt the US sufficiently to soften the terms of surrender. As Kecskemeti (1958) points out, the residual capacity to impose costs on the victor makes negotiation desirable by both sides before the ultimate showdown. It seems that in August 1945, the Japanese military finally realized that the estimates of their ability to inflict pain on the US were fantastically exaggerated. The entry of the USSR in the war and the consequent collapse of the vaunted Kwantung Army were perhaps the momentous events that revealed the hollow hopes of the “hawks.” This provided enough ammunition so that peace-seekers were able to overcome the opposition to unconditional surrender and finally end the war.

2. The Vietnam War

Recall that the inefficient (fighting) equilibrium is supported by strategies that require a deviating player to “suffer” the worst acceptable agreement. The Vietnam War provides an illustration of a case in which both sides believed that ending hostilities early would give the other an undue negotiating advantage. The leitmotif of the engagement under Johnson was that neither side made serious attempts to begin negotiations (recall that in the inefficient equilibrium, both sides make nonserious offers throughout the war). Both the United States and North Vietnam wanted to secure enough battlefield success to ensure a favorable outcome of diplomacy. However, the Jason Study indicated as early as 1966 that bombing had no measurable impact on the agricultural economy of North Vietnam and was unlikely to succeed in breaking Vietnamese morale. The Enthoven report of 1967 was even more frank in concluding that the Vietnamese strategy was to keep their losses at “a level low enough to be sustained indefinitely, but high enough to tempt us to increase our forces to the point of U.S. public rejection of the war” (Karnow 1983, p. 519).

General Giap, who was the principal architect of the 1968 Tet offensive, had a long-range strategy, whose principle objective “was to continue to bleed the Americans until they agreed to a settlement that satisfied the Hanoi regime" (Karnow 1983, p. 549). Interestingly enough, the US antiwar movement had played only a secondary role in Giap's planning, but once discovered, the additional means to impose costs on the US was exploited to the full. Giap remarked in the 1990s that he wanted to demonstrate "that if Vietnamese blood was being spilled, so was American blood [...] and more and more Americans renounced the war" (Karnow 1983, p. 557).

As Pape (1996, pp. 177-94) concluded, the failure of Rolling Thunder's (1965-8) alternation between a Schelling strategy of increasing risk to Hanoi's industrial base, a denial by interdiction, and a Douhet plan that focused on increasing current costs, was very simply due to North Vietnam's low vulnerability to these types of coercion strategies. It was not until 1972, when Hanoi switched to conventional warfare (which was militarily vulnerable to US bombing), that coercive air power could compel them to return to the negotiating table. However, even then Kissinger's first attempt failed when South Vietnam's Thieu refused to sign the agreement. The North Vietnamese used this as a pretext to back out of negotiations, which precipitated Nixon's "Christmas" bombing (Linebacker II) to break the deadlock and the attempt to blackmail a better agreement.

While the Communists believed that battlefield victories were the "key factor 'for the attainment of a political settlement'" and believed that they could force the US out by breaking its ties with the Saigon regime and exploiting the antiwar movement, Nixon was convinced that quitting unconditionally would spell the end of Saigon, and perhaps his own reputation, not to mention that it would damage the US international position. Despite Nixon's claims that the US had "finally achieved peace with honor," the Paris treaty of 1973 was an admission of defeat. Although the North Vietnamese demand for a coalition government was dropped, the agreement allowed their troops to remain on South Vietnam soil and virtually guaranteed the overthrow of Saigon which occurred two years later.

The US administration firmly believed that they could not quit the war because leaving Communist aggression unchecked would encourage similar movements elsewhere. It also worried about US political standing in the international arena, especially with regard to its European allies (Strong 1992, pp. 93-4). However, the relentless bombing could not coerce Hanoi into negotiations, and the domestic situation was hardly favorable for a full-scale land invasion. Saigon could not exist without US support, Vietnamization had failed, and the war had no end in sight. Thus, the ability of the United States to hurt North Vietnam was severely limited by the guerilla tactics of the Viet Cong, the agricultural economy (which, unlike industrialized economies cannot be hurt by destroying identifiable objects), the unsophisticated (and easily repairable) infrastructure, and the constant influx of supplies from the USSR and China.[21]

On the other hand, North Vietnam's power to hurt the United States increased with time and the extent of US involvement in the war. The war was costly, both in economic and humanitarian terms. The fearsome combination of rising taxes and a climbing death-toll was the nightmare of the administration, which was also harangued by the antiwar demonstrations. Without prospects of winning (thankfully ignoring Westmoreland's designs to deploy tactical nuclear weapons), and without the means to coerce Hanoi, the US could do little more than accept defeat and withdraw. Even though both sides believed that quitting early would give undue advantage to the enemy, the power to hurt rested with North Vietnam, and so did the eventual settlement.

VI. Conclusion

In this paper I argued that it is misleading to regard war as a costly lottery over exogenously fixed outcomes. I began by noting that both practitioners of war (Clausewitz 1832, Fuller 1961) and scholars (Kecskemeti 1958, Schelling 1966, Wagner 1994) have posited that war is a kind of bargaining process. I constructed a stylized model that incorporates the simultaneous occurrence of negotiation and fighting in order to allow players to condition their strategies on their past behavior. This approach seeks to account for the idea that war is part of a political process during which the participants may change objectives in response to varying military, domestic, and international factors. I completely stacked the model against conflict by assuming that peace can be supported in equilibrium and that fighting brings only costs and no benefits to both sides.

Even under these fairly restrictive assumptions, inefficient fighting can occur in equilibrium and there may be a delay in reaching an agreement. Significantly, this result does not require incomplete information and does not depend on any of the mechanisms from standard rationalist explanations of war. Instead, war becomes possible when states utilize conditional strategies and thus endogenize war aims. The existence of these inefficient equilibria in turn depends on the costs of war.

This article makes a distinction between two types of such costs. The first is the cost that a state can be made to pay when its opponent attempts to hurt it. The second is the cost that a state must pay in order to hurt its opponent. The power to hurt, which

[21] George Ball's memo to Johnson (July 1, 1965) is fairly incisive with respect to the prospects of victory and the costs of alternative outcomes. Excerpts in Strong (1992, pp. 108-9).

turns on the relative magnitude of these costs, and the conditional strategies open up a bargaining range, which can produce fighting in equilibrium under complete information.

These results yielded several implications hitherto underappreciated in research on the causes of war and its termination. First, states may fight as long as they believe that seeking peace early on would result in having to agree to a settlement they dislike. Second, since the power to hurt is a kind of bargaining power, the diminished capacity to inflict pain or suffer losses emerges as a major incentive to seek a negotiated settlement. The denial of such power to an opponent undermines the opponent's bargaining position.

It is important to emphasize that the power to hurt should not be treated in simple military terms. Throughout Napoleon's wars, England, which kept raising and resurrecting coalition after coalition against him, was his principal enemy (Fuller 1961, pp. 55-6). Without means of striking at the island itself, Napoleon resolved to strangle it economically, fully aware that if he could hurt British exports, he would undermine their credit and their ability to pay for the wars against him. The Continental System was designed for precisely that purpose. It is worth noting that lacking a capability to hurt his main enemy through direct miliary engagement, Napoleon correctly inferred that his troubles would never be over unless he found another way of doing so: the System was the weapon to do it. Thus, the power to hurt can take many forms, from military victory, to economic coercion, and humanitarian losses.

The idea that the power to hurt may not translate directly into factors commonly used to measure force, such as military capabilities, geo-political configurations, economic resources, and features of the political system, provides further clues as to why costly conflict may erupt between states with severe power asymmetries (as commonly measured). These conclusions also put in doubt common statistical models of war, which rely on aggregate military and economic capabilities of states. As demonstrated by the empirical cases, these rough indicators may be completely wrong in deducing the outcome of engagements. The power to hurt can take many forms.

This analysis exposes a flaw in standard theoretical explanations of war, which reduce the process to a lottery. Since there exist multiple equilibria that cannot be derived from such models, and since it appears that these equilibria provide useful theoretical and historical insights, one may conclude that the shortcomings of the usual approach are indeed troublesome. These results indicate that we should re-evaluate those theoretical models that treat war as an outcome, and instead think of war as a process where outcomes arise endogenously. Only in this way will it be possible to construct a theory of war termination, which is a necessary step toward a theory of war initiation.

VII. Appendix: Proofs

Proof of Proposition 3.1. In stationary SPE, which are independent of history and time, each player always makes the same proposal, chooses the same attack decision, and responds the same way. This implies that the attack decision must be a subgame perfect equilibrium (SPE) in Γ, for otherwise a player could deviate by choosing an alternative action and improve its payoff. Let (a_1^*, a_2^*) denote an arbitrary SPE in Γ, let $r_i^* \equiv r_i(a_1^*, a_2^*)$ be player i's payoff in that equilibrium outcome, and let w be as defined in the proposition. Note that because $s_1 + s_2 = \pi$ and the highest payoff attainable in equilibrium in Γ is s_i for each player, it follows that $w \geq 0$. The proposals in a stationary no-delay equilibrium must satisfy the equalities

$$\pi - x_2^* = (1 - \delta_1)r_1^* + \delta_1 x_1^*$$
$$\pi - x_1^* = (1 - \delta_1)r_2^* + \delta_2 x_2^*$$

because offering more than the minimum necessary to make the other player indifferent is not optimal. The system of equations has the unique solution stated in the proposition, which implies that there exists at most one stationary no-delay SPE. Note the following identity:

$$\pi - x_2^* = \pi - r_2^* - (1 - \delta_1)w = r_1^* + \delta_1(1 - \delta_2)w. \tag{4}$$

I now show that the strategies specified in the proposition in fact constitute a SPE.

Consider player 1's proposal at some arbitrary time $2t$. Following the equilibrium strategy yields x_1^*. If the player deviates and proposes some $x_1 < x_1^*$, then player 2 accepts, which leaves player 1 worse off. Thus, such deviation is not profitable. Suppose that player 1 deviates by offering $x_1 > x_1^*$, which player 2 rejects. There are three possible outcomes from any alternative strategy: (i) player 1 accepts a proposal by player 2 in some $\tau = 2t + 1$; (ii) player 2 accepts a proposal by player 1 in some $\tau = 2t$; or (iii) players perpetually disagree. Since player 2 always offers x_2^*, in any strategy that results in outcome (i), player 1 can do no better than accept x_2^*. The maximum payoff is then

$$(1-\delta_1^\tau)r_1^* + \delta_1^\tau(\pi - x_2^*) = (1-\delta_1^\tau)r_1^* + \delta_1^\tau(r_1^* + \delta_1(1-\delta_2)w) = r_1^* + \delta_1^{\tau+1}(1-\delta_2)w \leq x_1^*,$$

where the first equality follows from identity (4) and the definition of x_2^*. This implies that such deviation is not profitable. Since player 2 accepts only proposals $x_1 \leq x_1^*$, in any strategy that results in outcome (ii), player 1 can do no better than propose x_1^* at time τ. The maximum payoff is then $(1-\delta^\tau)r_1^* + \delta^\tau x_1^* \leq x_1^*$, where the inequality follows from $r_1^* \leq x_1^*$. This implies that such deviation is not profitable. Finally, in any strategy that results in outcome (iii), player 1's payoff is $r_1^* \leq x_1^*$, which means that such deviation is not profitable. This exhausts all alternative strategies, and hence, the proposal rule is optimal.

Consider now player 1's decision at some time $2t+1$ when it has to respond to a proposal by player 2. Suppose $x_2 < x_2^*$, in which case player 1's payoff is $\pi - x_2$ if it accepts. If player 1 deviates and rejects the proposal, the best payoff it can get is when it offers x_1^* (since we have already established the optimality of the proposal rule). But then $(1-\delta_1)r_1^* + \delta_1 x_1^* = \pi - x_2^* < \pi - x_2$, which implies that such deviation is not profitable. Suppose now $x_2 > x_2^*$ but player 1 deviates and accepts. The payoff then is $\pi - x_2 < \pi - x_2^* = (1-\delta_1)r_1^* + \delta_1 x_1^*$, and so the player is better off rejecting such a proposal and then offering x_1^* in the next period. This establishes the optimality of the acceptance rule.

Thus, player 1's strategy is optimal in every possible subgame given that player 2 uses the strategy described in the proposition. The proof for player 2 is equivalent, *mutatis mutandis*, and therefore the strategies do indeed constitute a stationary no-delay subgame perfect equilibrium of the bargaining game. Since I chose an arbitrary SPE in Γ, it follows there exists a unique stationary no-delay SPE for each SPE in Γ. The uniqueness of the SSPE can be demonstrated by a straightforward application of the method used by Shaked and Sutton (1984). Q.E.D.

Before I proceed with the proof of Proposition 3.4, it is necessary to go through more preliminaries. The following claim is the well-known result in game theory, which states that for each player, the payoff from any equilibrium is at least as large as its minimax payoff.

CLAIM 7.1. *Let $a^* \equiv (a_1^*, a_2^*)$ be an arbitrary Nash equilibrium of Γ. Then $r_i(a^*) \geq \underline{v}_i$.*

Proof. From the definition of the minimax payoff, for all a_2', $\underline{v}_1 \leq \max_{a_1} r_1(a_1, a_2')$. In particular, this implies that $\underline{v}_1 \leq \max_{a_1} r_1(a_1, a_2^*) = r_1(a_1^*, a_2^*)$. This establishes the claim for player 1. The proof for player 2 is analogous. Q.E.D.

The following lemma is a trivial consequence of the observation that each player can guarantee its minimax payoff by making nonserious offers (i.e. demanding π), rejecting every proposal, and using the strategy that maximizes its payoff in Γ while the other player is minimaxing.

LEMMA 7.1. *For all $u_i \in U_i$, $u_i \geq \underline{v}_i$.*

Thus, $\underline{v}_i$ is player i's reservation value because it is the smallest payoff that the other player can keep player i below. Notice in particular that the reservation value is strictly smaller than the status quo distribution of benefits. I now prove Lemma 3.3 for player 1 (a symmetric argument establishes the result for player 2).

Proof of Lemma 3.3[22]. Consider a subgame beginning with Γ followed by G_2 and an arbitrary SPE. Let $\tilde{r}_i \equiv r_i(\tilde{a}_1, \tilde{a}_2)$ be player i's payoff from the SPE profile for Γ, and $\tilde{u}_i$ be player i's payoff from the SPE beginning with G_2. Because this is an equilibrium, player 1 has no incentive to deviate, which necessarily implies that its payoff should be at least as good as the sum of the best outcome in Γ and the worst outcome in G_2, for otherwise player 1 would be able to obtain a better payoff by deviating in Γ. Thus, we have

$$(1-\delta_1)\tilde{r}_1 + \delta_1 \tilde{u}_1 \geq (1-\delta_1)\max_{a_1}\{r_1(a_1, \tilde{a}_2)\} + \delta_1 \mu_1,$$

where μ_1 is the infimum of the set of player 1's SPE payoffs in G_2. Notice that from the definition of the minimax payoff, it follows that $\mu_1 \geq (1-\delta_1)\underline{v}_1 + \delta_1 m_1$. This now implies that

$$(1-\delta_1)\tilde{r}_1 + \delta_1 \tilde{u}_1 \geq (1-\delta_1)\max_{a_1}\{r_1(a_1, \tilde{a}_2)\} + \delta_1\left[(1-\delta_1)\underline{v}_1 + \delta_1 m_1\right],$$

which, together with $\tilde{u}_1 + \tilde{u}_2 \leq \pi$, further implies that

$$(1-\delta_1)\tilde{r}_1 + \delta_1(\pi - \tilde{u}_2) \geq (1-\delta_1)\max_{a_1}\{r_1(a_1, \tilde{a}_2)\} + \delta_1\left[(1-\delta_1)\underline{v}_1 + \delta_1 m_1\right].$$

Solving this inequality yields

$$\tilde{u}_2 \leq \pi - (1-\delta_1)\underline{v}_1 - \delta_1 m_1 - \delta_1^{-1}(1-\delta_1)\Big[\max_{a_1}\{r_1(a_1, \tilde{a}_2)\} - \tilde{r}_1\Big],$$

which implies that player 2's payoff in this SPE has an upper bound:

$$\begin{aligned}(1-\delta_2)\tilde{r}_2 + \delta_2\tilde{u}_2 \leq\ & (1-\delta_2)\tilde{r}_2 + \delta_2[\pi - (1-\delta_1)\underline{v}_1 - \delta_1 m_1] \\ & - \delta_2\delta_1^{-1}(1-\delta_1)\Big[\max_{a_1}\Big\{r_1(a_1, \tilde{a}_2)\Big\} - \tilde{r}_1\Big]\end{aligned}$$

$$= (1-\delta_2)\left[\tilde{r}_2 - \frac{(1-\delta_1)\delta_2}{\delta_1(1-\delta_2)}\left(\max_{a_1}\left\{r_1(a_1,\tilde{a}_2)\right\} - \tilde{r}_1\right)\right] + \delta_2\left[\pi - \left((1-\delta_1)\underline{v}_1 + \delta_1 m_1\right)\right]. \tag{5}$$

Take $\overline{v}_2$ as defined in the lemma, and let $v_2^{**} \equiv \pi - \left((1-\delta_1)\underline{v}_1 + \delta_1 m_1\right)$ be player 2's payoff when player 1 is punished most severely. Since the SPE was arbitrary, (1) and (5) together imply that player 2's payoffs in any SPE are bounded above by $(1-\delta_2)\overline{v}_2 + \delta_2 v_2^{**}$. Therefore, player 2 accepts any proposal x_1 such that $\pi - x_1 > (1-\delta_2)\overline{v}_2 + \delta_2 v_2^{**}$, which implies that in any SPE player 1 cannot do worse than making such a proposal. Hence,

$$m_1 \geq \pi - (1-\delta_2)\overline{v}_2 - \delta_2 v_2^{**},$$

which, after rearranging terms and simplifying, yields the inequality stated in the lemma.[23]

Q.E.D.

Using this result, we can now find the upper bound on the SPE payoffs. The following lemma is stated in terms of player 2's payoffs; using a symmetric argument, we can establish the corresponding result for player 1.

LEMMA 7.2. *$M_2 \leq \overline{v}_2 + (1-\delta_1)w_1$, where w_1 and $\overline{v}_2$ are as defined in Lemma 3.3.*

Proof. Consider some subgame G_2. In any SPE, player 1's minimum payoff from rejecting a proposal by player 2 is $(1-\delta_1)\underline{v}_1 + \delta_1 m_1 \leq u_1$, where the inequality follows from Lemma 7.1 and the fact that $m_1 \leq u_1$ by definition. But now we have

$$\pi - (1-\delta_1)\underline{v}_1 - \delta_1 m_1 \geq \pi - u_1 \geq M_2,$$

where the second inequality follows from $u_1 + u_2 \leq \pi$ (otherwise there exists some $u_2 \leq M_2$ that violates this constraint). Substituting for m_1 from Lemma 3.3 and simplifying completes the proof. Q.E.D.

[23]The upper bound on player 2's payoffs is a convex combination of two terms. The first is (roughly) the difference in player 2's payoffs between Γ when both players follow their equilibrium strategies and Γ when player 1 deviates to a better for itself action (the difference for player 1 is usually called the "temptation" payoff). The second is the surplus when player 1 receives its smallest equilibrium payoff (this is the most player 2 can achieve by minimaxing in Γ and playing the SPE that yields player 1's worst payoff). To put it another way, the first term is player 2's payoff if it plays its equilibrium strategy but player 1 deviates to get its best payoff (and so player 2 suffers the "sucker" payoff in Γ), and the second term is player 2's payoff if it punishes player 1 for the deviation.

Together, the bounds on m_i and M_i represent the worst and best payoffs player i can (possibly) attain in any SPE of any subgame G_i. The next step is to show that the worst payoff for player 1 can be supported in a non-stationary SPE (and by a symmetric argument, an equivalent result for player 2). The two extremal SPE that support player 1 and player 2's worst payoffs are the key to proving the main proposition.

First, some preliminaries. The following claim establishes that in the bargaining model, there exists a pair of strategies that produce $\overline{v}_i$ in Γ.

Claim 7.2. *Let $\sigma^* \equiv (a_1^*, a_2^*)$ be an arbitrary Nash equilibrium of Γ. Then $r_i(\sigma^*) \leq \overline{v}_i$.*

Proof. Since for any Nash equilibrium, $\max_{a_1} \{r_1(a_1, a_2^*)\} - r_1(a_1^*, a_2^*) = 0$, it follows that $\overline{v}_2 \geq r_2(a_1^*, a_2^*)$. The proof for player 1 is symmetric. Q.E.D.

Claim 7.3. *Let σ^* be an arbitrary Nash equilibrium of Γ. Then $\underline{v}_i \leq r_i(\sigma^*) \leq \overline{v}_i$.*

Proof. Follows from Claim 7.1 and Claim 7.2. Q.E.D.

Since the peace outcome can be produced by a subgame perfect, and therefore Nash, equilibrium in the conflict game, Claim 7.2 implies that we can define $\overline{\sigma}$ as in Definition 3.3. Since any player can attack in the conflict game, and since fighting is costly, it follows that $\underline{v}_i < s_i = \overline{v}_i$. This fact establishes the discrepancy between the best attainable and the minimax payoffs that is essential to the main proposition. It is important to note that since $\overline{\sigma}$ is a subgame perfect equilibrium, it follows that no strategy profile (a_1, a_2) with $r_1(a_1, a_2) = \underline{v}_1$ and $r_2(a_1, a_2) = \overline{v}_2$ can be an equilibrium. The same holds for any strategy profile that produces the best payoff for player 1 and the minimax payoff for player 2.

Note that $\underline{\sigma}$ from Definition 3.3 may not be a subgame perfect equilibrium, and so $r_2(\underline{a})$ may not equal $\underline{v}_2$. I now turn to the proof of the main proposition of this section, which establishes that player 1's worst (respectively player 2's best) payoff can be supported in SPE.

Proof of Proposition 3.4. I shall establish the result for the extremal SPE that support player 1's worst payoff. The equivalent result for player 2 follows from a symmetric argument. The optimality of phase (B) in both cases follows from Proposition 3.1. In now prove the optimality of phase (A).

CASE 1: $\delta_1 \geq \delta_2$. First, note the following useful identities:

$$\pi - x_1^* = \pi - (\underline{v}_1 + (1-\delta_2)w_1) = \overline{v}_2 + \delta_2(1-\delta_1)w_1$$
$$\pi - x_2^* = \pi - (\overline{v}_2 + (1-\delta_1)w_1) = \underline{v}_1 + \delta_1(1-\delta_2)w_1$$

Consider player 1's strategy in Γ after player 1 rejects an offer. Given that player 2 is minimaxing, player 1 cannot improve its payoff, hence it follows from the definition of $\underline{v}_1$ that no one-shot deviation can be profitable. Consider now player 1's actions in Γ after player 2 rejects an offer. Since next period 2 offers $\pi - x_2^*$ regardless of 1's actions in Γ, and because the proposed strategy yields the highest payoff of $\overline{v}_1$ there, it follows that no deviation can be profitable. This establishes the optimality of the rule for play in Γ.

Consider now player 1's strategy in some G_1. If player 1 follows the proposed strategy, the payoff is x_1^*. To see that player 1 would not deviate by offering $x_1 < x_1^*$, notice that since player 2 accepts all such offers, player 1 is strictly worse off. Hence, such deviation is not profitable. Suppose player 1 deviates and offers some $x_1 > x_1^*$, which player 2 always rejects. Player 1's payoff from such deviation is

$$(1-\delta_1)\overline{v}_1 + \delta_1(\pi - x_2^*) = (1-\delta_1)\overline{v}_1 + \delta_1[\underline{v}_1 + \delta_1(1-\delta_2)w_1] \leq \underline{v}_1 + (1-\delta_2)w_1 = x_1^*$$

For the inequality to hold, it must be the case that

$$(1-\delta_1)\overline{v}_1 - \underline{v}_1 \leq (1-\delta_2)w_1 - \delta_1[\underline{v}_1 + \delta_1(1-\delta_2)w_1]$$

which simplifies to

$$\overline{v}_1 - \underline{v}_1 \leq \frac{(1+\delta_1)(1-\delta_2)}{1-\delta_1\delta_2}\left(\pi - \underline{v}_1 - \overline{v}_2\right) \tag{6}$$

Since $\pi - \overline{v}_2 = \overline{v}_1 \Rightarrow \pi - \underline{v}_1 - \overline{v}_2 = \overline{v}_1 - \underline{v}_1$, a sufficient condition for (6) to hold is

$$\frac{(1+\delta_1)(1-\delta_2)}{1-\delta_1\delta_2} \geq 1$$

which, after simplifying, becomes $\delta_1 \geq \delta_2$, which is true by assumption. Hence, such deviation is not profitable. This establishes the optimality of the proposal rule.

Consider now player 1's strategy in some G_2. Suppose player 1 has to respond to some $x_2 < x_2^*$. If it accepts, the payoff is $\pi - x_2$, if it deviates and rejects, the payoff is

$$(1-\delta_1)\underline{v}_1 + \delta_1 x_1^* = \underline{v}_1 + \delta_1(1-\delta_2)w_1 = \pi - x_2^* < \pi - x_2$$

and hence such deviation is not profitable. Suppose now player 1 has to respond to some $x_2 > x_2^*$, in which case the game switches to phase (B). If 1 deviates and accepts, the payoff is $\pi - x_2$. If it follows the proposed strategy, the payoff is at least

$$(1-\delta_1)\underline{v}_1 + \delta_1[\underline{v}_1 + (1-\delta_2)w_1] = \underline{v}_1 + \delta_1(1-\delta_2)w_1 = \pi - x_2^* > \pi - x_2$$

where the term in the square brackets is lower bound on the payoff player 1 can obtain in any SPE of G_1. Hence, such deviation is not profitable. This establishes the optimality of the acceptance rule.

Since this exhausts all possible subgames, the proposed strategy for player 1 is a best response to player 2's strategy. I now show that player 2's strategy is also optimal.

Consider player 2's strategy in Γ. If it follows the proposed strategy, the payoff in the next period is $\min\{x_2^*, \pi - x_1^*\} = \pi - x_1^*$. Therefore, 2's payoff from following the strategy is at least

$$(1-\delta_2)r_2(\underline{\sigma}) + \delta_2(\pi - x_1^*) = (1-\delta_2)r_2(\underline{\sigma}) + \delta_2\Big[\overline{v}_2 + \frac{\delta_2(1-\delta_1)}{1-\delta_1\delta_2}\Big(\pi - \underline{v}_1 - \overline{v}_2\Big)\Big] \quad (7)$$

If player 2 deviates, then the game switches to phase (B), where player 2's payoff is s_2. Hence, the deviation payoff is at most

$$(1-\delta_2)\overline{v}_2 + \delta_2 s_2 = \overline{v}_2 \quad (8)$$

In the limit, as $\delta_2 \to 1$, the expression in (7) converges to $\pi - \underline{v}_1 > \pi - \overline{v}_1 = \overline{v}_2$.[24] Therefore, for each δ_1 sufficiently high, there exists $\delta_2' \leq \delta_1$ such that for all $\delta_2 \in [\delta_2', \delta_1]$ deviation is not profitable. This establishes the optimality of the rule for play in Γ.

Consider now player 2's strategy in some G_2. If player 2 follows the proposed strategy, the payoff is x_2^*. To see that player 2 would not deviate by offering $x_2 < x_2^*$, note that since player 1 accepts all such offers, this would leave player 2 strictly worse off. Hence, such deviation is not profitable. Suppose player 2 deviates and offers some $x_2 > x_2^*$. Player 1 always rejects such proposals and the game switches to phase (B). Therefore, player 2's payoff from such deviation is at most

$$(1-\delta_2)r_2(\underline{\sigma}) + \delta_2 s_2 < \overline{v}_2 \quad (9)$$

where the inequality follows from $r_2(\underline{\sigma}) < \overline{v}_2$. Since x_2^* is strictly greater than (7), it

[24]If $\delta_1 = \delta_2$, then it converges to $\frac{1}{2}(\pi - \underline{v}_1 + \overline{v}_2) > \overline{v}_2$, where the inequality follows from $\pi - \underline{v}_1 > \overline{v}_2$.

follows that there exists some $\delta_2'' \in (0,1)$ such that for all $\delta_2 \in (\delta_2'', 1)$, x_2^* is strictly greater than (9). Hence, such deviation is not profitable. Taking $\underline{\delta}_2 = \max\{\delta_2', \delta_2''\}$ yields the discount factor stated in the proposition. This establishes the optimality of the proposal rule.

Consider now player 2's strategy in some G_1. Suppose player 1 offers some $x_1 < x_1^*$. If player 2 follows the proposed strategy and accepts it, the payoff is

$$\pi - x_1 > \pi - x_1^* = \overline{v}_2 + \frac{\delta_2(1-\delta_1)}{1-\delta_1\delta_2}(\pi - \underline{v}_1 - \overline{v}_2) \tag{10}$$

If it deviates and rejects the proposal, the game switches to phase (B), and its best payoff is given by equation (8), which is smaller than the RHS in (10), which implies it is strictly smaller than $\pi - x_1$. Hence, such deviation is not profitable. Suppose now player 1 offers some $x_1 > x_1^*$. If player 2 follows the proposed strategy and rejects it, the payoff is $(1-\delta_2)\overline{v}_2 + \delta_2 x_2^* = \overline{v}_2 + \delta_2(1-\delta_1)w_1 = \pi - x_1^* > \pi - x_1$, where the last term is the deviation payoff from accepting. Hence, such deviation is not profitable. This establishes the optimality of the acceptance rule.

Since this exhausts all possible subgames, the proposed strategy for player 2 is a best response to player 1's strategy. Therefore, the strategies stated in the proposition indeed constitute a subgame perfect equilibrium of the bargaining game.

Case 2: $\delta_1 \le \delta_2$. The optimality of player 1's play in Γ follows from the same argument as in Case 1.

Consider now player 1's strategy in G_1. Given player 2's strategy, if player 1 always makes nonserious offers and rejects all proposals, then player 1 can guarantee a payoff of $\overline{v}_1$ in any period, which begins with its offer, and $\underline{v}_1$ in any period, which begins with player 2's offer. Thus, player 1 can get at least

$$\begin{aligned}(1-\delta_1)(\overline{v}_1 + \delta_1\underline{v}_1 + \delta_1^2\overline{v}_1 + \delta_1^3\underline{v}_1 + \dots) &= (1-\delta_1)(\overline{v}_1 + \delta_1\underline{v}_1)(\delta_1^0 + \delta_1^2 + \delta_1^4 + \dots) \\ &= (1+\delta_1)^{-1}(\overline{v}_1 + \delta_1\underline{v}_1) \equiv \underline{x}_1\end{aligned}$$

Since player 2 only accepts $x_1 \le \underline{x}_1$, player 1 cannot improve its payoff from making a proposal that will be accepted. Therefore, player 1's strategy of making a nonserious offer is optimal.

Consider now player 1's strategy in G_2. Suppose player 2 offers $x_2 \le x_2^*$. If player 1 follows the proposed strategy and accepts, the payoff is at least $\pi - x_2^*$. If it is profitable to deviate and reject, then the following are true:

$$
\begin{aligned}
\pi - x_2^* &< (1-\delta_1)\underline{v}_1 + \delta_1(1-\delta_1)\overline{v}_1 + \delta_1^2(\pi - x_2^*) \\
(1-\delta_1^2)(\pi - x_2^*) &< (1-\delta_1)(\underline{v}_1 + \delta_1\overline{v}_1) \\
x_2^* &< (1+\delta_1)^{-1}(\pi - \underline{v}_1 + \delta_1\overline{v}_2)
\end{aligned}
$$

The last statement is false, which establishes the contradiction. Hence, deviation is not profitable. Suppose now player 2 offers $x_2 > x_2^*$, in which case the game switches to phase (B). If player 1 follows the proposed strategy and rejects, the payoff is $(1-\delta_1)\underline{v}_1 + \delta_1\overline{v}_1$. If it deviates and accepts, the payoff is $\pi - x_2 < \pi - x_2^* = (1+\delta_1)^{-1}(\underline{v}_1 + \delta_1\overline{v}_1) \leq (1-\delta_1)\underline{v}_1 + \delta_1\overline{v}_1$. Hence, such deviation is not profitable. Therefore, the acceptance rule is optimal.

Since this exhausts all possible subgames, the proposed strategy for player 1 is a best response to player 2's strategy. I now show that player 2's strategy is also optimal.

Consider now player 2's strategy in Γ. If it deviates, the game switches to phase (B), so the maximum payoff is $(1-\delta_2)\overline{v}_2 + \delta_2\overline{v}_2 = \overline{v}_2$. If player 2 follows the proposed strategy, then, since $\overline{v}_2 < x_2^* \Rightarrow (1-\delta_2)\overline{v}_2 + \delta_2 x_2^* \leq x_2^*$, the payoff is at least

$$(1-\delta_2)r_2(\underline{\sigma}) + \delta_2[(1-\delta_2)\overline{v}_2 + \delta_2 x_2^*]$$

which, as $\delta_2 \to 1$, converges to $x_2^* > \overline{v}_2$. Therefore, there exists $\delta_2' \in (0,1)$ such that for all $\delta_2 \in (\delta_2', 1)$, deviation is not profitable. This establishes the optimality of the rule for play in Γ.

Consider now player 2's strategy in G_1. Since player 1 never offers $x_1 < \underline{x}_1$, player 2 must decide whether to reject some $x_1 \geq \underline{x}_1$. Suppose it were profitable for player 2 to deviate and accept player 1's offer. Then, the following must be true

$$
\begin{aligned}
\pi - \underline{x}_1 &> (1-\delta_2)\overline{v}_2 + \delta_2 x_2^* \\
(1+\delta_1)^{-1}(\delta_1\pi - \delta_1\underline{v}_1 + \overline{v}_2) &> (1-\delta_2)\overline{v}_2 + \delta_2(1+\delta_1)^{-1}(\pi - \underline{v}_1 + \delta_1\overline{v}_2) \\
\delta_1\pi - \delta_1\underline{v}_1 + \overline{v}_2 &> (1+\delta_1)(1-\delta_2)\overline{v}_2 + \delta_2(\pi - \underline{v}_1 + \delta_1\overline{v}_2) \\
\delta_1\pi - \delta_1\underline{v}_1 &> \delta_1\overline{v}_2 + \delta_2(\overline{v}_1 - \underline{v}_1) \\
\delta_1(\overline{v}_1 - \underline{v}_1) &> \delta_2(\overline{v}_1 - \underline{v}_1) \\
\delta_1 &> \delta_2
\end{aligned}
$$

Because $\delta_1 \leq \delta_2$, the last statement is false, which establishes the contradiction. Therefore, player 2 will reject any $x_1 > \underline{x}_1$ and deviation is not profitable. Suppose now player 1 offers some $x_1 \leq \underline{x}_1$. If player 2 follows the proposed strategy and accepts, the payoff is

$$\pi - x_1 \geq \pi - \underline{x}_1 = (1+\delta_1)^{-1}[\overline{v}_2 + \delta_1(\pi - \underline{v}_1)] > (1+\delta_1)^{-1}(\overline{v}_2 + \delta_1\overline{v}_2) = \overline{v}_2$$

If it deviates and rejects, then the game switches to phase (B), where the payoff is $\overline{v}_2$. Hence, the largest deviation payoff is $(1-\delta_2)\overline{v}_2 + \delta_2\overline{v}_2 = \overline{v}_2$. Therefore, such deviation is not profitable. This establishes the optimality of the acceptance rule.

Consider now player 2's strategy in G_2. If it follows the proposed strategy, the payoff is x_2^*. Deviation by offering $x_2 < x_2^*$ is not profitable because player 1 accepts all such offers. Suppose player 2 deviates by offering $x_2 > x_2^*$, which player 1 always rejects. Since the game switches to phase (B), the payoff then is $(1-\delta_2)r_2(\underline{\sigma}) + \delta_2\overline{v}_2 \leq \overline{v}_2$. But since $\underline{v}_1 < \overline{v}_1 \Rightarrow \pi - \underline{v}_1 > \pi - \overline{v}_1 = \overline{v}_2$, we have $x_2^* = (1+\delta_1)^{-1}(\pi - \underline{v}_1 + \delta_1\overline{v}_2) > \overline{v}_2$. Hence, deviation is not profitable. This establishes the optimality of the proposal rule.

Since this exhausts all possible subgames, the proposed strategy for player 2 is a best response to player 1's strategy. Therefore, the strategies stated in the proposition indeed constitute a subgame perfect equilibrium of the bargaining game. Note that any offer $x_1 > \underline{x}_1$ will be rejected by player 2, and therefore there exists a continuum of SPE, in which player 1 makes different nonserious offers. All these SPE are payoff-equivalent. Q.E.D.

Bibliography

Art, Robert J. 1980. "To What Ends Military Power." *International Security* 4(Spring):4–35.

Bolt, Wilko. 1995. "Striking for a Bargain Between Two Completely Informed Agents: Comment." *The American Economic Review* 85(December):1344–47.

Brooks, Lester. 1968. *Behind Japan's Surrender: The Secret Struggle That Ended an Empire.* New York: McGraw-Hill.

Busch, Lutz-Alexander and Quan Wen. 1995. "Perfect Equilibria in a Negotiation Model." *Econometrica* 63(May):545–65.

Butow, Robert J. C. 1954. *Japan's Decision to Surrender.* Stanford: Stanford University Press.

Clausewitz, Carl von. 1832. *On War.* Princeton: Princeton University Press. Edited and translated by Michael Howard and Peter Paret, 1984.

Coase, Ronald. 1972. "Durability and Monopoly." *Journal of Law and Economics* 15(April):143–49.

Dunnigan, James F. and William Martel. 1987. *How To Stop a War: The Lessons of Two Hundred Years of War and Peace.* New York: Doubleday.

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

Fernandez, Raquel and Jacob Glazer. 1991. "Striking for a Bargain Between Two Completely Informed Agents." *The American Economic Review* 81(March):240–52.

Filson, Darren and Suzanne Werner. 2001. "Bargaining and Fighting: The Impact of Regime Type on War Onset, Duration, and Outcomes." Paper presented at the Political Economy of Conflict Conference, Yale University, March 23-24, 2001.

Fudenberg, Drew and Eric Maskin. 1986. "The Folk Theorem in Repeated Games with Discounting Or With Incomplete Information." *Econometrica* 54(May):533–54.

Fuller, J. F. C. 1961. *The Conduct of War, 1789-1961.* New Brunswick: Rutgers University Press.

Giangreco, D. M. 1997. "Casualty Projections for the U.S. Invasions of Japan, 1945-1946: Planning and Policy Implications." *Journal of Military History* 61(July):521–82.

Goemans, Hein E. 2000. *War and Punishment: The Causes of War Termination and the First World War.* Princeton: Princeton University Press.

Haller, Hans H. and Steinar Holden. 1990. "A Letter to the Editor on Wage Bargaining." *Journal of Economic Theory* 52:232–36.

Iklé, Fred Charles. 1971. *Every War Must End.* New York: Columbia University Press.

Karnow, Stanley. 1983. *Vietnam: A History.* New York: Penguin Books.

Kecskemeti, Paul. 1958. *Strategic Surrender: The Politics of Victory and Defeat.* Stanford: Stanford University Press.

Keegan, John. 1999. *The First World War.* New York: Knopf.

Kegely, Jr., Charles W. and Gregory A. Raymond. 1999. *How Nations Make Peace.* New York: St. Martin's Press.

Kennedy, Paul. 1983. *Strategy and Diplomacy, 1870-1945.* London: George Allen & Unwin.

Kim, Dong-Won. 2001. "Bargaining, Conflict, and War." Mimeo, Texas Tech University.

MacEachin, Douglas J. 1998. *The Final Months of the War With Japan: Signals Intelligence, U.S. Invasion Planning, and the A-Bomb Decision.* Washington, D.C.: Central Intelligence Agency.

May, Ernest R. 2000. *Strange Victory: Hitler's Conquest of France.* New York: Hill & Wang Publishers.

Miles Jr., Rufus E. 1985. "Hiroshima: The Strange Myth of Half a Million American Lives Saved." *International Security* 10(Fall):121–40.

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

Pape, Robert A. 1996. *Bombing to Win: Air Power and Coercion in War.* Ithaca: Cornell University Press.

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

Powell, Robert. 1997. "Bargaining in the Shadow of Shifting Power." Paper presented at the annual meeting of the American Political Science Association, Washington, D.C., August 28-31.

Powell, Robert. 2001. "Bargaining While Fighting." Mimeo, University of California, Berkeley. October 7, 2001 version.

Rosen, Steven. 1972. War Power and the Willingness to Suffer. In *Peace, War, and Numbers*, ed. Bruce M. Russett. Beverly Hills: Sage Publications pp. 167–83.

Rubinstein, Ariel. 1982. “Perfect Equilibrium in a Bargaining Model.” *Econometrica* 50:97–110.

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

Selten, Reinhard. 1975. “Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games.” *International Journal of Game Theory* 4(1):25–55.

Shaked, Avner and John Sutton. 1984. “Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model.” *Econometrica* 52(November):1351–64.

Sigal, Leon V. 1988. *Fighting to a Finish: The Politics of War Termination in the United States and Japan, 1945.* Ithaca: Cornell University Press.

Slantchev, Branislav L. 2002. “The Principle of Irreversibility in a Stochastic Model of Wartime Negotiations.” Mimeo, University of Rochester.

Smith, Alastair and Alan C. Stam, III. 2001. “Bargaining Through Conflict.” Paper presented at the Annual Meeting of the Public Choice Society, March 8-10, San Antonio, TX.

Smith, James D. D. 1995. *Stopping Wars: Defining the Obstacles to Cease-fire.* Boulder: Westview Press.

Stam, III, Alan C. 1999. *Win, Lose, or Draw: Domestic Politics and the Crucible of War.* Ann Arbor: The University of Michigan Press.

Strong, Robert A. 1992. *Decisions and Dilemmas: Case Studies in Presidential Foreign Policy Making.* Englewood Cliffs, NJ: Prentice Hall.

Wagner, R. Harrison. 1994. “Peace, War, and the Balance of Power.” *The American Political Science Review* 88(September):593–607.

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

Walter, Barbara F. 2001. “Explaining the Apparent Indivisibility of Territory.” Paper presented at the Political Economy of Conflict Conference, Yale University, March 23-24, 2001.