

# UC San Diego

## UC San Diego Previously Published Works

### Title

The Principle of Convergence in Wartime Negotiations

### Permalink

<https://escholarship.org/uc/item/87s3t16x>

### Journal

American Political Science Review, 47(4)

### ISSN

0003-0554

### Author

Slantchev, Branislav L

### Publication Date

2003-12-01

Peer reviewed

# The Principle of Convergence in Wartime Negotiations

BRANISLAV L. SLANTCHEV *University of California, San Diego*

*If war results from disagreement about relative strength, then it ends when opponents learn enough about each other. Learning occurs when information is revealed by strategically manipulable negotiation behavior and nonmanipulable battlefield outcomes. I present a model of simultaneous bargaining and fighting where both players can make offers and asymmetric information exists about the distribution of power. In the Markov perfect sequential equilibrium, making and rejecting offers has informational value that outweighs the one provided by the battlefield. However, states use both sources of information to learn and settle before military victory. The Principle of Convergence posits that warfare ceases to be useful when it loses its informational content and that belief in defeat (victory) is not necessary to terminate (initiate) hostilities. Thus, the standard puzzle in international relations that seeks to account for prewar optimism on both sides may not be that relevant.*

**W**hy do wars end? War is an instrument of policy and its goal is to achieve peace through violent diplomacy.<sup>1</sup> Attaining military victory is central for policy planners and military commanders and is explicitly stated as a goal in U.S. strategic doctrine (C. Powell 1992). However, total disarmament and complete overthrow of the opponent are quite rare (Pillar 1983). Wars most often terminate in negotiated settlements short of military collapse.

So why do opponents agree to terminate hostilities? War can be viewed as an organized coercive process through which opponents attempt to persuade one another to concede whatever is demanded by the other (Schelling 1960). Since this process is extremely costly, both sides have strong incentives to end it as soon as possible while conceding as little as they can. How long they hold out depends not only on their capabilities but also on the military situation and on what they expect to happen in the future.

Expectations are, in fact, central to explanations of rational war termination. Blainey (1988, 54) argued that the only surprise in war is that at least one side that expected to win actually lost, concluding that mutual optimism about military victory was necessary to start a war. The leading explanation of war as bargaining failure demonstrates how such optimism is possible when both players are rational (Fearon 1995; R. Powell 1996).

The emphasis on expectations about the *military* outcome, however, is misleading. War is coercive bargaining and ends because opponents succeed in coordinating their expectations about what each is prepared to concede.

---

Branislav L. Slantchev is Assistant Professor of Political Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0521 (slantchev@ucsd.edu).

I thank Randall Stone, Robert Powell, Curt Signorino, John Duggan, Erik Gartzke, and Jeffrey Lax for comments. I am also especially grateful to two anonymous reviewers for very helpful advice on how to better focus the material. Previous versions were presented at the 2002 Annual Meetings of the American Political Science Association and the Peace Science Society.

<sup>1</sup> This is derived from the classic dictum by von Clausewitz ([1832] 1984)—but is not equivalent, he likened war to a duel—and is frequently emphasized by some military thinkers like Fuller (1961). On war as pure coercion, see Calahan 1944, and for emphasis on the bargaining component, see Keckskemeti 1970 and Schelling 1966.

How do they coordinate these expectations? Although fighting can result in complete military victory, its more important function is coercive: to convince the opponent to accept a settlement. This happens after opponents learn enough about their prospects in war to decide that continuation is unprofitable. Warfare is transmission of information about these prospects. The Principle of Convergence states that once expectations converge sufficiently, war loses its informational content, and hostilities can terminate with a negotiated settlement.

I derive this principle from a formal model that treats war as a costly instrument of policy and allows unlimited diplomatic exchange. I analyze how rational players learn from two sources of information: negotiating behavior and the battlefield. These sources are subject to different degrees of strategic manipulability, and may possibly provide contradictory information.

Wagner (2000) emphasizes that fighting provides a means of revealing information that is “not available in the standard bargaining models” and is in some ways superior to inferring private information from negotiation behavior. But he may have overstated his case because (a) the “fog of war” makes interpretation of any event notoriously difficult (Iklé 1971), and (b) strategic bargaining behavior remains an important source of information (Iklé 1964). The model takes all these issues into account and speaks directly to the relevance of negotiations in warfare.<sup>2</sup>

States are uncertain about the probability of winning battles. As the war progresses, they observe the outcomes on the battlefield in addition to the bargaining history and evaluate their prospects for the future. Provided that states value the future sufficiently, they delay agreement in order to accrue enough information about their prospects and avoid settling prematurely on

---

<sup>2</sup> The principle is robust to alternate formal specifications and can be derived from other work, as shown in the discussion section where I relate the findings to models by R. Powell (2001), A. Smith and Stam (2001), and Filson and Werner (2002), all of which emphasize the informational approach to war. However, the analysis of wartime learning from divergent sources in a dynamically rich environment of simultaneous bargaining and fighting is new. How contradictory evidence is interpreted is crucial for the analysis of strategic assessment in war (Gartner 1997) but has been neglected. Indeed, many models are incapable of addressing it.

worse terms. This type of uncertainty can only be resolved *ex post*: The risk of war always exists and cannot be eliminated through prewar bargaining. The timing of the settlement depends on the rate at which information accrues but its terms also depend on the military position at the time agreement is struck, which explains the common last-minute jockeying for advantage prior to an armistice.

Fighting does reduce uncertainty, but the battlefield is a noisy source of information and not the only one. The strategic behavior of states at the bargaining table can be very revealing. Because the model constrains players to make counteroffers upon rejecting offers, information can accrue rather quickly and more precisely from their negotiating behavior. Since readiness to talk can be so revealing, it may provide a good rationale for delaying explicit diplomacy until after an armistice.

Military developments may provide information that contradicts the explicit bargaining behavior of the opponent. For example, making an unreasonable demand signals strength but defeat in battle reveals weakness. I explain how opponents interpret information coming from a *manipulable* source (negotiating table) and reconcile it with one from a *nonmanipulable* source (battlefield). Any theory of war termination has to account for the process of convergence of expectations that ends fighting and, so, must explain how opponents interpret new information.

The substantive findings that emerge from this analysis have theoretical implications and empirical significance. In equilibrium, as in the historical record, total military victory is a rare occurrence. A common explanation of why wars end is that both sides agree as to who the eventual winner will be. This is not necessary: War can end when both sides agree on the relative likelihood of various outcomes. This is a much weaker requirement and explains cases where wars were settled before it became clear that one side would emerge victorious for sure.

Thus, *belief in eventual military defeat is not necessary for war termination*. This implies that *belief in eventual victory is not necessary for war initiation*. The standard puzzle about how both sides can be optimistic about their chances of military victory is therefore less relevant.

Why do weak states sometimes attack stronger ones even when it is clear that they have no chance of victory? One argument forces us to assume irrational expectations or resolve. This is not necessary: As long as the stronger state believes that its opponent is a little stronger than it actually is, the weaker state can benefit from fighting a short war and settling. Or as von Clausewitz ([1832] 1984, 92) put it, if the weaker states succeeds in giving the stronger one “doubts about the future,” it can hope to profitably exploit its fear of prolonged conflict.

## RELATED WORK

Three models explore the interdependence of bargaining and fighting in different environments. Powell's

(2001) formalization is based on the Rubinstein (1982) bargaining model with inside options (Muthoo 1999) and one-sided incomplete information. Only the uninformed player can make offers, and every time an offer is rejected, players can fight. Fighting may result in the collapse of either state, and the probability of collapse is exogenously specified. The equilibrium exhibits the “skimming property,” where the uninformed state screens out its opponent by making progressively larger offers.

Filson and Werner (2002) offer a richer battlefield environment where fighting can shift the relative military advantage of each state because every time a battle is fought, states expend resources of which they have limited amounts. The bargaining protocol is also one-sided and only the uninformed state can make offers. They analyze a two-period special case by assuming severe resource constraints and find a logic analogous to the skimming property.

A. Smith and Stam (2001) embed the one-sided bargaining protocol in a random walk model of warfare based on work by A. Smith (1998). In each period where states disagree, a costly fight ensues and improves one state's chances for military victory. With time, disagreement over the probability of eventual victory disappears as both sides update their beliefs using the information revealed by battlefield outcomes. Expectations converge on stalemate and war ends with a settlement. However, since players are nonrational in their model, it is difficult to relate their results to others.

My model is closely related to this formalization but makes several crucial modifications and extensions. First, the bargaining protocol is richer than any of the available models: Both sides can make unlimited numbers of offers, allowing for screening and signaling behavior. Second, players discount the future and suffer per-period costs instead of only paying fixed costs. This is not a trivial improvement because the results hinge upon how much players value the future. Third, players use all the information available, not only the battlefield. The uninformed state learns about its opponent by observing both its strategic behavior at the bargaining table and the nonmanipulable battlefield performance.

## THE MODEL

Two players,  $i \in \{1, 2\}$ , bargain over a two-way partition of a flow of benefits with size  $\pi$ . An agreement is a pair  $(x, y)$ , where  $x$  is player 1's share, and  $y$  is player 2's share. The set of possible pairs is  $\mathcal{X} = \{(x, y) \in \mathcal{R}^2 : x + y = \pi\}$  and  $0 \leq x, y \leq \pi$ . Players have strictly opposed preferences and each is concerned only with the share of benefits it obtains from the agreement. Because a share  $x$  identifies a distribution uniquely, let  $x$  be equivalent to the pair  $(x, \pi - x)$ , and  $y$  be equivalent to the pair  $(\pi - y, y)$ . The status quo distribution of benefits is  $(s_1, s_2)$  with  $s_1 + s_2 = \pi$ .

The two players bargain according to the alternating-offers protocol (Rubinstein 1982). Players have a common discount factor  $\delta \in (0, 1)$ , and act in discrete time

with a potentially infinite horizon and periods indexed by  $t(t = 0, 1, 2, \dots)$ . In even-numbered periods, player 1 proposes a division  $x \in \mathcal{X}$  to player 2. If player 2 accepts that proposal, an agreement is reached, and the game ends with players receiving their shares in  $(x, \pi - x)$ .<sup>3</sup> If player 2 rejects the proposal, then players fight a costly engagement, which may improve the relative military position of a player, and the period ends. Player 2 makes a counteroffer  $y \in \mathcal{X}$  in the next period. If player 1 accepts, the game ends and players receive their payoffs from the agreement  $(\pi - y, y)$ ; if player 1 rejects, they fight another military engagement. The game continues until an agreement is struck or until one of the players is decisively defeated. If a player decisively defeats the other, then it obtains the entire flow of benefits  $\pi$ . Each military engagement is costly, and states suffer a constant per-period loss of utility, reducing their instantaneous per-period wartime payoffs to  $b_i < s_i$  (and so  $b_1 + b_2 < \pi$ ).

War is modeled as a stochastic process of attrition. It is a homogeneous Markov chain with two absorbing states: victory and loss.<sup>4</sup> The *current military position* of a player at time  $t$  captures the player's relative overall success from all engagements that have occurred up to time  $t$ . Let  $N \geq 2$  denote the finite number of military objectives and let  $k$  be the number of objectives achieved by player 1. The set of possible states is  $K = \{0, 1, \dots, N\}$ . At time  $t$ , player 1's current military position,  $k_t \in K$ , is the difference between the total number of its victories and that of its losses in battles that have occurred in periods  $(0, 1, \dots, t - 1)$ . The state variable  $k_t$  is an indicator of relative military advantage at time  $t$  and summarizes the whole history of the war up to that point in time;  $k_0$  is the position at the outset of war.<sup>5</sup>

One battle over one objective occurs in each period. Player 1 wins the fight with probability  $p$  and loses with probability  $1 - p$ .<sup>6</sup> If player 1 wins the battle at time  $t$ , then  $k_{t+1} = k_t + 1$ , and if it loses, then  $k_{t+1} = k_t - 1$ . If

<sup>3</sup> This is common to these types of models and assumes away possible enforcement problems. In particular, if a player accepts a share that leaves it worse off and military capability is derived from the share, then the other player is facing a commitment problem because it will be unable to credibly promise not to extract further concessions in the future. This can only make settlement more difficult today.

<sup>4</sup> See Grimmett and Stirzaker 1992 for random processes.

<sup>5</sup> Until a settlement is reached (or a military victory obtains), there is no explicit relationship between the distribution of (nonmilitary) benefits and the military position. One way to think about  $k_0$  is in terms of the relative strategic circumstances, that is, how much effort would be necessary to expend to achieve military victory while holding the probability of winning individual battles constant. For example, the military position of Germany vis-à-vis Czechoslovakia in 1937 favored Germany less than the post-Munich one following the surrender of the Sudetenland along with its impressive fortifications. In terms of the model,  $k_0(1937) < k_0(1938)$ , making military victory more likely in the latter case even if the probability of winning battles remained the same. In this case, Germany's benefits also increased.

<sup>6</sup> It is possible to relax this assumption in two ways: The probabilities of victory and defeat need not sum to 1; the probabilities of victory and defeat vary according to the current military position. It is not clear *a priori* how probabilities should vary with battlefield success. For example, sometimes early failure mobilizes the will to fight (e.g., Britain after Dunkirk in May 1940) but other times it does not (e.g., France in June 1940).

$k_t = 0$ , player 1 is militarily defeated and the game ends with player 2 imposing the settlement  $(0, \pi)$ . If  $k_t = N$ , player 2 is militarily defeated and player 1 imposes the settlement  $(\pi, 0)$ .<sup>7</sup> Players maximize the time-averaged discounted sum of per-period payoffs,  $(1 - \delta) \sum_{t=0}^{\infty} \delta^t r_t^i$ , where  $r_t^i$  is player  $i$ 's instantaneous payoff in period  $t$  and equals  $b_i$  if players disagree, 0 if player  $i$  loses the war,  $\pi$  if it wins, and  $i$ 's share of benefits if players terminate the war with a settlement.<sup>8</sup>

This model avoids some common pitfalls. Unlike the costly lottery approach, it does not reduce war to a single-shot event and permits analysis of dynamics. Unlike the infinitely repeated game approach, it does not go against the intuition that the process does not last indefinitely, or even a large number of periods (Rubinstein 1991, 918). Instead, this model captures the dynamic nature of the process without either fixing an arbitrary number of periods or allowing it to extend indefinitely, while incorporating the time dependence of each state. Finally, war is completely instrumental: It only serves as advancing players closer to victory or defeat, but players do not benefit from fighting itself; they only care about the political deal. This differs from A. Smith 1998, and the results reflect this.

### COMPLETE INFORMATION

Let  $W_i^k : K \rightarrow \mathcal{R}$  denote player  $i$ 's expected payoff from fighting to the finish starting in state  $k$ . For  $0 < k < N$ , this function is defined recursively as

$$W_0^1 = W_N^2 = 0,$$

$$W_N^1 = W_0^2 = \pi,$$

$$W_k^i = (1 - \delta)b_i + \delta[pW_{k+1}^i + (1 - p)W_{k-1}^i].$$

The functions  $W^i$  are second-order linear recurrence relations and have closed forms. It can be shown that for all  $k$ ,  $W_k^i \in [0, \pi]$  and  $W_k^1 + W_k^2 < \pi$ . That is, there exists no state where players lack incentives to bargain.

The set of Nash equilibria of the negotiation game is very large (Slantchev 2002). The set of equilibrium payoffs consists of all payoffs that are at least as good as fighting to the end from the starting state:  $[W_{k_0}^1, \pi - W_{k_0}^2]$ .

<sup>7</sup> This implies that players only pay costs while fighting continues and war does not permanently shrink the resource base. This can be justified empirically by the Phoenix factor: It does not take states that long to recover from war. Allowing for a longer finite cost-decay period will not alter the results. I conjecture that assuming permanent destruction of resources will make players more reluctant to prolong the learning process.

<sup>8</sup> The formal description of strategies and payoff functions is cumbersome because the payoffs reflect the fact that bargaining may be ended by the exogenous stochastic process. The formalization, along with several computer programs (in C++ and Gauss) for numerical computations are available from the author. Merlo and Wilson (1995) provide a general  $n$ -player infinite-horizon complete information model, in which the identity of the proposer and the size of the pie follow a general Markov process. Their results cannot be used here because in my case the Markov process eventually terminates the bargaining.

Because these equilibria rely on incredible threats to sustain optimal behavior, I require that the strategies be subgame perfect (Selten 1975). In addition, given the structure of warfare, it is natural to restrict attention to a class of strategies where behavior depends on the military position. Strategies that condition behavior on payoff-relevant history are called Markov. Since the only variable that influences future payoffs is the military position, Markov strategies depend only on  $k_t$  to determine optimal behavior at time  $t$ . A subgame perfect equilibrium in Markov strategies is called Markov perfect (MPE).<sup>9</sup> A *stationary* MPE is one where players always make the same state- dependent offers and responses. That is, offers differ by state but are time-invariant. A *no-delay* MPE is one where players' optimal offers are immediately accepted.

**Proposition 1.** *The stochastic bargaining game with complete information has a unique stationary no-delay Markov perfect equilibrium, in which player 1's first state-dependent offer is immediately accepted by player 2, and no fighting occurs.*

For every possible military position, there exists a unique optimal offer a player can make that will be accepted by its opponent in equilibrium. This offer is calibrated to make the other player indifferent between accepting it and delaying for one period in order to have its optimal offer accepted then. To anticipate how this result will be used in the sections that follow, the terms of the offer also depend on  $p$  in the intuitive way: Player 1's optimal proposal is strictly increasing in  $p$ , while player 2's optimal proposal is strictly decreasing.

This result establishes an important result: Modeling warfare as a probabilistic process does not lead to inefficient behavior. It is worth contrasting this result with the one obtained by A. Smith (1998), where both players prefer fighting in some states. The difference stems from the assumption that players derive direct utility from their military position, which is not the case here. When war is instrumental, the stochastic element is not sufficient to produce inefficiency.

### ASYMMETRIC INFORMATION

We now assume that player 1 is uncertain about the distribution of power. While player 2 knows the true value of  $p$ , player 1 is asymmetrically informed and believes that  $p$  may be low ( $p_L$ ), medium ( $p_M$ ), or high ( $p_H$ ), with  $0 < p_L < p_M < p_H < 1$ . That is, player 2 is weak, denoted  $2_w$ , when  $p = p_H$ ; moderately strong, denoted  $2_m$ , when  $p = p_M$ ; and very strong, denoted  $2_s$ , when  $p = p_L$ . Let  $T = \{w, m, s\}$  denote the set of type indicators. Player 1 initially believes that player 2 is weak with probability  $q^w > 0$ , strong with probability

$q^s \in (0, 1 - q^w)$ , and moderately strong with probability  $q^m = 1 - q^w - q^s > 0$ .<sup>10</sup>

The set of Bayesian equilibria (Harsanyi 1968) is very large, as is the range of payoffs that can be supported in equilibrium. Like Nash equilibrium, this solution concept ignores the dynamic structure of the game. Bayesian equilibria may rely on noncredible threats to sustain optimal behavior because strategies are only required to be best responses at the beginning of the game, and players do not learn from history. An equilibrium refinement that overcomes these shortcomings is necessary. Learning in the stochastic negotiations model is complicated because there exist two sources of information: (i) the battlefield, which is nonmanipulable, and (ii) the strategic behavior of the opponent, which is highly manipulable. The uninformed player must take both into account when updating its beliefs about the distribution of power.

### Involuntary Revelation of Information

Suppose that all types of player 2 reject some proposal. Following such rejection, the support of player 1's beliefs will remain the same, that is, if it believed that it might be facing each type with positive probability, it will continue to do so. However, depending on the outcome of the battle, player 1 will update the probability associated with each type. Intuitively, winning a battle should make player 1 more optimistic about the chance of facing a weak opponent, while losing a battle should make player 1 more pessimistic.

To formalize this intuition, let  $I_t$  be an indicator that equals 1 if player 1 wins the battle at time  $t$  and 0 otherwise. Suppose that all types of player 2 reject the initial offer. Using Bayes' Rule, the posterior belief is then, for  $\tau \in T$ ,

$$\Pr(2_\tau | I_0) = \frac{\Pr(2_\tau)\Pr(I_0 | 2_\tau)}{\sum_{i \in T} \Pr(2_i)\Pr(I_0 | 2_i)}$$

Since the number of victories,  $v$ , in  $n$  battles is a binomially distributed random variable given the probability of winning, we have

$$\Pr(v, n | p) = \binom{n}{v} p^v (1 - p)^{n-v},$$

and thus, the posterior  $q_1^s \equiv \Pr(2_s | I_0)$  is

$$\frac{q^s p_L^{I_0} (1 - p_L)^{1-I_0}}{q^w p_H^{I_0} (1 - p_H)^{1-I_0} + q^m p_M^{I_0} (1 - p_M)^{1-I_0} + q^s p_L^{I_0} (1 - p_L)^{1-I_0}}$$

Algebraic manipulation shows that extending the argument to  $v$  victories after  $n$  battles allows the posterior belief  $q_n^s(v)$ , where  $v = I_0 + I_1 + \dots + I_{n-1}$ , to be expressed directly in terms of the initial belief and the

<sup>9</sup> A pair of strategies forms MPE if, and only if, each player's strategy maximizes its intertemporal payoff at any time  $t$ , given  $k_t$  and assuming that henceforth each player conforms to its strategy (Fudenberg and Tirole 1991, ch. 13).

<sup>10</sup> This stylization of incomplete information reflects closely the discussions in the informal literature that stress disagreements about the relative strength of opponents (Blainey 1988). Usually, this is taken to mean that countries disagree about the eventual military outcome of war (Keckskemeti 1958; Pillar 1983).

number of victories. The posteriors for the other types are defined analogously.

Because player 1 updates its beliefs through a non-strategic mechanism that is beyond the control of its opponent, the possibilities for strategic dissimulation by player 2 are limited. Even if weak opponents try to imitate the strategic behavior of their stronger counterparts, they cannot do so for long because their poor performance on the battlefield will reveal information that will gradually convince player 1 that it is facing a weak opponent, not a strong one. On the other hand, the battlefield is a noisy source, and in the absence of strategic behavior, player 1 will never be able to eliminate the possibility that it is facing a particular type of opponent. That is, although the probabilities associated with some type can become arbitrarily close to zero, they never equal it; the support consists of all three types. Strategic bargaining, however, changes that.

### Properties of Sequential Equilibria

The usual solution concept for dynamic games of incomplete information is Kreps and Wilson's (1982) sequential equilibrium, which elevates beliefs to the level of strategies and specifies a method for updating these beliefs. This solution concept requires that strategies are sequentially rational (that is, they are consistent with the beliefs), and beliefs are derived from strategies and updated via Bayes' Rule whenever possible. The vagueness of the last requirement stems from existence of zero-probability events (events that should never occur in equilibrium), where the rule cannot be applied.

Sequential equilibrium assigns explicit but unrestricted by theory conjectures that players use to update their beliefs following zero probability events. These conjectures are crucial because the equilibrium outcomes are quite sensitive to their specification (Rubinstein 1985b). Following Grossman and Perry (1986a), who call their solution concept *perfect sequential equilibrium*, I postulate the following credible conjectures for player 1: If there exist some types of player 2, for which a deviation would be profitable, player 1 updates to believe that player 2 is among these types. This requirement eliminates sequential equilibria in which the credibility of behavior is established through incredible beliefs. Also, as in Rubinstein (1985a), I require that if player 1 becomes convinced that its opponent is of some type with probability 1, then it never revises this belief. I shall therefore look for perfect sequential equilibria in Markov strategies (MPSE), where player 1 updates its beliefs credibly based on possible optimality of deviations of its opponent, and where beliefs and offers depend on the outcome of fighting during disagreement periods.

Once all information is revealed, the game becomes equivalent to the one analyzed in the previous section, with starting state set at the military position at the time that no more private information remains. This means that we can use the solution to the complete information model for the subgames of the asymmetric information model that follow full revelation.

Recall that by Proposition 1, there exists a unique vector of state-dependent proposals that make players indifferent between accepting the proposal and delaying agreement by one period in order to make a proposal themselves. Let  $(V_s^1(k_0), V_s^2(k_0))$  denote the unique complete information MPE offers from Proposition 1 when player 1's opponent is of type  $2_s$ .  $V_s^1(k)$  is player 1's complete information MPE payoff when player 1 is the proposer, the current state is  $k$ , and the opponent's type is  $2_s$ . Equivalently,  $V_s^2(k)$  is the payoff to player  $2_s$  when it is the proposer in state  $k$ . Note that  $V^1(k)$  is strictly increasing in  $p$ , which implies that  $V_s^1(k)$  is strictly larger than the analogous optimal offers made if player 2's type is moderately strong or weak. No sequential equilibrium outcome can be better (worse) for player 1 than the MPE outcome in the complete information game with  $2_w(2_s)$ .

From Proposition 1, the best that player 1 can obtain when its opponent is  $2_s$  in the complete information MPE is  $V_s^1(k)$ . Since it does not pay to delay such agreement, it follows that if after some history player 1 believes with probability 1 that its opponent is  $2_s$ , then it will always offer  $V_s^1(k)$ , which player  $2_s$  accepts. As the following lemma shows, this implies that all types of player 2 will accept such an offer.

**Lemma 1.** *If  $2_s$  accepts some offer  $x$  in a sequential equilibrium, then  $2_w$  and  $2_m$  accept that offer also.*

*Proof.* Suppose that in equilibrium  $2_s$  accepts  $x$  but one or both of the other types reject it. Following that rejection player 1 concludes that the probability of facing  $2_s$  is 0 and will, therefore, offer player 2 at most  $\pi - V_m^1(k) < \pi - V_s^1(k) \leq \pi - x$ . Both  $2_m$  and  $2_w$  are better off by accepting  $x$  instead.  $\square$

**Lemma 2.** *If in a sequential equilibrium player 1 offers some  $x$  and player 2 rejects it, then player 2 either makes a unique acceptable offer or counters with an unacceptable one.*

*Proof.* Suppose that some type of player 2 counters with  $y_1$  and another counters with  $y_2 \neq y_1$ . If player 1 accepts both offers, then the type which made the higher offer can profitably deviate by offering the other.  $\square$

The set of sequential equilibria is large, and many of these equilibria can be supported with optimistic conjectures, in which player 1 threatens player 2 that if it deviates from the proposed strategy, 1 will conclude that its type is  $2_w$ . However, these equilibria are eliminated by the perfectness restriction on conjectures used in the construction of the unique MPSE in the next section.

### Markov Perfect Sequential Equilibrium

To construct the MPSE, I first solve for the equilibrium in which player 1 has good information about its opponent and then solve for equilibrium in games where its information is slightly worse. Since information only improves over time, we shall use the results from the

first step in the solution for the next one.<sup>11</sup> Since there are only three types of opponents that player 1 might be facing, and because it does not pay to delay once the type is known, I shall look for equilibria with the most rapid disclosure of information. To solve for this type of equilibrium, we induct backward on both beliefs and strategies, starting with the last period, where, if players adhere to their equilibrium strategies, player 1 knows with probability 1 that it is facing  $2_s$ . This belief holds regardless of the intermediate losses and victories because in such equilibrium the weaker types have screened themselves out earlier. The construction of this equilibrium is quite involved and is relegated to the Appendix.

**Proposition 2.** *When players are sufficiently patient, there exists a generically unique MPSE with the following structure. In period  $t = 0$ , player 1 makes an initial offer that only  $2_w$  accepts; in period  $t = 1$ , player  $2_m$  makes an offer which player 1 accepts, and player  $2_s$  makes a non-serious offer, which player 1 rejects; and in period  $t = 2$ , player 1 makes an offer which all types of player 2 accept.*

Finding the equilibrium involves searching for a strategy for player 1 that can profitably induce the three types of player 2 to separate in equilibrium, as well as a strategy for the strong type that induces it to signal its strength by making a nonserious offer. The separating equilibrium exists only for sufficiently high discount factors. When  $\delta$  is lower, then semiseparating and pooling equilibria appear as well.

Conditioning behavior on the battlefield outcomes presents players with a complicated mix of incentives. Both players prefer to settle as early as possible, but player 1 does not want to offer more than strictly necessary to induce its opponent to accept. This requires that it makes some offers that might be rejected, resulting in inefficiency in the process.

Equilibrium play is strongly conditioned by the military position at the time offers are made, accepted, or rejected. It is preferable to win a battle prior to sitting at the negotiating table. Victories make player 1 more optimistic about its chances even though at  $t = 1$  it knows for sure that it is facing either a moderately strong or a strong opponent. Consequently, if it wins the fight, it demands more than it does if it loses it.

Players can deal with contradictory information. By rejecting player 1's initial offer, its opponent signals its strength, but if it then loses the battle, player 1 becomes more optimistic again. The value of the discount factor necessary to induce player  $2_s$  to signal its strength in period 1 depends on player 1's belief about the likelihood of facing the strong type, which in turn depends on the battlefield outcome. This belief is more optimistic if player 1 wins the battle and so the discount factor has to be smaller than the one required to sustain separation after defeat.

There exist discount factors such that  $2_s$  pools with  $2_m$  after player 1 is defeated but makes a nonserious offer

after player 1 wins. A strong state may refuse to settle after a victory for its opponent while still preferring to settle after its defeat. The existence of these equilibria is a good reason for the often observed behavior of attempting to gain a military advantage immediately prior to making a proposal and of strong states refusing to settle if the opponent has gained a military advantage.

The level of uncertainty about the type of opponent can be expressed as a combination of each type's probability of winning and beliefs about the three types. For simplicity, define the level of uncertainty as the variance of the probabilities of winning. When the three types are more or less equally likely to prevail in battle, then player 1 is quite certain about the strength of its opponent. Conversely, when the three probabilities are very different, player 1 can be said to be quite uncertain about the type of opponent it is facing.

With little uncertainty, the discount factor required to sustain separation increases quite dramatically. This implies that we are more likely to see partially or fully pooling equilibria, where players agree either immediately or, at most, with one period delay, shortening the expected duration of war. Conversely, when there is a lot of uncertainty, separation becomes easy, and wars should be longer. This result follows from the requirement that stronger types must have incentives to signal, and that player 1 must have incentives to screen. When the types are comparatively equal in strength, the necessary gains are negligible, and so the incentive to delay disappears. More uncertainty about the distribution of power results in longer duration of conflict.

## DISCUSSION

### Warfare as Information Transmission

Stating that war is the pursuit of political objectives, although better than regarding it as the "untrammelled manifestation of violence," is still imprecise. In its most common form, this approach states that war is a way to secure political objectives by force. As Hobbs (1979, 46) notes, "The war aim of strategy is to clinch a political argument by force instead of words." This is further elaborated by Wagner (1994, 603), who claims that "the primary function of force in bargaining is to improve one's bargaining position by increasing the costs of disagreement for one's adversary."

The analysis here, as in the three related models, departs from this approach. The ability to increase the costs of disagreement may entail improvement of one's position but does not explain why fighting occurs. With complete information, states still settle immediately. On the other hand, the presence of asymmetric information does not simply generate risk of war, as in Fearon 1995. In this model war does not arise as a result of breakdown of bargaining. Instead, war occurs when players are sufficiently patient and prefer to engage in strategic screening and signaling, that is, in transmission of information. War is bargaining, and bargaining is transmission of information.

<sup>11</sup> The analysis generally follows the method used by Fudenberg, Levine, and Tirole (1985) and Grossman and Perry (1986b).

Warfare only has informational content while uncertainty about the opponent persists. Fighting becomes irrelevant (in the informational sense) once players learn enough about their opponents.<sup>12</sup> The uncertainty reduction interpretation has become quite prominent in recent formal work, whose common theme is that fighting provides information about the relative strength of opponents and allows them to arrive at more congruent estimates of their chances of success, thus enabling them to conclude bargains (Wagner 2000). As Reiter (2003) points out, this contradicts the earlier interpretation which treated fighting as detrimental to bargaining because of the expectation that the winning side inevitably expands its demands (Wittman 1979).

R. Powell's model cannot address this because it is impossible for the uninformed player to become optimistic with time, which is a consequence of the static distribution of power (probability of collapse is exogenous and fixed). My model, along with the other two, shows that these trends are not mutually exclusive. The current military situation influences the proposals and responses. It is preferable to have won the last round of fighting prior to concluding an agreement.<sup>13</sup> However, reduction of uncertainty has enormous implications for equilibrium behavior. The model demonstrates quite clearly that both effects are at work at the same time. Although the eventual bargain does depend quite significantly on the current military position, the important information transmission happens through the strategic behavior of opponents which provides more precision than the crude fighting mechanism. Combat itself may reveal less than the willingness of opponents to engage in it.

This is a strong qualification of the existing arguments because it shifts the emphasis of war termination back to the political realm. That negotiating behavior has such dramatic implications is recognized by many diplomats and practitioners, for example, Nicolson (1954) and Iklé (1964), but has been neglected in the study of war termination.<sup>14</sup> Focusing squarely on military developments as the most important source of information, as Wagner (2000) does, may not be helpful because of the "fog of war" that makes for wildly divergent estimates of battlefield performance (Gartner 1997; Iklé 1971). Diplomatic exchange remains an important tool to influence expectations of the opponent, which is probably one reason governments are reluctant to engage in it while the war continues.

Unfortunately, sometimes force is necessary to convey sufficient information to induce a revision of beliefs. This is not because force itself convinces, but because

the willingness to use it, and suffer its costs, distinguishes between different types of opponents. It is not clear that there is a way to avoid this because weak players always have incentives to dissemble as being strong and only a costly delay may persuade their opponents otherwise.

### The Impact of Uncertainty

It is relatively straightforward conceptually (but a bit demanding technically) to extend the analysis to any finite number of types. This will result in additional delay in the separating equilibrium because it will take more fighting to distinguish among them.

Except for A. Smith and Stam (2001), who ignore strategic behavior altogether, all other models exhibit the screening property. However, since neither R. Powell nor Filson and Werner allow the informed player to make counteroffers, it is impossible to analyze signaling behavior.<sup>15</sup> This leads to questions about this player's willingness to delay agreement if a richer communication tool is at its disposal. However, as the results show, the ability to send more complicated messages only sometimes translates into ability to terminate the war. Although a moderately strong state will make an acceptable counteroffer, a strong one will still make an unacceptable one to demonstrate its strength.

Uncertainty benefits the weak and hurts the strong. The asymmetry of information is always detrimental to the uninformed party, who tries to screen out its opponent when under complete information it would settle immediately. A strong opponent also does worse because it has to engage in costly delays to signal its strength and separate itself from weak types who have incentives to claim they are strong but who cannot afford a prolonged fight. For both, the ex ante payoff under uncertainty is lower than the payoff under complete information.

Weaker opponents, on the other hand, do better, sometimes significantly so, than they would have done under complete information. This is because the offers player 1 must make are conditioned on its estimate that it might have to make more concessions should player 2 turn out to be strong. Because delay is costly, player 1 will not engage in it forever and, so, will make concessions that balance the need to minimize the current offer and the possibility of having it rejected for a lower settlement. Invariably, this exceeds what it would have offered had it known for sure that its opponent was weak.

### The Principle of Convergence

The model shows that a change in war aims (expectations) is necessary for war to end (Fox 1970, 7). War is not so much about twisting arms, but about influencing expectations. How can players systematically and

<sup>12</sup> From the Myerson and Satterthwaite (1983) result, we know that uncertainty about whether gains from a negotiated outcome exist necessarily prevents full ex-post efficiency. This is not the reason here, as gains from settlement always exist, regardless of how strong the opponent is.

<sup>13</sup> See Forster 1941 and Goemans 2000 for examples of how both sides modified demands depending on recent success or failure.

<sup>14</sup> In this regard, Chapter 11 in Iklé's (1964) book is especially informative, particularly the section on the function of proposals, which states that "your proposals must change your opponent's expectations" (194).

<sup>15</sup> Some signaling does occur in their models also because the rejection of an offer conveys information, albeit in a rather crude way. The ability to make explicit offers permits a much wider range for strategic behavior.



strategically influence each others' beliefs? The tension between the desire to find a settlement and the desire to give up as little as possible produces inefficiency as opponents learn more about each other.

Learning occurs through two channels: a non-strategic, nonmanipulable, and involuntary one—the battlefield—and a strategic, manipulable, and voluntary one—the negotiation table. The first is imprecise and noisy, and although players can infer something about the distribution of power, it is not sufficient to ensure convergence in beliefs. The strategic channel is more useful because offers, counteroffers, and rejections are all rational decisions that reveal information about the privately known parameter. However, as noted before, some of the strength of its impact is due to the requirement that players always make counteroffers.

Since negotiations can be so revealing, it is perhaps not surprising that empirically warring parties try to avoid initiating talks until after an armistice.<sup>16</sup> In the cases where explicit bargaining was contemporaneous with fighting, the offers and counteroffers were heavily dependent on dramatic developments of the battlefield but as war progressed, convergence did occur despite lack of decisive engagements (Pillar 1983).

The *Principle of Convergence* posits that warfare ceases to be useful when it loses its informational content, which occurs when strategic and involuntary revelations make beliefs “irreversible” in the sense that both sides can agree on the relative likelihoods of different outcomes.

While variants of the principle can be derived from the other models, it exhibits slightly different dynamics. In A. Smith and Stam 2001, players only take into account battlefield outcomes and ignore the manipulable source. Their model cannot address the issue of wartime negotiations at all. Filson and Werner solve their model for an artificially restricted case that limits possible fighting to only two battles, and the one-sided bargaining protocol (shared with R. Powell's model) ignores the possibility that an informed player may signal strategically. It is, however, important to note that this principle appears quite robust to different specifications, which is usually not the case because bargaining models are quite sensitive to the precise protocol.

J. Smith (1995) states that the main reason for war continuation is belief in victory, and so one of the most important requirements for a cease-fire agreement is a clear military trend. Similar arguments abound, and all share the conclusion that a convergence of expectations about the military attrition trend are a necessary condition for termination of armed conflict (Calahan 1944; Foster and Brewer 1976; Kecskemeti 1970). The principle of convergence is much weaker for it only requires that opponents agree on the relative likelihoods

of different outcomes, not on who the eventual winner is going to be.

### A Substitute for Victory

Traditionally, the purpose of the military instrument has been to secure victory on the battlefield. As General Douglas MacArthur claimed, “War's very objective is victory—not prolonged indecision. In war there is no substitute for victory.” The thinking is reflected in American strategic doctrines and permeates theories of war that invariably focus on its outcome expressed in strictly military terms: victory, defeat, or stalemate.

The Principle of Convergence suggest that this view is flawed because a crucial result is that *belief in defeat (or victory) is not a necessary condition to agree to terminate a war*, which lends support to von Clausewitz's ([1832] 1984, 91–92) claim that “not every war need be fought until one side collapses. . . if one side cannot completely disarm the other, the desire for peace on either side will rise and fall with the probability of further successes and the amount of effort these would require.”

The presence of uncertainty allows weaker states to exploit informational asymmetries because they know their opponent has incentives to settle as soon as possible and so will make offers that exceed the ones under complete information. In other words, weak adversaries can actually profit from fighting a little and then settling, even though they know very well that in the long run they will inevitably lose.

Although it is possible for war to end with the complete military defeat of one side, such instances will be rare. This is what we observe empirically as well: Most wars do not terminate with the obliteration of the losing side but are settled long before that. This follows immediately from the informational role of warfare—once opponents learn “enough” about each other, they can find a mutually acceptable settlement. How much “enough” is and how long it takes to learn it depend on prior beliefs, the battlefield performance, and the speed with which information is revealed during the war.

There are substitutes for victory in war, and these are the political settlements that terminate the war. Since it is not necessary to win a war in order to end it, it is not necessary to agree on who the eventual winner will be either.

### On War Initiation

The Principle of Convergence also has startling implications for the initiation of war. If it is not necessary to win a war in order to profit from it, it is not necessary to believe in victory to start one. The traditional theoretical puzzle is framed as one about divergent optimistic expectations about the outcome of war, and the literature has sought ways to account for such discrepancies in rationalist terms (Fearon 1995).

My results suggest that this places unnecessary demands on optimism: It is quite rational to start a war a state expects to lose as long as its opponent believes that this state is stronger because this induces the opponent

<sup>16</sup> Refusing to talk can be a signal itself. However, since it is relatively easy to obtain inefficiency in bargaining models where players can delay making offers (Admati and Perry 1987), requiring that they always do make offers increases the hurdle the model must pass. This assumption increases the rate of information transmission, but it is useful because it highlights one reason why explicit intrawar diplomacy can be so difficult.

to offer better terms. It is not necessary for both sides to be optimistic about the *military* outcome.

This further implies that asymmetric conflict may not be as puzzling as once thought. It does not require suicidal or desperate rulers for a weak state to challenge a strong one, although it is necessary for the strong one to believe its adversary is not as weak as it actually is.

Filson and Werner find that war can begin when the initiator underestimates the strength of the defender, who then fights to demonstrate that it does not have to concede as much. This also happens in the present model, but the additional rationale for a weak state fighting an opponent because the opponent may overestimate its strength cannot be derived from their model. This is a consequence of limiting the types of opponents to only two, which deprives moderately strong players of the ability to separate from the weak while pooling with the strong, which is what generates their higher payoff in the present model.

### Duration of Total and Restricted Wars

Scholars often make a conceptual distinction between *restricted* and *total* war (Kecskemeti 1958; Manwaring 1987; Wagner 2000).<sup>17</sup> The difference between the two is mainly in the way they end or are expected to end. The former type is seen as achieving limited goals with qualified military engagements. Restricted wars usually end with a settlement and both sides retain fighting capacity. The latter type is seen as seeking the complete destruction of the military capability of the defeated state.

The Principle of Convergence challenges the distinction between wars fought “solely for the purpose of revealing information and wars fought to disarm the adversary” (Wagner 2000). In the model war can end with the complete military defeat of one state before a bargain is struck, but the outcome arises from the same mechanism that produces the political settlements. Moreover, it is very unlikely that such an outcome will occur because, all things equal, war results in relatively quick disclosure of information, and so is wont to lose its informational content without significant delay. Extreme uncertainty can produce delays that are long enough for parties to fail to reach an agreement. However, no one starts out intending to wage an absolute war, it just may turn out that way, as it happened in the Second World War.<sup>18</sup>

The *ceteris paribus* qualification is important. Innovations can significantly alter the prewar capabilities

while fighting lasts. Depending on the identity of the successful innovator, this may result in further delays—a state that has become stronger demands more and rejects terms—or dramatic shortening—a weak state finds it no longer can extract concessions from a now significantly stronger opponent.

### The Shadow of the Future

States must value the future sufficiently for war to occur because only when the settlement is important enough to the players do they have incentives to bear costly delays to secure better outcomes. When players discount the future too much, only semiseparating and pooling equilibria exist instead. This implies that more stable governments will be harder to settle with, which also helps explain why it is often the case that the parties who begin the war are not the same ones that finish it, especially on the losing side (Calahan 1944).

In international relations theory, the shadow of the future is usually thought to have a benign effect because it makes cooperative behavior possible in equilibrium. However, in a situation where uncooperative behavior in the short term may secure better benefits in the long run, the shadow of the future has the exact opposite effect. The more patient players are, the more incentives they have to delay agreement and fight.<sup>19</sup>

### CONCLUSION

If warfare is purely instrumental, then its role as an information transmitter is paramount. War results not from bargaining failure but from incentives to determine the type of opponent and obtain a better negotiated settlement. In a costly, but noisy, fighting environment such information takes time to accrue, and only when players learn sufficiently about their prospects in war will they agree to a settlement that reflects these expectations. The Principle of Convergence posits that warfare ceases to be useful when it loses its informational content.

These results have significant implications for the theory of war because they show the standard puzzle about both sides being optimistic about victory to be irrelevant. States do not have to believe in victory to engage in war, and neither do they have to believe in defeat to end it. As long as both sides want to settle as quickly as practicable, weaker states can benefit from uncertainty and obtain a deal through fighting that they would not have been able to obtain under complete information.

More attention must be paid to the different sources of information in war. The battlefield is a noisy but involuntary source, while the bargaining table is a precise but manipulable one. There are other sources of information between the two extremes, like public opinion and intelligence. It will be worth investigating how

<sup>17</sup> I use the term “restricted” war to denote the distinction between fighting to disarm an opponent (military victory) and fighting to achieve a political objective (which does not require disarming). This is different from *limited* war, which refers to opponents refraining from using all their capabilities from fear of a potentially disastrous escalation.

<sup>18</sup> This war also furnishes an example of intraalliance politics that made it hard to settle before the complete disarmament of the opponent. The fear that allowing for anything less than an unconditional surrender by Germany might split the uneasy alliance with the Soviets contributed to prolonging the war until Berlin itself fell.

<sup>19</sup> This result corresponds with the findings in Fearon 1998; and R. Powell 1993. Even though the latter arises from a very different model, the rationale is the same.

these can be utilized and how states interpret information acquired during war.

A number of testable hypotheses can be derived from the model. The skimming property of the equilibrium implies that as war progresses, the outcome becomes less advantageous for the worse informed party. Since it takes optimism to select oneself into war, it is likely that this party will be the initiator.

Information acquired during the war outweighs pre-war information derived from capabilities and economic resources. If the war begins with a series of early victories against a strong opponent, the screening process slows down because the proposer fails to reduce its demands fast enough to induce its opponent to quit. A series of initial defeats accelerates the screening process and will lead to shorter wars. Thus, we can address the question of how initial performance in war influences its duration.

The military position achieved immediately prior to negotiating has a strong effect on the bargaining outcome in the sense that a state may prefer to delay agreement following a victory when it will prefer to settle following defeat.

Since behavior at bargaining table has such important consequences for the termination of war, states will try to manipulate the prospect of negotiations by refusing to come to the table.

Because it is the strong that are most hurt by the informational asymmetries, powerful states will seek to reveal their strength, and moderately strong or weak ones will oppose such transparency.

The Weinberger–Powell Doctrine places rather stringent restrictions on the use of force because it views warfare through the prism of military victory. This is counterproductive because it may lead to failure to engage in situations where a determined, yet limited, application of force can yield satisfactory results. The emphasis on the use of overwhelming force, on the other hand, seems well placed, both because it discourages adversaries from costly delays and because it enhances the deterrent posture as long as the use of force is credibly contemplated.

Insofar as the new National Security Strategy of the Bush administration signifies increased willingness to use force, it should provide for better deterrence against traditional adversaries. The strategy implicitly seeks to shorten the time horizon of the opponent (regime change for state actors, preemptive strike against non-state enemies), a distinctly novel approach in contrast to (sometimes limited) *détente* with status quo powers like the Soviet Union during the Cold War. Since long time horizons are a major reason for engaging in conflict, this shift should produce desirable effects by making adversaries less willing to engage in costly contests with the United States.

## APPENDIX

In any stationary no-delay MPE, player 1 must offer player 2 at least what that player expects to obtain by rejecting a proposal. Since in this equilibrium player 2's offer is immediately accepted (or else the game ends), player 1's offer must satisfy,

for  $1 < k < N - 1$ ,

$$\begin{aligned} \pi - x_1^* &= (1 - \delta)b_2 + \delta[py_2^* + (1 - p)\pi], \\ \pi - x_k^* &= (1 - \delta)b_2 + \delta[py_{k+1}^* + (1 - p)y_{k-1}^*], \\ \pi - x_{N-1}^* &= (1 - \delta)b_2 + \delta(1 - p)y_{N-2}^*. \end{aligned} \tag{1}$$

The corresponding equations for player 2's offers are

$$\begin{aligned} \pi - y_1^* &= (1 - \delta)b_1 + \delta px_2^*, \\ \pi - y_k^* &= (1 - \delta)b_1 + \delta[px_{k+1}^* + (1 - p)x_{k-1}^*], \\ \pi - y_{N-1}^* &= (1 - \delta)b_1 + \delta[p\pi + (1 - p)x_{N-2}^*]. \end{aligned}$$

This defines a system of  $2(N - 1)$  simultaneous equations. Fix some arbitrary  $N \geq 2$  and let  $n = N - 1$ . Label the equilibrium offers such that  $x_1^*, x_2^*, \dots, x_n^*$  correspond to  $z_1, z_2, \dots, z_n$ , and  $y_1^*, y_2^*, \dots, y_n^*$  correspond to  $z_{n+1}, z_{n+2}, \dots, z_{2n}$ . Construct the  $2n \times 2n$  matrix  $\mathbf{A}$  of coefficients in the usual way and let  $\mathbf{w} > \mathbf{0}$  be the corresponding RHS vector. The following lemma establishes that there is a unique solution to the system of equations.

**Lemma 3.** *There exists a unique  $\mathbf{z}^* = \mathbf{A}^{-1}\mathbf{w}$ .*

*Proof.* Since  $\mathbf{w} \neq \mathbf{0}$ , it is sufficient to establish that  $\mathbf{A}^{-1}$  exists.  $\mathbf{A}$  can be partitioned into four square submatrices:

$$\mathbf{A} = \begin{pmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{M} & \mathbf{I} \end{pmatrix},$$

where  $\mathbf{I}$  is the identity matrix of size  $n$  and  $\mathbf{M}$  is an  $n \times n$  matrix whose diagonal elements are 0, immediate lower off-diagonal elements are  $\delta(1 - p)$ , immediate upper off-diagonal elements are  $\delta p$ , and everything else is 0. Like  $\mathbf{M}$ , each element of  $\mathbf{M}^2$  is nonnegative, and the sum of entries in each column is less than one. By Theorem 8.13 in Simon and Blume 1994, 175, this implies that  $(\mathbf{I} - \mathbf{M}^2)^{-1}$  exists. But since  $\det \mathbf{A} = \det(\mathbf{I} - \mathbf{M}^2)$ , it follows that  $\det \mathbf{A} \neq 0$  as well.  $\square$

*Proof of Proposition 1.* Consider the following strategies. Player 1 always offers  $x_k^*$ , accepts all offers  $x \geq x_k^*$ , and rejects all offers  $x < x_k^*$ , where  $k$  is the realization of the stochastic process and  $x_k^*$  is the  $k$ th element of  $\mathbf{z}^*$  from Lemma 3. Player 2's strategy is defined analogously. It is trivial to verify that these strategies are subgame perfect. Since the vector with proposals is unique, there exists at most one stationary no-delay MPE. Agreement is immediate on  $x_{k_0}^*$ .  $\square$

*Proof of Proposition 2.* In the proposed equilibrium, the game in period  $t = 2$  is equivalent to the complete information game with  $p = p_L$  and starting state  $k_2$ . By Proposition 1, this game has a unique stationary no-delay MPE in which player 1 offers and player 2 accepts  $V_s^1(k_2)$ .  $\square$

## The Two-Period Game

Let  $k_1$  denote the realization of the state variable in period  $t = 1$ , and let  $q_1^w = 0$ ,  $q_1^m > 0$ , and  $q_1^s = 1 - q_1^m > 0$  denote player 1's prior (i.e., at the beginning of the period, before player 2 makes an offer) belief that its opponent is of type  $2_w$ ,  $2_m$ , and  $2_s$ , respectively. The probability that the opponent is weak is zero because players follow equilibrium strategies. Let  $p_1 = q_1^s p_L + (1 - q_1^s) p_M$  be player 1's ex ante expectation that it will win a fight if it rejects an offer in period 1. Player 2's unique (by Lemma 2) equilibrium offer that is accepted by player 1 is

$$\begin{aligned} y^*(q_1^s, k_1) &= \pi - (1 - \delta)b_1 - \delta[p_1 V_s^1(k_1 + 1) \\ &\quad + (1 - p_1)V_s^1(k_1 - 1)]. \end{aligned} \tag{2}$$

Since  $2_s$  makes an unacceptable offer, it must be the case that

$$y^*(q_1^s, k_1) < (1 - \delta)b_2 + \delta[\pi - p_L V_s^1(k_1 + 1) - (1 - p_L)V_s^1(k_1 - 1)],$$

which, together with condition (2), yields the additional constraint

$$p_M - p_L > \left[ \frac{1 - \delta}{\delta(1 - q_1^s)} \right] \left[ \frac{\pi - b_1 - b_2}{V_s^1(k_1 + 1) - V_s^1(k_1 - 1)} \right]. \quad (3)$$

**Lemma 4.** *Let  $\underline{\delta}_1(q_1^s)$  be the smallest discount factor that solves (3) for some  $q_1^s$ . Whenever  $\delta \geq \underline{\delta}_1(q_1^s)$ , player  $2_s$  strictly prefers to make an unacceptable offer in period  $t = 1$  to making an acceptable one.*

This lemma implies that the optimal strategy for  $2_w$  and  $2_m$  in period  $t = 1$  is to demand, and receive,  $y^*(q_1^s, k_1)$ . Player  $2_s$  demands, but does not receive,  $y > y^*(q_1^s, k_1)$  as long as  $\delta \geq \underline{\delta}_1(q_1^s)$ . Player 1 accepts all  $y \leq y^*(q_1^s, k_1)$ , and rejects everything else. Let  $x_1^*(q_1^s, k_1) = \pi - y^*(q_1^s, k_1)$  denote player 1's smallest expected payoff in  $t = 1$ .

**Beliefs Following a Battle.** In equilibrium, an acceptance of the first offer by player 2 signals unambiguously that its type is  $2_w$ , and therefore a rejection signals that its type is either  $2_m$  or  $2_s$ . The prebattle probability that the type is  $2_s$  equals  $q^s / (q^s + q^m)$ . The post-battle posterior is then

$$q_1^s(I) = \frac{q^s p_L^I (1 - p_L)^{1-I}}{q^s p_L^I (1 - p_L)^{1-I} + q^m p_M^I (1 - p_M)^{1-I}}, \quad (4)$$

where  $I$  is the battle indicator that equals 1 if player 1 won or 0 if it lost.

**The Three-Period Game.** Let  $p_0 = (q^m p_M + q^s p_L) / (q^m + q^s)$  be player 1's expectation that it will win a fight if  $2_m$  and  $2_s$  reject its offer. Since player  $2_w$  accepts player 1's offer, the optimal offer in  $t = 0$  is

$$x_0^* = \pi - (1 - \delta)b_2 - \delta[p_H y^*(q_1^s(1), k_0 + 1) + (1 - p_H)y^*(q_1^s(0), k_0 - 1)]. \quad (5)$$

Player 1's ex ante payoff from a strategy that induces only  $2_w$  to accept with probability one is at least

$$x^* = q^w x_0^* + (1 - q^w)[(1 - \delta)b_1 + \delta(p_0 x_1^*(q_1^s(1), k_0 + 1) + (1 - p_0)x_1^*(q_1^s(0), k_0 - 1))]. \quad (6)$$

In equilibrium this must be at least as good as the payoffs from inducing only  $2_s$  to separate, or from settling immediately with all three types.

If player 2 follows the equilibrium strategy, the only way player 1 can induce only  $2_s$  to reject the initial offer is to satisfy player  $2_m$ , which implies that player 1 can propose at most

$$\hat{x}_0 = \pi - (1 - \delta)b_2 - \delta[p_M y^*(q_1^s(1), k_0 + 1) + (1 - p_M)y^*(q_1^s(0), k_0 - 1)].$$

If both  $2_w$  and  $2_m$  accept the initial offer,  $q_1^s = 1$ , and Lemma 4

has no bite:  $2_s$  will not delay but will demand instead  $V_2^s(k_1)$ , to which player 1 will agree. Therefore, player 1's ex ante payoff from inducing both  $2_w$  and  $2_m$  to accept the initial offer is

$$\hat{x} = (1 - q^s)\hat{x}_0 + q^s[(1 - \delta)b_1 + \delta(\pi - p_L V_s^2(k_0 + 1) - (1 - p_L)V_s^2(k_0 - 1))].$$

If player 2 follows the equilibrium strategy, the only way player 1 can guarantee that all three types will accept the initial proposal is to satisfy  $2_s$ , which implies that player 1 can propose at most

$$\hat{x}_0 = (1 - \delta^2)(\pi - b_2) + \delta^2[p_L^2 V_s^1(k_0 + 2) + 2p_L(1 - p_L)V_s^1(k_0) + (1 - p_L)^2 V_s^1(k_0 - 2)],$$

whose ex ante expected value is  $\hat{x} = \hat{x}_0$  because the offer is accepted with probability one.

To establish optimality of player 1's screening strategy, we wish to show that it would not want to deviate by either settling with all three types immediately or separating only  $2_s$  for one period. The strategy will be optimal as long as  $\hat{x} \leq x^*$  and  $\hat{x} \leq x^*$ .

After algebraic manipulation, we find that  $\hat{x} < x^*$  whenever

$$p_M - p_L > \left[ \frac{q^s(1 - \delta)}{\delta(1 - q^s)} \right] \left[ \frac{\pi - b_1 - b_2}{V_s^2(k_0 - 1) - V_s^2(k_0 + 1)} \right]. \quad (7)$$

**Lemma 5.** *Let  $\underline{\delta}_0$  be the smallest discount factor that solves (7). Whenever  $\delta \geq \underline{\delta}_0$ , player 1 strictly prefers to delay agreement for one period in order to separate  $2_s$  from the other types to settling immediately.*

Showing that  $\hat{x} \leq x^*$  is a bit involved, but the following result can be established using numerical methods.

**Lemma 6.** *Let  $\underline{\delta}_2$  be the smallest discount factor such that  $\hat{x} < x^*$ . Whenever  $\delta \geq \underline{\delta}_2$ , player 1 strictly prefers to delay agreement for up to three periods in order to separate each of the three types to settling in any of the prior periods.*

With these results, the proof of the proposition is straightforward. It is always the case that  $\underline{\delta}_1(q_1^s(1)) < \underline{\delta}_1(q_1^s(0))$ . Let  $\underline{\delta} = \max\{\underline{\delta}_0, \underline{\delta}_1(q_1^s(0)), \underline{\delta}_2\}$ , and take any  $\delta \in (\underline{\delta}, 1)$ . Consider the following strategy for player 1: in  $t = 0$ , offer  $x^*$  from (6); in  $t = 1$ , accept any  $y \leq y^*(q_1^s(I_0), k_1)$  from (2) and reject anything else; in all even periods  $t \geq 2$ , offer  $V_s^1(k_t)$ ; in all odd periods  $t \geq 2$ , accept any  $y \leq V_s^2(k_t)$  and reject anything else. In period 1 update to believe that the probability of  $2_s$  is  $q_1^s(I_0)$ , as defined in (4), and the probability of  $2_w$  is zero. In periods  $t \geq 2$  update the probability of  $2_s$  to 1.

Consider the following type-dependent strategy for player 2. Type  $2_w$  accepts any  $x \leq x_0^*$  in period  $t = 0$ , and follows  $2_m$ 's strategy henceforth. Type  $2_m$  accepts any  $x \leq \hat{x}_0$  in period  $t = 0$ , offers  $y^*(q_1^s(I_0), k_1)$  in period  $t = 1$ , and follows  $2_s$ 's strategy henceforth. Type  $2_s$  accepts any  $x \leq \hat{x}_0$  in period  $t = 0$ , demands  $\pi$  in period  $t = 1$ , and for all  $t \geq 2$  accepts  $x \leq V_s^1(k_t)$  in even-numbered periods and demands  $V_s^2(k_t)$  in odd-numbered periods.

Proposition 1 and lemmas 4, 5, and 6 guarantee that the constructed Markov strategies and player 1's credible beliefs constitute a perfect sequential equilibrium. The construction demonstrates that there is one type of MPSE but it is only generically unique because player  $2_s$  can make different non-serious offers in period  $t = 1$ . There exists a continuum of equilibria of this type, but they are all payoff-equivalent.  $\square$

## REFERENCES

- Admati, Anat R., and Motty Perry. 1987. "Strategic Delay in Bargaining." *Review of Economic Studies* 54 (July): 345–64.
- Blainey, Geoffrey. 1988. *The Causes of War*. 3rd ed. New York: Free Press.
- Calahan, H. A. 1944. *What Makes a War End?* New York: Vanguard Press.
- Cimbala, Stephen J., and Keith A. Dunn, eds. 1987. *Conflict Termination and Military Strategy: Coercion, Persuasion, and War*. Boulder: Westview Press.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (Summer): 379–414.
- Fearon, James D. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52 (2): 269–305.
- Filson, Darren, and Suzanne Werner. 2002. "A Bargaining Model of War and Peace: Anticipating the Onset, Duration, and Outcome of War." *American Journal of Political Science* 46 (October): 819–38.
- Forster, Kent. 1941. *The Failures of Peace: The Search for a Negotiated Peace During the First World War*. Washington, DC: American Council on Public Affairs.
- Foster, James L., and Garry D. Brewer. 1976. "And the Clocks Were Striking Thirteen: The Termination of War." *Policy Sciences* 7 (June): 225–43.
- Fox, William T. R. 1970. "The Causes of Peace and Conditions of War." *Annals of the American Academy of Political and Social Science* 392 (November): 1–13.
- Fudenberg, Drew, and Jean Tirole. 1991. *Game Theory*. Cambridge, MA: MIT Press.
- Fudenberg, Drew, David Levine, and Jean Tirole. 1985. "Infinite-Horizon Models of Bargaining with One-Sided Incomplete Information." In *Game-Theoretic Models of Bargaining*, ed. Alvin Roth. New York: Cambridge University Press, 73–98.
- Fuller, J. F. C. 1961. *The Conduct of War, 1789–1961*. New Brunswick, NJ: Rutgers University Press.
- Gartner, Scott Sigmund. 1997. *Strategic Assessment in War*. New Haven, CT: Yale University Press.
- Goemans, Hein E. 2000. *War and Punishment: The Causes of War Termination and the First World War*. Princeton, NJ: Princeton University Press.
- Grimmett, G. R., and D. R. Stirzaker. 1992. *Probability and Random Processes*. 2nd ed. Oxford: Oxford University Press.
- Grossman, Sanford J., and Motty Perry. 1986a. "Perfect Sequential Equilibrium." *Journal of Economic Theory* 39: 97–119.
- Grossman, Sanford J., and Motty Perry. 1986b. "Sequential Bargaining under Asymmetric Information." *Journal of Economic Theory* 39: 120–54.
- Harsanyi, John C. 1968. "Games with Incomplete Information Played by 'Bayesian' Players." *Management Science* 14 (March): 159–82, 320–34, 486–502.
- Hobbs, Richard. 1979. *The Myth of Victory: What Is Victory in War?* Boulder, CO: Westview Press.
- Iklé, Fred Charles. 1964. *How Nations Negotiate*. New York: Harper & Row.
- Iklé, Fred Charles. 1971. *Every War Must End*. New York: Columbia University Press.
- Keckskemeti, Paul. 1958. *Strategic Surrender: The Politics of Victory and Defeat*. Stanford, CA: Stanford University Press.
- Keckskemeti, Paul. 1970. "Political Rationality in Ending War." *Annals of the American Academy of Political and Social Science* 392 (November): 105–15.
- Kreps, David, and Robert Wilson. 1982. "Sequential Equilibrium." *Econometrica* 50: 863–94.
- Manwaring, Max G. 1987. "Limited War and Conflict Control." In *Conflict Termination and Military Strategy: Coercion, Persuasion, and War*, ed. Stephen J. Cimbala and Keith A. Dunn. Boulder, CO: Westview Press, 59–76.
- Merlo, Antonio, and Charles Wilson. 1995. "A Stochastic Model of Sequential Bargaining with Complete Information." *Econometrica* 63 (March): 371–99.
- Muthoo, Abhinay. 1999. *Bargaining Theory with Applications*. Cambridge: Cambridge University Press.
- Myerson, Roger B., and Mark A. Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory* 28: 265–81.
- Nicolson, Harold G. 1954. *The Evolution of Diplomatic Method*. London: Constable.
- Pillar, Paul R. 1983. *Negotiating Peace: War Termination as a Bargaining Process*. Princeton, NJ: Princeton University Press.
- Powell, Colin L. 1992. "U.S. Forces: Challenges Ahead." *Foreign Affairs* 71 (Winter): 32–45.
- Powell, Robert. 1993. "Guns, Butter and Anarchy." *American Political Science Review* 87 (March): 115–32.
- Powell, Robert. 1996. "Stability and the Distribution of Power." *World Politics* 48 (January): 239–67.
- Powell, Robert. 2001. "Bargaining While Fighting." Mimeo, University of California, Berkeley. October 7.
- Reiter, Dan. 2003. "The Bargaining Model of War." *Perspectives on Politics* 1 (March): 27–43.
- Roth, Alvin, ed. 1985. *Game-Theoretic Models of Bargaining*. New York: Cambridge University Press.
- Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50: 97–110.
- Rubinstein, Ariel. 1985a. "A Bargaining Model with Incomplete Information About Time Preferences." *Econometrica* 53 (September): 1151–72.
- Rubinstein, Ariel. 1985b. "Choice of Conjectures in a Bargaining Game with Incomplete Information." In *Game-Theoretic Models of Bargaining*, ed. Alvin Roth. New York: Cambridge University Press, 99–114.
- Rubinstein, Ariel. 1991. "Comments on the Interpretation of Game Theory." *Econometrica* 59 (July): 909–24.
- Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
- Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: 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.
- Simon, Carl P., and Lawrence Blume. 1994. *Mathematics for Economists*. New York: W. W. Norton.
- Slantchev, Branislav L. 2002. "Markov Perfect Equilibrium in a Stochastic Bargaining Model." Typescript. University of California, San Diego.
- Smith, Alastair. 1998. "Fighting Battles, Winning Wars." *Journal of Conflict Resolution* 42 (June): 301–20.
- Smith, Alastair, and Allan C. Stam, III. 2001. "Bargaining through Conflict." Paper presented at the Annual Meeting of the Public Choice Society, San Antonio, TX.
- Smith, James D. D. 1995. *Stopping Wars: Defining the Obstacles to Cease-Fire*. Boulder, CO: Westview Press.
- von Clausewitz, Carl. (1832) 1984. *On War*, ed., transl. Michael Howard and Peter Paret. Princeton, NJ: Princeton University Press.
- Wagner, R. Harrison. 1994. "Peace, War, and the Balance of Power." *American Political Science Review* 88 (September): 593–607.
- Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44 (July): 469–84.
- Weinberger, Caspar W. 1984. "The Uses of Military Power." Remarks by Secretary of Defense Weinberger to the National Press Club, Washington, DC, November 28.
- Wittman, Donald. 1979. "How a War Ends: A Rational Model Approach." *Journal of Conflict Resolution* 23 (December): 743–63.