The Armed Peace: A Punctuated Equilibrium Theory of War

Bahar Leventoğlu
Department of Political Science, Stony Brook University

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

April 6, 2006

Abstract. According to a leading rationalist explanation, war can break out when a large rapid shift of power renders unbelievable a rising state's promise to compensate its declining opponent, causing the latter to attack preventively. This mechanism does not provide a complete and coherent explanation of war because it does not specify how inefficient fighting resolves this commitment problem. We present a complete information model of war as a sequence of battles and show that although opportunities for a negotiated settlement arise throughout, the very desirability of peace creates a commitment problem that undermines its likelihood. Because players have incentives to settle as soon as possible, they cannot credibly threaten to fight long enough if an opponent launches a surprise attack. This decreases the expected duration and costs of war and causes mutual deterrence to fail. Fighting's sheer destruction improves the credibility of these threats by decreasing the benefits from continuing the war. Equilibrium fighting may involve escalating costs that exceed the value of the stakes by the time peace is negotiated and that leave both players worse off than when the war began.

*Author names in alphabetical order. Slantchev is the corresponding author. We are grateful to participants of seminars at Stanford University, Dartmouth College, Texas A&M University, and the Conflict and Cooperation Conference at the Kellogg School of Management, Northwestern University for comments. This research was supported by the National Science Foundation (grant SES-0518222 to Slantchev and SES-0518185 to Leventoğlu). Presented at the 2005 annual meeting of the American Political Science Association (Washington, DC), and the 2006 meeting of the Midwest Political Science Association (Chicago).

To understand why wars begin, we have to know why they end—the termination of war must involve the resolution of its causes (Blainey 1988, x). Imagine a situation in which a powerful state faces a rising opponent. The declining state could wage costly preventive war while it is still strong or it can wait and be forced to accept an unpalatable revision tomorrow. Because the challenger is still weak, it would rather live in peace, and so it must compensate the declining state for foregoing an advantageous war. If the amount of this compensation exceeds the total benefits states can divide today, the challenger must promise to deliver the rest in the future. If the power shift is sufficiently rapid and large, however, the now dominant challenger will renege on that promise. As a result, the declining state cannot believe it today and goes to war.

This commitment problem is one of the leading rationalist explanations of war (Fearon 1995), the underlying mechanism is quite general (Powell 2004*b*), and there are very good substantive reasons to insist on a theory that does not rely on incomplete information to explain war (Powell 2006). Unfortunately, the theory does not specify how inefficient fighting can resolve this commitment problem short of the total obliteration of one of the warring states. Because outcomes of such finality are very rare, the theory does not provide a satisfactory answer to this puzzle. But if fighting does not resolve the commitment problem, it is pointless to bear its costs and risks in this context: in the end, the commitment problem will still exist and both states would have suffered tremendously. Moreover, if the opponents can terminate the war without resolving the commitment problem, then what does it mean to say that the war was caused by this problem in the first place?

We have several goals in this article. First, we explain why it is imperative that our explanations of war are complete and coherent, and then discuss how the commitment problem mechanism fails these requirements and why. This leads us to a model with stylized representation of war that simultaneously addresses our objections to the traditional approach and allows for the commitment problem to arise. Our second goal is to demonstrate using this model that large rapid shifts of power are only necessary but not sufficient to cause war. Instead, peace fails when states cannot credibly threaten to impose large enough costs to deter each other from trying to exploit the advantages of surprise attack. The credibility problem arises from the very desirability of peace; states cannot threaten to prolong fighting more than absolutely necessary. We then use these results to show how rational actors can escalate and incur costs far out of proportion to the gains, why wars tend to be settled when actors are nearing exhaustion, and why peace is only possible at specific junctures during the conflict.

1 The Credible Commitment Problem

The *bargaining model of war* posits a fundamental puzzle that any rationalist theory of war must be able to answer. Because fighting destroys resources, the benefits that players can divide after a war are always less than the benefits they could have divided prior to it. This means that it should be possible to locate *ex ante* agreements that would leave both players better off compared to what they can expect to get from fighting. The *inefficiency puzzle* is why actors are unable to reach such an agreement even when their existence is common knowledge.[1]

While most of recent work has focused on informational asymmetries as a cause of war, there are compelling reasons to consider causal mechanisms that do not depend on incomplete information.[2] First, uncertainty is ubiquitous and is present in almost all crises, and yet only few of these escalate into war. Second, the information revelation mechanism cannot deal very well with long wars: to think that it takes many years of near constant interaction for opponents to learn enough about each other is surely stretching the theory. Third, Leventoğlu and Tarar (2005) show that private information by itself often merely leads to delay in reaching a negotiated settlement preferable to war, and for bargaining to break down in war requires that players are impatient enough.

This means that we need an approach that does not rely exclusively on incomplete information. The major explanation of war under complete information is the *credible commitment problem* (CCP) caused by large, rapid shifts of power. Imagine an environment in which an actor that is weak today will become strong tomorrow. Because the opponent is relatively strong now, its expected payoff from fighting is rather high, so the transfer that the temporarily weak actor must agree to has to be correspondingly large. With resource constraints, this amount may exceed the benefits currently available, and so the weak actor must commit to transfer the remainder over several periods. However, when power shifts to that actor's advantage, its incentives to follow through on that agreement will be undermined, rendering today's commitment incredible. In such a situation, the weak may be unable to prevent its opponent from resorting to force (Organski 1968, Fearon 1995).

Powell (2004*b*) derives a condition that ensures that large rapid shift of power will be sufficient to undermine the credibility of the promises of the temporarily weak actor and cause war. Powell (2006) further demonstrates that bargaining indivisibilities, first-strike advantages (Fearon 1995) and bargaining over objects that are sources of military power (Fearon 1996) also turn out to be manifestations of this fundamental mechanism. In other words, a great many complete-information explanations for war involve this commitment problem at their hearts.[3]

The CCP mechanism, however, raises another fundamental puzzle: *how does fighting resolve the commitment problem?* To see why this is important, note that any rationalist explanation of war must meet two minimalist criteria. It must be *complete*, which means

[1]Fearon (1995) was the first to formulate the problem in this way. See Powell (1999) for this inefficiency puzzle, and Powell (2002) for an overview of this literature.

[2]Briefly, the informational story is that leaders possess private information about their expected payoffs from war and peace, and they have incentives to misrepresent this knowledge to extract bargaining advantage. War can break out when actors bargain in the shadow of power and pursue strategies with which they run a slightly higher risk of war in return for obtaining slightly more at the bargaining table. For more on the risk-return trade off as a cause of war, see Fearon (1995), Powell (1999), and Slantchev (2003*b*). Powell (2006) provides reasons the informational approach may be somewhat limited substantively.

[3]This mechanism is not exhaustive. If arming is costly, then defending the status quo by deterring the opponent may be costlier than fighting. If peace is too expensive, bargaining can break down in war (Powell 1993). Slantchev (2005) shows how private information may lead to arming decisions that can cause war in this way. Another explanation focuses on coordination problems: when there are multiple equilibria, actors may fail to coordinate on an efficient one. Slantchev (2003*a*) analyzes a situation where fear of early settlement drives inefficient behavior. Finally, Langlois and Langlois (2005) propose another mechanism that relies on strategic uncertainty induced by mixed strategies.

it must account for war's outbreak and termination; and it must be *coherent*, which means that its account of termination must explain how fighting has resolved the cause (Author). In other words, a theory of war is complete and coherent if it can identify a cause of war that fighting then resolves. Whereas it can be shown that the asymmetric information explanation is both coherent and complete, the CCP is neither because it is silent on how inefficient fighting can resolve the commitment problem.[4]

One trivial answer is that it does so by eliminating one of the opponents. However, this is empirically untenable: many wars end in negotiated settlements. For example, out of 104 interstate wars between 1816 and 1991, 67 (or 64%) ended short of total military victory for one side.[5] Most wars end because both sides *agree* to cease hostilities while both could fight on. That is, they strike a bargain where previously they have been unable to do so. Furthermore, if we are to sustain CCP as a *cause* of war, then it must be the case that when players agree to settle, this problem is resolved. If this were not true, then peace would be possible even in the presence of the commitment problem, which would then imply that this problem could not have been the cause of fighting in the first place. Gartzke (1999, pp. 571-72) notes this shortcoming of CCP and even goes on to argue that because it cannot be overcome, any rationalist explanation of war necessarily requires incomplete information. We want to contest this claim by showing at least one way in which war can resolve the commitment problem.

We now enumerate several minimal features that the model must have to address the puzzle we have identified. First, we must represent war as a sequence of costly engagements rather than a game-ending costly lottery over exogenous outcomes because we wish to trace the effect of fighting. We must allow for negotiations during fighting, and we must allow for the game to end with a military victory in addition to a peace settlement. Second, we must incorporate large rapid shifts of power to create an environment where the original CCP can manifest itself. Third, contrary to most existing bargaining models, we should not assume that an agreement automatically ends the game with the division of benefits. Rather, we must allow for the possibility that actors can renege on the agreement and attempt to use the newly acquired resources to extract further concessions. In other words, peace, if it happens, must be endogenous to the model. Finally, ideally we would not want our results to be dependent on a particular choice of a bargaining protocol, and the model must have complete information.

[4]To see that the informational story is complete and coherent, observe that under the risk-return trade off points to the inability to ascertain the truthfulness of opponent's claims without risking war. This suggests that for peace to occur, they must be able to learn enough about each other so that this is no longer an issue. Slantchev (2003*b*) provides just such a rationale for war termination: battlefield outcomes and intrawar diplomacy allow actors to acquire information that eventually causes their expectations to converge sufficiently to permit an agreement. By eliminating private information, fighting essentially resolves the original cause of war. This convergence of expectations during war can be obtained in other settings. See Filson and Werner (2004), Powell (2004*a*), and Smith and Stam (2004).

[5]Our calculations using Slantchev's (2004) data set. which separates war outcomes into those where one side is unable to continue the military contest and those where both sides agree to a settlement while still physically able to continue fighting. For more on war outcomes, see Kecskemeti (1958), Iklé (1971), and Pillar (1983).

2 The Model

Two players, each initially endowed with some capital $K_i > 0$, dispute a prize worth $v \geq 0$.[6] Each period $t(t = 1, 2, \ldots)$ consists of two rounds: bargaining and fighting. Let $k_i(t) > 0$ denote the amount of resources available to player i at the beginning of period t. In the bargaining round, players negotiate over the distribution of resources comprising the entire surplus $S(t) = v + k_1(t) + k_2(t)$. If players reach an agreement, (x_1, x_2) such that $x_i \in [0, S(t)]$ and $x_1 + x_2 = S(t)$, they implement it and $k_i(t) = x_i$ (resources become immediately available for fighting). If they fail to reach an agreement, they keep whatever resources they had at the beginning of the period. After the bargaining round, players simultaneously choose whether to fight. If both choose not to fight, the game ends and the negotiated distribution of resources remains. If at least one player chooses to fight the other, a battle occurs and each loses $c_i \geq 1$ units of his resources, with $C = c_1 + c_2$ being the total cost. The game transits to the next period with probability $(1 - p)$, or one of the players collapses with probability $p < 1$. How this probability is distributed between the two players depends on their actions during the period: if they choose to fight, then either player 1 collapses with probability p_2 or player 2 collapses with probability $p_1 = p - p_2$; if i chooses to fight but j chooses not to, then j collapses with probability p, and i collapses with probability 0 (that is, if there is going to be a battle, it pays to participate in it).

How long players can survive fighting depends on their resource endowment. Let $K_i > 0$ be player i's initial capital stock. Because capital stocks are finite, the game ends in a finite number of periods. The payoffs are as follows. A player who collapses fighting in period t derives utility 0, and the surviving player gets $S(t) - C$; that is, the victor absorbs the loser's remaining resources. If player i's capital stock falls to zero in t, then i collapses automatically without an additional battle with a payoff of 0, and j's payoff is $S(t)$. If both players run out of resources at the same time, they collapse simultaneously and split the total surplus equally, each obtaining $S(t)/2$. If both players choose peace at time t, the game ends and each keeps the $k_i(t)$ as agreed upon. Since there is no confrontation in this case, there is no capital loss.

To specify the strategies formally, let $(x_1(t), x_2(t)) \in [0, S(t)]^2$ denote a proposal in the bargaining round of period t, $a_i(t) \in \{\text{accept, reject}\}$ denote i's decision to agree to a division, and $f_i(t) \in \{\text{fight, do not fight}\}$ denote i's decision to attack in the fighting round. A history $h_t = \{(x_1(s), x_2(s), a_1(s), a_2(s), f_1(s), f_2(s)) : s = 1, \ldots, t-1\}$ in period t is the list of proposals and actions up to that period. History ends when peace is reached or one of the players collapses, either militarily or through attrition. A bargaining protocol, $B = (b_1(h_t), b_2(h_t))$ determines the proposals in the bargaining round after each possible history h_t. We require that $b_1(h_t) + b_2(h_t) = S(h_t)$ for every history where $S(h_0) = S(0)$ at the beginning of the game. This formalization allows for any bargaining protocol that distributes all available resources in finite time in each bargaining round.

A pair of strategies $\sigma = (\sigma_1, \sigma_2)$ determines the players' actions after every possible history. Given a bargaining protocol, a strategy profile σ is a Nash equilibrium if σ_i is a best response to σ_j for $i, j \in \{1, 2\}$ and $i \neq j$. Given a bargaining protocol, a Nash equilibrium is subgame-perfect (SPE) if the subgame strategies induced by σ constitute

[6] We refer to player 1 (or a generic player i) as "he," and player 2 as "she."

a Nash equilibrium in every subgame. Subgame perfection eliminates incredible threats in the sense that no player would unilaterally deviate from the actions prescribed by the strategy when he is called to carry out a threat. As we shall see, this notion of credibility is too weak given the flexibility of the model that allows players to renegotiate even the bargaining protocol.

It is worth discussing briefly some of the assumptions in this model and how they relate to the requirements we outlined in the previous section. First, war is a sequence of engagements rather than an one-shot event. We shall refer to these engagements as battles with the understanding that we mean any period of time during which fighting occurs and actors do not negotiate (e.g., campaigns). Military victory can be achieved in two ways: either by causing the military collapse of the opponent during a campaign or by exhausting the opponent's resource base. However, the war can end as soon as both actors agree to a division of the benefits and as long as neither then engages in additional fighting after the distribution. The model thus allows for negotiated outcomes without assuming away the commitment to uphold the resulting distribution in peace. In addition, we have assumed that when actors bargain, they can use their resources for side-payments too. We allow for a general bargaining protocol as long as it ends in finite time and does not artificially preclude settlement (we explain later what we mean by that). Furthermore, by allowing actors to turn around and use the resources obtained by negotiation for further war, we have effectively endogenized peace: any equilibrium that involves a successful negotiated settlement will necessarily incorporate disincentives to renege from it.

Second, we have created an environment where the original CCP can arise by allowing for large rapid shifts of power when an actor surprises its opponent by attacking when the other is not. The shift of power arises from the increased probability of military collapse of the actor caught by surprise. This is very similar to the first-strike advantage notion that Fearon (1995) uses. The difference is that instead of conferring an advantage for the entire war, it only does so temporarily on the tactical level. We find this assumption much more tenable for two reasons. Strategic surprise, although possible to achieve, is often indecisive for the entire war (the two most famous examples are the Japanese attack on Pearl Harbor in 1941 and the Egyptian/Syrian attack on Israel in 1973). On the other hand, tactical surprise often is decisive for the particular engagement, and such an engagement could end the war. For example, the surprise Spartan destruction of the Athenian fleet at Aegospotami ended the Peloponnesian War, the surprise Soviet invasion of Manchuria coupled with the dropping of the American atomic bombs ended the War in the Pacific, and the surprise Israeli crossing of the Suez Canal with the resulting encirclement of the Egyptian army in the Sinai ended the Yom Kippur War. The assumption that the side which achieves surprise has zero probability of collapse in that engagement is made for simplicity and does not affect the results as long as that probability is well below the one for the surprised opponent. Finally, it is worth stressing that the costs are per engagement, not for the entire war. This makes the model a bit more attractive on substantive grounds: we do not have to assume, like most models do, that war costs are relatively small. Indeed, as we shall find, cumulative war costs here can be gargantuan, far exceeding the value of the prize actors are fighting over.

3 Total War

Total war occurs when players fight until one of them collapses from exhaustion. In this section, we establish the conditions for such an equilibrium. (All proofs are in Appendix A.) We first show that if there exists a period from which players will fight at least one battle, then they will also fight in all previous periods provided S is large enough. We then prove that players will fight to the end if at least one of them is sure to collapse after one battle. Together, these results imply that if S is large enough, total war will be inevitable: in any SPE, players will fight without redistributing resources until one of them collapses from exhaustion.

LEMMA 1. *If players fight for $T \geq 1$ periods when the resources are (k_1, k_2, S), then they fight $T+1$ periods when the resources are $(k_1 + c_1, k_2 + c_2, S + C)$ provided that S and p are large enough.*

It is worth elaborating what this lemma means. We are not claiming that if players fight for some $(k_1, k_2; S)$, then they would fight for an arbitrary distribution $(\hat{k}_1, \hat{k}_2; S+C)$ where $\hat{k}_i \neq k_i + c_i$. Rather, we prove that if the resource distribution $(k_1, k_2; S)$ is such that players fight, then players will also fight when the resource distribution is exactly $(k_1 + c_1, k_2 + c_2; S + C)$ as long as S and p are large enough (observe that a large v will be sufficient to ensure that his will happen). The next step is to locate a period in which players will fight for sure.

LEMMA 2. *Consider a period with $(k_1, k_2; S + C)$. If $k_i \in [c_i, 2c_i)$ and $k_j \geq 2c_j$, then players fight in this period provided that $p_j S > C$.*

In words, if one of the players has just enough resources for exactly one additional battle, then no peace is possible. If players have not redistributed resources until the period prior to i's imminent collapse, then $S > v$, and hence $p_j v > C$ will be sufficient to ensure war in that period. This means that if v is large enough and players do not redistribute, then the last battle is inevitable. We now show that if the condition of Lemma 1 is satisfied, then players would not, in fact, redistribute before the penultimate period, which implies that they will fight a total war.

PROPOSITION 1 (Total War). *If the conditions of Lemma 1 and Lemma 2 are satisfied, then players never redistribute resources and fight until the weaker one collapses from exhaustion.*

This result gives us the first cut at a solution to the commitment problem: war resolves it by eliminating one of the opponents militarily. As Stalin famously (reportedly) quipped, "Death solves all problems—no man, no problem." In this total war equilibrium, the war may end short of one player getting exhausted if some player collapses in a battle. In this, the result is similar to the long civil wars one can observe in the equilibrium of Fearon's (2004) model, where this outcome is likewise probabilistic. From our bargaining perspective, this solution does not help answer the puzzle because players do not *choose* to end the war but are rather forced to by military exigencies. A nontrivial explanation must involve them choosing actions such that they fight and then settle on the equilibrium path, all with complete information.

4 Limited War

Limited war occurs when players fight for some length of time and then settle on a negotiated redistribution. This is the most interesting case because it involves inefficient use of power under complete information: the opponents waste resources and then mange to negotiate the peace even though they agree in expectation on how the war will evolve from the very beginning. The results in this section demonstrate that peace is crucially dependent on the ability to threaten war, and in particular, the ability to threaten to impose sufficient costs by prolonging the fight. The problem actors face when negotiating peace is that these threats may not be credible: when it is common knowledge that negotiations will be available in the future, and actors will be tempted to reach a peaceful agreement then, the incentive to prolong the war in the future is undermined, and hence the threat to impose costs today becomes unbelievable.

We now proceed in several steps. First, we show that for peace to occur, actors must credibly threaten to fight if it is violated. This suggests that the most permissive conditions for peace are those where players can credibly make the most deterrent threats, that is, threats that would impose the highest costs on the opponent. Since the most deterrent threats are those where players fight until the bitter end (when one of them collapses from exhaustion), the second step is to derive the equivalent to Powell's (2004*b*) sufficiency condition for war: if peace cannot be attained when even the most deterrent threats are credible, then peace cannot be attained in any SPE. We demonstrate that fighting can lead to violation of this condition, opening up the road to peace. The next step is to show that peace can be attained once this condition fails. We construct an SPE in which peace can be sustained by threats to fight to the finish. We then demonstrate that even though these threats are subgame-perfect, they are not credible because players have incentives to renegotiate rather than fight a total war. We then construct an SPE that is immune to such renegotiations and investigate the conditions under which peace can be sustained.

4.1 Strongest Deterrent Threats

We begin with the following lemma, which makes it very clear that peace is sustained by the threat of fighting off the equilibrium path.

LEMMA 3. *Suppose* (x_1, x_2) *is a peaceful bargain when resources are* $(k_1, k_2; S)$ *with* $S = v + k_1 + k_2$*. Then there is no peaceful bargain when resources are* $(x_1 - c_1, x_2 - c_2; S - C)$ *and* S *is large enough.*

Lemma 3 shows that if players are able to conclude a bargain that is peaceful in equilibrium and one of them tries to deviate and attack after the redistribution of resources according to that agreement, then the next period surely involves fighting as well. The longer the fighting one can threaten with, the higher the expected costs for both players, and hence the less each would be willing to accept at the negotiating table, and the better the prospects for peace. The most deterrent threat is to fight to the bitter end. The minmax strategies to fight to the end do form a Nash equilibrium. Furthermore, these strategies can also be subgame-perfect depending on the bargaining protocol.[7] In other words, *if we*

[7]For example, suppose B requires players to submit simultaneous demands, with agreement obtaining if,

allow players to commit to playing the Nash/SPE equilibrium that involves fighting to the end, then we are allowing them to make the most severe threats possible, which creates the strongest incentive to make peace today. If for some period peace is impossible with the commitment to fight to the end should they fail to reach an agreement, peace certainly will not be possible in any subgame perfect equilibrium either.

We now turn to investigation of the conditions for war provided players can commit to their most deterrent threats. Suppose that given the current stocks $(k_1, k_2; S)$, the war can last at most T more periods. That is, if both players fight in each of the following periods without reallocating, then at least one of them will collapse after $T = \min\{T_1, T_2\}$ fights, where $T_i = \lfloor k_i / c_i \rfloor$.

Denote the current period as the first and let $F_i(t|T)$ denote player i's expected payoff from rejecting all offers and fighting to the end starting in period t and fighting up to, and including, period T. For example, suppose that one of the players would collapse after $T = 7$ battles (so if they start fighting now, he would collapse in period $t = T + 1 = 8$). Then $F_1(3|7)$ would denote player 1's expected payoff in period 3 from fighting to this end (that is, fighting five more battles). When $T = T_i < T_j$, if players never reallocate and fight in each period, eventually player i will collapse in period $T + 1$. Hence, $F_i(T|T) = p_i(v + k_1 + k_2 - TC) = p_i(v + k_j - Tc_j)$ and $F_j(T|T) = (1 - p_i)(v + k_j - Tc_j)$. For $t \in \{1, 2, \ldots, T - 1\}$, define the following recursive equation:

$$\begin{aligned} F_i(t|T) &= p_i(S - tC) + (1 - p)F_i(t + 1) \\ &= p_i \sum_{n=0}^{T-t-1} (1 - p)^n \big[S - (n + t)C\big] + (1 - p)^{T-t} F_i(T|T). \end{aligned} \tag{1}$$

This is player i's expected payoff in period t if both players fight without redistribution until the weaker player collapses. It is also player i's reservation value: the payoff that can be unilaterally guaranteed. The lower this payoff for player i, the more deterrent j's threat. Since fighting is inefficient, $\sum_i F_i(1|T) < S$, so bargains that improve on the minmax payoffs always exist. However, as we shall now see, this is not enough for peace to occur. Since we want to find the most permissive condition for peace, we need to derive the most deterrent threats for both players. To do this, we show that the joint expected payoff from fighting until one player collapses from exhaustion is strictly decreasing in the number of potential fights:

LEMMA 4. *When resources are redistributed to increase the number of potential fights, the sum of the players' payoffs from fighting to the end decreases.*

In other words, if the distribution (x_1, x_2) enables players to fight T periods at most, and the distribution (y_1, y_2) enables them to fight one more period, then the sum of their expected fighting payoffs under (y_1, y_2) is strictly worse, and hence (y_1, y_2) involves more deterrent threats and is more conducive to peace. As the proof shows, the joint loss if

---

and only if, they sum up to $S(t)$. Then, the strategies "demand twice the entire surplus and fight regardless of outcome in the bargaining round" are subgame-perfect: given the strategy of j, agreement can never be reached regardless of the proposal i makes, and given j's unconditional attack, i's best response is to fight as well. We thank Bob Powell for suggesting this.

peace fails is precisely $(1-p)^{T-t+1}C$, which is the cost of the additional battle times the probability of having to fight it.

The result in Lemma 4 is intuitive: the longer players expect the war to last, the more resources they expect to waste waging it. Hence, their joint payoffs must necessarily decrease in expected duration. The most deterrent threat a player can make is to fight to the end. Given total resources S, the maximum number of battles players can fight after some distribution is $\overline{T} = S/C$. The largest number of battles under $(x_1, x_2; S)$ is $T_1 = x_1/c_1 = x_2/c_2 = T_2 \Rightarrow x_i = c_i S/C$, where we used $x_1 + x_2 = S$. Therefore, if players equalize resources, they can make the most mutually deterrent threats possible, and the resulting environment is most conducive to peace.

4.2 The Sufficient Condition for War

Simply offering a player his minmax payoff is not enough to induce him to agree to peace. Peace requires that *both players forego the advantages of surprise attack*. To see this, suppose a player accepted a division that exactly matched his minmax payoff, $x_i = F_i(1|T)$, and suppose in equilibrium players achieve peace immediately. Since they achieve peace, j's strategy after this distribution must be not to attack. But then surprise attack yields i at least $p(S-C)+(1-p)F_i(2|T) > p_i(S-C)+(1-p)F_i(2|T) = x_i$. Therefore, attacking is a best response for i, which contradicts the assumption that x_i is accepted in a peaceful equilibrium. Because a peaceful bargain must deter surprise attacks, it must exceed the minmax payoffs.

To see the minimum demands that players would make, suppose they have divided everything such that $(x_1, x_2; S)$ with $x_1 + x_2 = S = k_1 + k_2 + v$. Let $T = \min\{x_1/c_1, x_2/c_2\}$ denote the largest number of battles they can fight under the new distribution without reallocation until one of them collapses from exhaustion. Player i's payoff from sneak attack after the negotiated division is *at least* $A_i(x_1, x_2) = p(x_1+x_2-C)+(1-p)F_i(2|T)$. As we have established, peace requires that surprise attacks are not profitable. Therefore, fighting a battle is going to be unavoidable if there are not enough resources to satisfy minimum deterrent demands. That is, players are sure to fight if:

$$A_1(x_1, x_2) + A_2(x_1, x_2) > S \quad \text{for all feasible } (x_1, x_2; S). \tag{2}$$

This is the logic of Powell's (2004*b*) sufficiency condition which guarantees that complete-information bargaining will break down in any stochastic game, a general category that encompasses our model. To see that this is the case, note that F_i denotes player i's minmax payoff in any period t because i can always guarantee himself this payoff by rejecting all offers and fighting in each period. Further, since $A_i = p_j(S-C) + F_i(1|T)$, $p_j(S-C)$ reflects the increase in i's payoff if he catches his opponent by surprise and results from the temporary "power shift" in i's favor whenever his opponent is expected not to fight. Hence, in any peaceful equilibrium, i must obtain at least A_i, leaving at most $S - A_i$ to meet the minimal demand of the other player. In a peaceful equilibrium both players must have their minimal demands satisfied, yielding the condition in (2).

To ensure that peace will not be possible, we have to establish that no distribution can violate (2). By Lemma 4, $\sum_i F_i(2|T)$ is minimized by taking the largest number of battles, which implies that $\sum_i A_i \geq 2p(S-C) + (1-p)\sum_i F_i(2|\overline{T})$. Since the right-hand side of this inequality is the *worst* players can jointly expect, if this amount exceeds the available surplus, then fighting is guaranteed. In other words, we obtain a sufficient condition of war in the current period: $2p(S-C)+(1-p)\sum_i F_i(2|\overline{T}) > S$, which reduces to:

$$p^2\left(\overline{T}-1\right)+(1-p)^{\overline{T}} > 1. \tag{P}$$

Condition (P) is sufficient to guarantee that peace will not be possible in period $t=1$. It is analogous to Powell's (2004*b*), and we have emphasized this with our choice of nomenclature.

4.3 Peace With Threats of Total War

Whereas condition (P) shows that limited war is possible in principle, it does not prove that it can happen. With a sufficient condition for fighting, we only have a *necessary condition* for peace in its converse. Although we know that this condition for peace can be achieved through fighting, we do not know whether players will be able to commit to peace once it is satisfied. Perhaps whenever they fight in equilibrium, they always end up in a total war? If this is the case, then our arguments do not take us very far. Therefore, it is imperative to demonstrate that limited war can happen in equilibrium. That is, that players fight and then settle, all with complete information.

Condition (P) is defined entirely in terms of the fixed exogenous parameters and the total resources available at the beginning of the period. This means that we can apply this condition to each period of the game by taking $S_t = S-(t-1)C$ to be the surplus in period t, and $\overline{T}_t = S_t/C$ to be the maximum number of battles that can be fought until some player collapses from this period on. The following lemma shows that fighting can lead to violation of that condition.

LEMMA 5. *Condition* (P) *is satisfied if, and only if,* $\overline{T}$ *is sufficiently high.*

Since $\overline{T}$ depends on the amount of available resources, $S(t)$, which is decreasing as fighting continues, Lemma 5 states that even if war is certain under the initial resource distribution, the squandering of resources fighting entails will eventually make settlement possible. In other words, destruction opens up the road to peace.

We now wish to see whether this implies that players can, in fact, achieve peace in SPE provided (P) fails. Since our model allows players to bargain over the protocol they use in addition to the distribution they achieve, all we have to do is show that peace can be sustained in SPE for some protocol that makes threats of total war upon deviation subgame-perfect. As we noted before, these threats can be SPE, and we now use this result to support a limited war equilibrium.

PROPOSITION 2. *Players can achieve peace with threats of total war in SPE in any period in which condition* (P) *is not satisfied.*

This result gives us a first cut at a real solution to the commitment problem through fighting. Suppose the initial distribution of resources is such that condition (P) is satisfied. Players are guaranteed to fight and by Lemma 5 the destruction of resources eventually leads to violation of that condition. By Proposition 2, players then end the war immediately. This peace is sustained by threats to fight to the end if any player violates it with a sneak attack. Whereas such a threat is not capable enough to deter players while the resource base is very large, it does become sufficient as fighting shrinks that base. This happens because the power shift resulting from a sneak attack gets smaller and hence the temptation to benefit from a deviation decreases. Since the original commitment problem depends on large, rapid shifts of power, this result shows that *fighting resolves the commitment problem by reducing the size of the power shift*.

It is natural to ask whether the mechanism through which this is achieved is reasonable. In some sense, it appears to be: after all, the threats are subgame-perfect, and hence no actor has an incentive to change strategy unilaterally. However, this is not enough to make them credible in a more intuitive sense.

4.4 Credibility and Renegotiation Incentives

The result in Proposition 2 depends on finding an SPE in which fighting to the end is subgame-perfect. The key to constructing such an equilibrium is the fact that players make their attack decisions simultaneously and in effect cannot reciprocate unexpected peace feelers. To see what this means, suppose that the bargain protocol is such that players redistribute resources such that one of them, say player 1, obtains a share that is strictly better than what he is getting in the SPE they are playing but player 2's strategy is to fight at such an allocation. Suppose further that given player 2's strategy, player 1's best response is to attack as well. Since the bargaining outcome represents a deviation, players fight and the game continues. Now suppose that player 2 deviates and does not attack. Player 1's optimal course of action would be to reciprocate because peace then prevails and his payoff under the new allocation is strictly better. If he could condition his attack decision on what player 2 does, then subgame-perfection would indicate that fighting is not credible. However, the extensive form does not allow the player to condition in that way and since he is forced to make the attack decision "in the dark," subgame-perfection does not have a bite.

On one hand, this may appear to be an artifact of the extensive form which we can alleviate by making fighting decisions sequential. On the other hand, the essence of a sneak attack seems to be that an actor is making the relevant decision in the dark. The modeling choice seems appropriate. However, there is still something unsatisfying about the threats that support the limited war SPE in Proposition 2.

Observe that the strategy to deter deviations requires punishments that hurt both players. In the example SPE, the punishment is extreme: total war. Suppose now that the players find themselves off the equilibrium path because someone has deviated and they are doomed by their SPE strategies to fight to the bitter end. The question then arises: given that players have found themselves in an inefficient situation would they renegotiate to get out of it? That is, if there exists an SPE such that each player gets at least his total war payoff and at least one gets a strictly better payoff, it is reasonable to expect that players will then renegotiate their original "agreement" that has now led them to this mutually hurtful situation.

Out model allows players to negotiate the bargaining protocol, which implies they can renegotiate the equilibrium they play. For example, given a bargaining protocol, they may look for an alternative SPE that Pareto-dominates the one induced by the original strategies.

Alternatively, they may look for a different bargaining protocol that allows them to achieve another such SPE. In any case, we should expect players to renegotiate their original positions and switch to the new equilibrium. But if such alternatives exist, then the credibility of the original SPE is undermined and the threat becomes incredible: there is no reason players should expect to fight to the end if they can always go back to the negotiation table and fix their predicament. We should look for SPE that are immune to such renegotiations because these are the only ones that actors can reasonably expect to play. Hence, we focus on SPE that are *renegotiation-proof* in the sense of Farrell and Maskin (1989) as the only plausible candidates for solutions to our model.

Formally, a bargaining protocol and an induced SPE pair (B, σ) is *renegotiable* if there exists a subgame, a bargaining protocol B' starting at that subgame, and an SPE σ' induced by B' in that subgame that is weakly preferred by each player and strictly preferred by at least one player to the SPE induced by (B, σ) in that subgame. We say that a bargaining protocol and an induced SPE pair, (B, σ), is renegotiation-proof if it is not renegotiable.

Given that players should be expected to search for a better SPE if they ever find themselves in an inefficient equilibrium and bearing in mind the fact that peace requires deterrent threats, the strongest threats they can make and still be believed are the ones that can be sustained in a renegotiation-proof SPE (RPSPE). We refer to threats that can be sustained in a RPSPE as *credible* keeping in mind that such a notion of credibility is more demanding than that required by subgame-perfection.

4.5 Peace With Credible Threats

Intuitively, the strongest credible threat is the one that involves the most periods of fighting provided players have no incentives to renegotiate during any of these periods. In other words, suppose players can achieve peace in some period T. Then, the strongest credible threat they can make in $T-1$ is to fight one battle. If they have no incentive to renegotiate in that period either, the strongest credible threat they can make in $T-2$ is to fight two battles, and so on. This immediately suggests that threats of total war may be incredible, which in turn means that condition (P) may not be able to pin down whether peace can be obtained in a limited war RPSPE. If players cannot commit to total war but to some (much smaller) number of battles, the total costs they can credibly impose on each other are also smaller. Since peace depends on the severity of punishment of sneak attacks, if this punishment is not that costly, the deterrent effect is so much weaker, and prospects for peace so much gloomier. The question then becomes: can players achieve peace if they can use only credible threats?

Letting $V_i(x_1, x_2; S)$ denote player i's RPSPE payoff in the continuation game given a feasible distribution $(x_1, x_2; S)$, the necessary and sufficient condition for fighting at this distribution is $\widehat{A}_i(x_1, x_2) = p(S-C) + (1-p)V_i(x_1 - c_1, x_2 - c_2; S - C) > x_i$ for some i. Hence, the sufficient condition for fighting using only credible threats is:

$$\sum_i \widehat{A}_i(x_1, x_2) = 2p(S-C) + \sum_i (1-p)V_i(x_1 - c_1, x_2 - c_2; S - C) > S. \qquad \text{(W)}$$

Because $\sum V_i(x_1 - c_1, x_2 - c_2; S - C) \geq 0$ and $S > C$, it follows that for p and S large enough, (W) will be satisfied. Bargaining in this period will break down and at least one battle will be guaranteed in any RPSPE. However, whereas (P) implies (W), there will be instances where (W) is satisfied but (P) is not. In other words, if condition (P) fails, we are not guaranteed that peace will occur because (W) may still be satisfied. If this is the case, then (P) does not provide a compelling resolution to the commitment puzzle because threats of total war are not credible. If, however, we can find an equilibrium in which fighting leads to violation of (W), then we do have such a mechanism.

We now use condition (W) to construct an example SPE that uses only credible threats and that demonstrates the main theoretical results that we use for our substantive discussion. Assume that players are *symmetric*, that is, $K_1 = K_2 = K$, $c_1 = c_2 = c$, and $p_1 = p_2 = p/2$. Suppose the game begins with $v > 0$. After T battles without redistributing, each player has $k_i = K - Tc$ resources left, so the total is $S = 2k_i + v$. If they distribute now, players face a situation with surplus S and $v = 0$, and therefore in any SPE the relevant behavior is that in these continuation games. Consequently, we now explore games with symmetric players and $v = 0$. We first show that in this setup, subgames with equalization of resources are particularly helpful because they capture all relevant aspects and are easy to analyze. We then use these results to construct our numerical example.

The following lemma proves that in games with symmetric players who have equal resources, we do not lose any generality by restricting analysis to subgames in which players do not redistribute along the equilibrium path. Using this result, the lemma further derives the necessary and sufficient condition for fighting in any arbitrary period.

LEMMA 6. *Assume $v = 0$ and symmetric players with $k_i = nc$, where $n \geq 1$. Then, without loss of generality, the players do not redistribute in equilibrium. Suppose that they find a peaceful settlement when $k_i = nc$. Then:*

(A) they fight with $k_i = (n + 1)c$ if, and only if, $pn > 1$;

(B) if they fight with $k_i = (n + t)c$, $t = 1, \ldots, T$, then they fight with $k_i = (n + T + 1)c$ if, and only if,

$$p^2(n + T) + (1 - p)^{T+1} > 1. \tag{3}$$

Observe now that by Lemma 6, if players are symmetric and have equal resources, there is no loss of generality if we consider only SPE where players do not redistribute. This now means that if they cannot achieve peace in equilibrium with (nc, nc), they cannot achieve it under any alternative allocation. This yields the following helpful result:

COROLLARY 1. *Suppose players are symmetric and cannot achieve peace if they redistribute such that $k_i = nc$ and $S = 2k_i$. Then they cannot achieve peace under any alternative distribution.*

This is a powerful result: for any distribution (x_1, x_2) with $v = 0$ and $S = x_1 + x_2$, we only need to check if players can achieve peace by equalizing their shares. That is, we use the conditions in Lemma 6 to check if players can achieve peace had they distributed $(S/2, S/2)$. This saves us a lot of work because otherwise we had to compute continuation values for any subgame with shares $x_i - c$, rather than just with $S/2 - c$. But since in any period with $v > 0$, the credibility of peace will depend on what happens after they redistribute, this result provides the key to unraveling the SPE in the entire game. The algorithm we use is to take any period, suppose they equalize resources, and check if they will still fight using Lemma 6. If they do, then no alternative distribution can produce peace (by Corollary 1), and we know a battle is inevitable in any SPE. If they do not, then they will certainly achieve peace in this period (because we have identified at least one distribution that can do it and because we assumed that players will be able to utilize the opportunity).

PROPOSITION 3 (Limited War). *Assume* $v = 12$ *and* $p = .18$*, and symmetric players with* $K_i = 30$*,* $c_i = 1$*, and* $p_i = p/2$*. The players fight 12 battles without distributing resources and achieve peace in* $t = 13$ *provided neither collapses in the interim.*

Since the construction is illustrative, we provide the proof here. Letting S denote the surplus in period t and noting that since players do not redistribute in any period in which they fight, we can use the time index to denote the continuation values, and so $V_i(t)$ is player i's expected equilibrium payoff in period t. Condition (W) then is $S < .36(S - 2) + .82 \sum_i V_i(t+1)$. If (W) fails in some period t, then $\sum_i V_i(t) = S$ because the entire surplus is peacefully distributed by our requirement that fighting cannot continue beyond periods in which peace is possible under some bargaining protocol and induced SPE. This effectively limits the duration of war to which players can credibly commit to that period. If, on the other hand, (W) is satisfied, then peace is impossible in this period and $\sum_i V_i(t) = p(S - C) + (1 - p) \sum_i V_i(t+1)$.

Recall from the proof of Proposition 1 that if players are sure to fight in some period t, they cannot improve matters by redistributing, and therefore there is no strict incentive to do so in equilibrium. In other words, we can restrict attention to SPE where players do not redistribute in any period in which they expect to fight for sure. This now allows us to backward-induct along the no-distribution path using condition (W) and Lemma 6.

Consider now the no-redistribution path at $t = 30$, where players have fought 29 battles already, and so $S = 14$. If they distribute and equalize resources now, they could fight $n = 7$ more battles. By part (A) of Lemma 6, players will not fight with any $n \leq 6$ but will fight with $n = 7$ (because not fighting with $n = 6$ results in $p = .18 > .17 = 1/n$ in the current period). Therefore, players will surely fight at $t = 30$ and settle in the next period (by collapsing simultaneously if they do not redistribute now or by redistributing if they do), and so $\sum_i V_i(30) = p(14 - 2) + (1 - p)(12) = 12$.

Going to the previous period along the no-distribution path, we have $t = 29$ with $S = 16$. Equalization would yield $n = 8$, and since players fight with $n = 7$ but settle with $n = 6$, part (B) of Lemma 6 applies with $T = 1$ and $n = 6$. Solving (3) shows that players can achieve peace in this period. Therefore, $\sum_i V_i(29) = 16$. Continuing in this way, we construct Table 1, which shows the SPE outcomes for all periods of the game along the no-distribution path.

Observe that (P) is satisfied for all $t \leq 5$. This implies that peace is impossible in the first four periods even according to the stricter criterion. However, even though (P) fails for all $t > 5$, there are many periods where peace cannot be achieved because (W) still holds. If, for example, this game started out with $K_i = 25$, then (P) would have no bite at all but we will still get 7 battles in equilibrium. This illustrates our claim that with its reliance on incredible threats, Proposition 2 does not provide a compelling reason to expect players to achieve peace, and therefore is not a persuasive solution of the commitment problem.

t	$(k_1, k_2; S)$	$\sum_i \widehat{A}_i$	Lemma 6	(P)	outcome	$\sum_i V_i(t)$
1	(30, 30; 72)	74.516	1.226	1.135	fight	61.92
2	(29, 29; 70)	72.381	1.214	1.103	fight	60.14
3	(28, 28; 68)	70.296	1.207	1.070	fight	58.42
4	(27, 27; 66)	68.271	1.204	1.038	fight	56.75
5	(26, 26; 64)	66.320	1.209	1.006	fight	55.16
6	(25, 25; 62)	64.459	1.221	0.974	fight	53.66
7	(24, 24; 60)	62.707	1.244	0.942	fight	52.27
8	(23, 23; 58)	61.088	1.278	0.910	fight	51.01
9	(22, 22; 56)	59.632	1.327	0.879	fight	49.91
10	(21, 21; 54)	58.375	1.394	0.847	fight	49.02
11	(20, 20; 52)	57.360	1.482	0.816	fight	48.36
12	(19, 19; 50)	56.640	4.320	0.785	fight	48.00
13	(18, 18; 48)	47.939	0.994	0.754	**peace**	48.00
14	(17, 17; 46)	46.187	1.017	0.723	fight	38.27
15	(16, 16; 44)	44.568	1.051	0.693	fight	37.01
16	(15, 15; 42)	43.112	1.100	0.663	fight	35.91
17	(14, 14; 40)	41.855	1.167	0.634	fight	35.02
18	(13, 13; 38)	40.840	1.256	0.606	fight	34.36
19	(12, 12; 36)	40.120	3.060	0.579	fight	34.00
20	(11, 11; 34)	33.672	0.971	0.553	**peace**	34.00
21	(10, 10; 32)	32.415	1.037	0.528	fight	27.02
22	(9, 9; 30)	31.400	1.126	0.505	fight	26.36
23	(8, 8; 28)	30.680	2.340	0.483	fight	26.00
24	(7, 7; 26)	25.335	0.940	0.465	**peace**	26.00
25	(6, 6; 24)	24.320	1.029	0.449	fight	20.36
26	(5, 5; 22)	23.600	1.800	0.437	fight	20.00
27	(4, 4; 20)	19.600	0.964	0.429	**peace**	20.00
28	(3, 3; 18)	18.880	1.440	0.427	fight	16.00
29	(2, 2; 16)	14.880	0.899	0.431	**peace**	16.00
30	(1, 1; 14)	14.160	1.080	0.444	fight	12.00

Table 1: Limited War SPE (No-distribution Path).

5 Discussion

We now turn to the substantive implications from the analysis of the Limited War SPE in Proposition 3. One very general result is a vindication for Vegetius' dictum "if you want peace, prepare for war." As Lemma 3 shows, a peaceful settlement must be sustained by the threat to punish attempts to exploit it; that is, peace necessarily involves a credible deterrent to surprise attacks. This is in keeping with results in Powell (1993) and Slantchev (2005) who show that peace may require substantial military investments to produce such mutual deterrence. That the notion of the *armed peace* has emerged in three rather different stylized conflict environments and that the logic behind it is essentially equivalent across them signify that it may be a very general phenomenon.

5.1 How War Resolves the Commitment Problem

Recall Fearon's (1995, 402-4) discussion of how first-strike advantages can close the bargaining range and cause war. As Powell (2006) has shown, the mechanism that causes inefficiency in that model is equivalent to the general commitment problem resulting from large rapid shifts of power. Intuitively, foregoing the advantages of a first strike produces a power shift in favor of one's opponent. The "declining" actor needs to be compensated for not striking first but the "rising" actor cannot credibly promise to deliver the rest of the compensation tomorrow when he finds himself in a strong position. As a result, the declining actor wages preventive war today. The logic of the general power shift mechanism implicitly relies on a stylized representation of war that abstracts from the effect of fighting on the shift itself. If states go to war because of a commitment problem caused by first-strike advantage, then how do they agree to end the war while this technology is still present?

Observe that the power shift reflects the change in continuation payoffs, and these depend not just on the first-strike advantage that comes from surprise but also on the subsequent behavior of the actors. Whereas the advantage expressed as an increased probability of military victory in a battle remains constant throughout, the continuation payoffs change as the resources shrink with the duration of fighting. In other words, the *size* of the power shift varies even though the technology that makes it possible does not. Hence, *in general the solution to the commitment problem involves playing strategies that minimize the power shift to the point where the incentives for surprise attack disappear.*

The question then becomes: under what conditions would players choose strategies that undermine these incentives? Condition (P) shows that there are situations in which no such strategies exist, and so players cannot avoid fighting. However, as Proposition 2 makes clear, the resulting destruction eventually opens up the road to peace and players can achieve a peaceful settlement provided they can threaten to punish attempts to undermine it by total war. We argued, however, that such threats are not credible because they commit players to a painful equilibrium even though both have incentives to renegotiate and find a better one. Since our model provides opportunities for such renegotiations, it is reasonable to expect players to utilize them.

Unfortunately, this ability to find mutually better solutions reduces the severity of the threats players can credibly make, which in turn undermines the deterrent effect that is supposed to maintain the peace settlement. If players expect to be able to renegotiate and end the war as soon as possible, then they cannot threaten to fight to the bitter end. At best, they can threaten to prolong fighting until the next such opportunity presents itself. This now implies that they may not be able to reduce the size of the power shift sufficiently to avoid fighting. In other words, *war is caused by the inability of players to commit credibly to punish a sneak attack with sufficient severity to deter it, and this inability arises from their incentives to seek peace at first opportunity.*

The example in Table 1 helps follow the logic. At $t = 1$, players could fight 30 battles without redistributing. Unfortunately, even threatening to fight all of them until they collapse from exhaustion cannot prevent fighting in the first five periods where condition (P) holds. A rather dark implication of this analysis is that there may exist situations where even a credible threat to fight to the end may not avert war. Sometimes the stakes can be so high that neither player can impose enough costs on its opponent to deter him from risking a few battles to win them.

From $t = 6$ onward, however, players could achieve peace in any period if they only could threaten to fight to the bitter end. These threats are not credible because players know that if the war does not end by $t = 13$, they will renegotiate and achieve peace then. To wit, surprise attack at, say $t = 7$, does not risk a total war but a limited one, and both players know this. Because of this, neither player can impose sufficient costs on the opponent to deter him from sneak attack, and the incentives to strike a bargain in all prior periods dissipate. Players are not credibly prepared for war, and therefore cannot obtain peace.

The difference between the two conditions is intuitive: whereas (P) uses the largest number of battles, (3) only uses the number that players are actually expected to fight *in an equilibrium with credible threats*. In the latter, $T+1$ is the largest number of battles players expect to fight until they can achieve peace. As such, it is equivalent to $\overline{T}$ in (P). Hence, the crucial difference between the two conditions is in the n term, which one can interpret as the number of fights players would have been able to fight if they could commit credibly to do so. Since they cannot, increasing this term leads to (3) being satisfied even under conditions where (P) would fail. The larger this discrepancy, the worse the prospects for peace.

This highlights our main conclusion: the ability to achieve peace critically depends on the credibility of the mutually deterrent threats that players can make. Even if (P) could be satisfied with Nash/SPE threats, players have no reason to believe them but will instead only take into account how many battles the opponent is actually prepared to fight before renegotiating. As we have seen, this unhappy calculation can undermine the incentives for peace today. Conversely, if one can credibly threaten to punish an opportunistic move by imposing very large costs on the attacker, then peace can be sustained.

How does fighting resolve this war commitment problem then? Unhappily, it does so by war's very nature, its sheer destructiveness. Initially, both players are rich in resources and the stakes are high. What are a few battles compared to the possibility of obtaining these riches should a military operation prove successful? As war progresses, the pie shrinks and continuation becomes less and less tempting, which in turn means that it takes weaker threats to deter participants from surprise aggression. Eventually, players can credibly commit to fight a handful of battles and this minimize the power shift to the point that peace can be maintained. Every war carries the seeds of its own peace.

5.2 War Costs and Rational Escalation

Can two players rationally escalate war and end up agreeing to peace terms that are worse than what each originally started the war with? That is, can rational players pay war costs that exceed the value of the prize?

In the Dollar Auction game (Shubik 1971), two players alternate in bidding for a prize of one dollar. The highest bidder wins the prize but both have to pay their bids. O'Neill (1986) analyzes a discrete complete-information version of the dollar auction with budget constraints. He finds a unique subgame perfect equilibrium in which no escalation occurs—player 1's initial bid forces player 2 out of the game.[8] This result differs markedly from experimental studies of the Dollar Auction in which players often escalate, sometimes bidding more than the value of the prize itself (Teger 1980). O'Neill concludes that such escalatory behavior must be irrational.

The Dollar Auction has been used extensively to shoehorn interstate crises and wars, among other events, into its interpretive framework of loss-avoidance, a case of "we have invested too much to quit now." For many analysts, the game provides an especially apt analogy that illustrates the pernicious consequences of various forms of irrationality. They see escalation as pathological, and the policy prescriptions derived usually take the form of a wishlist: if only players could foresee..., if only they knew..., if only they could admit their mistakes....[9] The general implication is clear: if only players were fully rational and knew everything about each other, then we would never see "senseless" escalation that ends with both of them paying more than the prize itself is worth.

This conclusion is implicitly endorsed by a great many formal models of conflict which assume that the costs of war are less than the value of the prize. In fact, even the canonical bargaining model of war is forced to make this assumption if conflict is ever to occur with positive probability in equilibrium. After all, if the costs of war are expected to exceed the value of victory, then a rational player would never go to war (Bueno de Mesquita 1981).[10] In other words, most of our rationalist theories of war tacitly agree with the view that escalation of the sort that happens in experimental plays of the Dollar Auction is inherently irrational.

Leaving aside the question of whether the Dollar Auction is actually a good model of crises and wars (after all, it admits no bargaining and no negotiated outcomes), we argue that excessive escalation can happen with fully informed rational players even when they are virtually unconstrained by the bargaining protocol. While we do not mean to discount psychological explanations, we want to emphasize that there is no need to resort to irrationality to understand situations in which both players end up paying more than the prize itself is worth.

When we replace the heroic assumption that war costs are low relative to the value of the stakes with the milder one that the stakes are worth at least one battle, the tragedy of war reveals itself: *If players survive to make a peaceful settlement, war is sure to have become a net loss for both!* When players settle at $t = 13$ in the Limited War SPE, the amount to be divided is $S = 48$, whereas it started out at 72. They have collectively paid war costs of 24. Observe now that *at the time of peace,* players accept shares that leave them worse off relative to just conceding the prize from the outset. For example, if players share equally at $t = 13$ (not unlikely given their symmetry), each would obtain a payoff of 24. This is worse than immediate concession at the outset and it is less than the initial resources the player had. Each has paid war costs of 12 for the dubious privilege of obtaining a benefit of 6. Furthermore, since (P) would hold if we increase the size of initial stocks, richer players would fight longer wars. For example, letting $K = 35$ results in five additional battles, and

[8] See Leininger (1989) for a more comprehensive analysis.

[9] See, for example, Poundstone (1992, 261-65) and Teger (1980) for some examples. Deutsch's (1973, 356-57) analysis of U.S. involvement in Vietnam essentially follows the same logic.

[10] We thank Ken Schultz for drawing our attention to this shortcoming of existing models.

total individual costs of 17 at the time of negotiated peace.

War becomes unprofitable very quickly: from $t = 7$ on, at least one of the players cannot recover the resources he started the war with. In our example, negotiated peace can occur only after war has lasted nearly twice the duration that could be potentially profitable for one of the players. However, the piling costs do not deter players from fighting because they are sunk, and hence all that matters are the expected future gains and losses. It is that forward-looking aspect of war that may make it a dead loss to both players if they fight into peace. Pillar (1983, 173) argues that it can be rational to continue fighting even after the war "has already escalated well out of proportion to the value of the objectives at stake" because one cannot manipulate past costs, only future ones. The model vindicates this logic even in an environment where all cost manipulation is only implicit in the threat to continue to fight. It also provides rationalist theoretical foundations the empirical findings by Orme (2004) and what he terms the "paradox of peace": peace is most likely when the threat of costly conflict is greatest.

These results show that it is rational to risk small escalatory steps that eventually may accumulate enough costs to exceed the value of the issue at stake. This highlights the problematic assumption in traditional models of war and suggests that we may need to rethink some of the causal mechanisms derived from such theories.

5.3 War as Punctuated Equilibrium

Our analysis suggests that one can usefully view war as a mutually coercive process that involves continuous fighting punctured by occasional opportunities for peace. Even though peace negotiations are available throughout the war, a credible commitment to a settlement is only possible at specific junctures. There are specific *windows of opportunity* to end the conflict, and if such a window closes, players are stuck fighting until the next one comes along. In a sense, it seems true that conflicts have to be "ripe for resolution" (Zartman 1985). However, in our formulation ripeness is not a battlefield property of conflict that appears when players reach a "mutually hurting stalemate" (Zartman and Berman 1982). Rather, it is a function of the credibility of threats to punish attempts to take advantage of the peace negotiations. If players fail to cease the fleeting opportunity to end the war, they will be condemned to fight it out until another window presents itself.

In our example, should players for some reason be unable to negotiate at $t = 13$, they have to carry on the war for seven more periods until $t = 20$ opens up the possibility for peace again. At this puncture in fighting, the terms of peace are much worse for both: players would have jointly paid costs of 40 to divide the prize, and each can hope to live with a little more than a third of his original resources. The desire to avoid this additional fighting and worse outcome in turn induces players to agree to peace at $t = 13$. Observe, however, that nowhere in our model are players stalemated (they can always risk a battle that gives a chance of outright victory) and neither does inability to negotiate a termination of war hinge on problems with perception.

These windows of opportunities are rarer when the stakes are higher. The more resource-rich the warring parties, the longer the fighting spells between these windows. Their frequency, however, increases the longer the war lasts, and their closure gets ever shorter as opponents approach exhaustion. That is, the more weakened the actors are from fighting, the more willing to negotiate they become. As they approach collapse, the terms of peace begin to approximate the expected payoff from continued fighting, obviating the incentive to risk it. As the terms of peace deteriorate, so does the expected payoff from prolonging the war, and hence the prospects for war termination improve. Ironically, the better the expected terms of the settlement, the worse the prospects for immediate peace. This is because the peace settlement itself is a function of the available benefits to be divided and since fighting may secure these benefits completely, the stronger the incentive to risk it.

It is this pattern of windows of opportunities for peace, which cluster toward the military end of war, that leads us to view war as a *punctuated equilibrium*.

6 Conclusion

One of the canonical rationalist explanations of war is that opponents cannot credibly commit themselves to follow through on the terms of agreement because a change in relative power renders such promises against their interests. Actors then may prefer to start a war today rather than face the unpalatable consequences of peace tomorrow. However fundamental and intuitive, this mechanism is incomplete and incoherent because it does not explain how fighting alleviates that commitment problem. We argued that unless we view war as a process that traces the effect of fighting, we will not be able to resolve this puzzle.

The analysis uncovered a subtlety that essentially turns the original commitment problem on its head. In our account, an actor's inability to promise credibly to fight for long lowers the costs of war and causes his opponent to demand so much today that he prefers to continue fighting rather than concede. The credibility problem arises from the opportunities for peace in the future: when both actors know that they want to settle the costly conflict as soon as possible, threats to extend fighting beyond such an opportunity for peace become unbelievable. Actors are tempted to risk some more fighting because they cannot deter each other by threatening not to negotiate in the future. In showing how fighting resolves that problem, we provide a complete and coherent rationalist explanation of war that does not require asymmetrically informed players.

Ironically, the very desirability and possibility of peace make war more likely because they decrease its expected duration and costs. An obvious tactic then suggests itself: if one could conceal such temptation to negotiate and somehow commit not to seek peace until a military resolution of the conflict, the likelihood of being able to negotiate an early termination will increase. Of course, it should also be obvious how difficult it will be to pull such a trick: one must simultaneously demonstrate complete resolve to fight to the bitter end and willingness to negotiate peace. The problem becomes worse in societies where leaders might be constrained by the public to fight short wars: all else equal, democracies may be unable to mount credible threats to fight to the end, and this may embolden their opponents and needlessly prolong the wars they fight. After all, it is not at all clear that democratic leaders cannot mobilize for a long haul in spite of widespread opposition. But then again, neither it is clear that these leaders will persevere against popular opinion for too long.

A Proofs

Proof of Lemma 1. Let $V_i(k_1, k_2; S)$ denote i's equilibrium payoff in the game that begins with total surplus S and capital stocks (k_1, k_2) such that $k_1 + k_2 < S$. Suppose that when the resources are $(k_1, k_2; S)$, players will fight for $T \geq 1$ more periods. Consider now the distribution $(k_1 + c_1, k_2 + c_2; S + C)$ and any feasible (x_1, x_2) such that $x_1 + x_2 = S + C$ and $x_i \geq 0$. If i rejects this distribution and fights this period, then they continue with $(k_1, k_2; S)$ and fight T periods by our supposition. Let $V_i(T+1) = V_i(k_1 - Tc_1, k_2 - Tc_2; S - TC)$ be the short-hand expression for the payoff players expect to obtain in the period after T fights starting tomorrow. Recalling that today's surplus is $S + C$, fighting today and then for another T periods starting tomorrow can be compactly expressed as: $W_i(1|T+1) = p_i S + p_i \sum_{n=1}^{T}(1-p)^n(S - nC) + (1-p)^{T+1} V_i(T+1) = p_i \sum_{n=0}^{T}(1-p)^n [S - nC] + (1-p)^{T+1} V_i(T+1)$. Player i can expect $W_i(1|T+1)$ if he rejects a deal and fights today. Hence, if (x_1, x_2) is a peaceful bargain, it must be the case that:

$$x_i \geq W_i(1|T+1) \text{ for } i = 1, 2. \tag{4}$$

We want to show that for any feasible distribution (x_1, x_2) that satisfies (4), there exists i such that:

$$x_i < pS + (1-p)V_i(x_1 - c_1, x_2 - c_2; S). \tag{5}$$

That is, such a distribution is not peaceful because at least one player expects to do better by deviating and fighting at least one battle. We proceed by contradiction. Suppose that there does exist (x_1, x_2) such that $x_1 + x_2 = S + C$, and (4) is satisfied, but (5) is not. Suppose now that $(x_1 - c_1, x_2 - c_2)$ is a peaceful bargain for $(x_1 - c_1, x_2 - c_2; S)$. In this case, $V_i(x_1 - c_1, x_2 - c_2; S) = x_i - c_i$, which means that $x_i \geq pS + (1-p)(x_i - c_i)$ by our original supposition. We now obtain: $\sum_i x_i = S + C \geq 2pS + (1-p)(x_1 - c_1 + x_2 - c_2) = 2pS + (1-p)S \Rightarrow C \geq pS$, which is violated for v large enough. Therefore, $(x_1 - c_1, x_2 - c_2)$ cannot be a peaceful bargain when resources are $(x_1 - c_1, x_2 - c_2; S)$. That is, it must be the case that for some i,

$$x_i - c_i < p(S - C) + (1-p)V_i(x_1 - 2c_1, x_2 - 2c_2; S - C). \tag{6}$$

We now show that if S and p are large enough, then (6) implies (5), which will establish our claim. Because $(x_1 - c_1, x_2 - c_2)$ yields a fight, it follows that: $V_i(x_1 - c_1, x_2 - c_2; S) = p_i(S - C) + (1-p)V_i(x_1 - 2c_1, x_2 - 2c_2; S - C)$. Combining this with (6) yields: $V_i(x_1 - c_1, x_2 - c_2; S) > p_i(S - C) + x_i - c_i - p(S - C) = x_i - c_i - p_j(S - C)$. Multiplying both sides by $(1-p)$ and adding pS to each produces the equivalent inequality: $pS + (1-p)V_i(x_1 - c_1, x_2 - c_2; S) > pS + (1-p)\left[x_i - c_i - p_j(S - C)\right]$. We now show that $pS + (1-p)\left[x_i - c_i - p_j(S - C)\right] > x_i$, which will complete the proof because it will establish (5) directly. Using some algebra to simplify the last expression, we obtain: $(p_i + pp_j)S + p_j(1-p)C > px_i + (1-p)c_i$. From (4), we have $x_j \geq W_j(1|T+1)$, which implies that $x_i \leq S + C - W_j(1|T+1)$, and so $px_i + (1-p)c_i \leq p\left[S + C - W_j(1|T+1)\right] + (1-p)c_i$. Hence, showing that $(p_i + pp_j)S + p_j(1-p)C \geq p\left[S + C - W_j(1|T+1)\right] + (1-p)c_i$ would be sufficient. Rearranging terms yields the inequality we need to establish:

$$pW_j(1|T+1) \geq p_j(1-p)(S - C) + pC + (1-p)c_i. \tag{7}$$

Simplifying $W_j(1|T+1)$ yields:

$$W_j(1|T+1) = p_j\left(\frac{1-(1-p)^{T+1}}{p}\right)(S-C) + (1-p)^{T+1}V_j(T+1) + p_j\left(\frac{2p-1+(1-p)^{T+1}[p(T-1)+1]}{p^2}\right)C.$$

Comparing the coefficients on $(S-C)$ in (7) produces $p \geq (1-p)^{T+1}$, which holds for any T as long as $p \geq 1/2$. (Note that as T increases, the smallest value of p that satisfies the inequality decreases.) Consequently, (7) is satisfied provided this is the case and S is sufficiently large. $\square$

Proof of Lemma 2. Let the distribution be $(k_1, k_2; S+C)$. Without loss of generality, let $k_1 \in [c_1, 2c_1)$. That is, without a redistribution of resources player 1 has enough to fight just one last battle. Assume that $p_2S > C$, and note that this implies $pS > C$ as well. Assume also that $k_2 \geq 2c_2$, so that player 2 will outlast player 1 if they fight.

We now establish the bounds on player 1's share in any equilibrium bargaining division, and then show that no peaceful division exists within these bounds. Consider some arbitrary division of the surplus (x_1, x_2) such that $x_1 + x_2 = S + C$. If 2 rejects this and fights, her expected payoff is $p_2S + (1-p)S = (1-p_1)S$. Hence, player 2 will not accept any share $x_2 < (1-p_1)S$, which implies $x_1 \leq S + C - (1-p_1)S = p_1S + C$. Similarly, player 1 can guarantee p_1S by rejecting any division and fighting. Hence, $x_1 \in [p_1S, p_1S + C]$ for any feasible equilibrium bargaining division.

Suppose now that a peaceful bargain (x_1, x_2) exists. Then $x_1 \geq pS + (1-p)V_1(x_1 - c_1, x_2 - c_2; S) \geq pS$. Since $x_1 \leq p_1S + C$, it follows that $p_1S + C \geq pS \Leftrightarrow C \geq p_2S$ must hold. This contradicts our assumption and hence no peaceful (x_1, x_2) exists. In this case, player 2 is better off without a redistribution that expands player 1's expected life span, and player 1 will not accept any distribution that decreases it. Hence, after any potential distribution, player 1's life span must remain unchanged, and so we have $(\hat{k}_1, \hat{k}_2)$ with $\hat{k}_1 \in [c_1, 2c_1)$. If this distribution is peaceful, then $\hat{k}_1 \geq pS$ and $\hat{k}_2 \geq pS + (1-p)S = S$. Adding these inequalities yields $S + C \geq (1+p)S \Rightarrow C \geq pS$. This contradicts our assumption, and therefore players fight. The proof for the case where both players are about to collapse simultaneously is analogous. $\square$

Proof of Proposition 1. Consider the extensive form of the game and take the path where players fight in each period without redistributing resources. Its terminal node is the collapse from exhaustion of one of the players, say player 1. Hence, at the penultimate node resources are $(k_1, k_2; S+C)$ such that $k_1 \in [c_1, 2c_1)$ and $k_2 \geq 2c_2$, with $S+C = v+k_1+k_2$. By Lemma 2, players will fight at that node and will not redistribute resources. Consider now the node prior to that, with resources $(k_1+c_1, k_2+c_2; S+2C)$. Since players fight one battle with $(k_1, k_2; S+C)$, Lemma 1 implies that they will fight at that node as well. Note that the lemma does not say that they will not redistribute, only that they will fight even if they redistribute.

We now show that players will not redistribute at that node. For any arbitrary $(k_1, k_2; S)$, a redistribution (x_1, x_2) is feasible if, and only if, $V_i(x_1, x_2; S) \geq V_i(k_1, k_2; S)$ for $i =$

1, 2 because otherwise at least one player will not accept it and will instead continue with $(k_1, k_2; S)$. Now consider some $(k_1, k_2; S)$ that is sure to result in a fight. We claim that if (x_1, x_2) is a feasible bargain for $(k_1, k_2; S)$, then $(x_1 - c_1, x_2 - c_2)$ is a feasible bargain for $(x_1 - c_1, x_2 - c_2; S - C)$ as well, and so players do not have a (strict) incentive to redistribute today. To see this, consider any (x_1, x_2) that is feasible for $(k_1, k_2; S)$ that does not have a peaceful solution. Then $V_i(x_1, x_2; S) = p_i(S - C) + (1 - p)V_i(x_1 - c_1, x_2 - c_2; S - C)$ and $V_i(k_1, k_2; S) = p_i(S - C) + (1 - p)V_i(k_1 - c_1, k_2 - c_2; S - C)$ because neither (k_1, k_2) nor (x_1, x_2) can avoid a fight today. Combining these inequalities with the fact that (x_1, x_2) is feasible yields $V_i(x_1 - c_1, x_2 - c_2; S - C) \geq V_i(k_1 - c_1, k_2 - c_2; S - C)$. That is, $(x_1 - c_1, x_2 - c_2)$ is a feasible bargain for $(k_1 - c_1, k_2 - c_2; S - C)$. Hence, players will not redistribute at $(k_1, k_2; S)$ if they are sure to fight there.

Returning to our argument, players will fight without redistribution at the node with $(k_1 + c_1, k_2 + c_2; S + 2C)$. Repeated application of that lemma to all preceding nodes along this path unravels the game back to the initial node and ensures that players will not attempt redistribution there either. In other words, the only equilibrium outcome is total war. □

Proof of Lemma 3. Seeking a contradiction, suppose $(x_1 - c_1, x_2 - c_2; S - C)$ admits a peaceful bargain. Since (x_1, x_2) is peaceful $x_i \geq p(S - C) + (1 - p)V_i(x_1 - c_1, x_2 - c_2; S - C)$ for $i = 1, 2$. This implies $x_1 + x_2 = S \geq 2p(S - C) + (1 - p)\sum_i V_i(x_1 - c_1, x_2 - c_2; S - C) = 2p(S - C) + (1 - p)(S - C) = S - C + p(S - C)$. This now implies $C \geq p(S - C)$. Noting that $S = k_1 + k_2 + v$ and $k_1 + k_2 \geq C$, it follows that $S - C = k_1 + k_2 + v - C \geq v$, and so the previous inequality implies $C \geq p(S - C) \geq pv$, which fails for v large enough. Hence, there is no distribution of resources that can sustain peace when the surplus is $S - C$ if players can agree to a peaceful bargain when the surplus is S. □

Proof of Lemma 4. We want to show that $\sum_i F_i(t|T)$ is strictly decreasing in T. Note that $\sum_i F_i(T|T) = v + K_1 + K_2 - TC = S - TC$, where S denotes the total amount of resources available in the first period. We can then write the sum of expected fighting payoffs after division in period t as: $\sum_i F_i(t|T) = p\sum_{\tau=0}^{T-t-1}(1 - p)^\tau\big[S - (\tau + t)C\big] + (1 - p)^{T-t}(S - TC)$. We now show that this expression is strictly decreasing in T. Take an arbitrary $T \geq 2$ and consider another division that enables players to fight one more period. Then $\sum_i F_i(T + 1|T + 1) = v + K_1 + K_2 - (T + 1)C = S - (T + 1)C$ and: $\sum_i F_i(t|T + 1) = p\sum_{\tau=0}^{T-t}(1 - p)^\tau\big[S - (\tau + t)C\big] + (1 - p)^{T-t+1}\big[S - (T + 1)C\big]$. Therefore, $\sum_i F_i(t|T + 1) - \sum_i F_i(t|T) = -(1 - p)^{T-t+1}C < 0$. Since we picked T arbitrarily, this establishes the claim. □

Proof of Lemma 5. We want to show that $h(x) = p^2(x - 1) + (1 - p)^x > 1$ holds if, and only if, x is sufficiently high. To do this, we show that $h(\overline{x}) = 1$ has a unique solution $\overline{x}$ and that $h'(x) > 0$ for all $x \geq \overline{x}$. The derivative of $h(x)$ with respect to x is $h'(x) = p^2 + (1 - p)^x \ln(1 - p)$; it is increasing in x when $x < 1$ and decreasing when $x \geq 1$. Let x^* solve $h'(x^*) = 0$, and so $x^* \geq 1$ is equivalent to $p^2/(1 - p) \geq \ln[1/(1 - p)]$. We know that $h(\overline{x}) = 1$ has a solution because h is continuous, $h(1) < 1$ and $\lim_{x\to\infty} h(x) = \infty$. To see that this solution is unique, suppose first that $x^* \geq 1$. Then $h'(x) < 0$ for $x \in [1, x^*]$, and hence $h(x) < 1$ in this interval, and h increases monotonically to ∞ afterwards, so there exists a unique solution $\overline{x} > x^* \geq 1$. Suppose now that $x^* < 1$, which implies that $h'(x)$ increases on $[x^*, 1]$, and then decreases but $h'(x) \geq p^2$. Hence, $h(x)$ is a monotone increasing function of x for $x \geq x^*$, and there exists a unique solution $\overline{x} > 1 > x^*$. Therefore, $h(x) > 1$ holds if, and only if, $x > \overline{x}$ where $\overline{x}$ is the unique solution to $h(x) = 1$. □

Proof of Proposition 2. Let σ_f denote an SPE in which players never redistribute and fight until one collapses from attrition. An example of such an equilibrium can be obtained under the bargaining protocol that prescribes simultaneous submission of demands and agreement if, and only if, the demands sum up to the total available resources. Then the SPE is given by the strategies to demand twice the surplus and fight regardless of the bargaining outcome. (This exploits the simultaneity of moves in the fighting round.) Consider now any period t in which (P) does not hold. This implies that there exist distributions that can deter sneak attacks provided the threats to fight to the end are credible. Let $(\tilde{x}_1(t), \tilde{x}_2(t))$ denote demands such that $\tilde{x}_1(t) + \tilde{x}_2(t) = S(t)$, $\tilde{x}_i(t) \geq A_i(\tilde{x}_1(t), \tilde{x}_2(t))$, and $\tilde{x}_i(t) \geq F_i(t|T)$, where $T = \lfloor \min\{k_1(t)/c_1, k_2(t)/c_2\} \rfloor$. Because (P) is not satisfied, at least one such distribution exists. Consider now the following strategies for each player: in period t, demand $\tilde{x}_i(t)$ and do not fight if the negotiated distribution is $(\tilde{x}_1(t), \tilde{x}_2(t))$; otherwise fight in the current period and then play the strategies prescribed by σ_f in all future periods. Using this strategy, players will achieve peace at $(\tilde{x}_1(t), \tilde{x}_2(t))$ in period t with payoffs $\tilde{x}_i(t)$ each. To check for subgame-perfection, observe that deviation to any other demand results in fighting to the end without redistribution, and since $\tilde{x}_i(t) \geq A_i(\tilde{x}_1(t), \tilde{x}_2(t))$ it is not profitable. Note that the latter inequality also implies that sneak attack is not profitable when the equilibrium peace distribution obtains. Furthermore, since $\tilde{x}_i(t) \geq F_i(t|T)$, accepting the offer is at least as good as fighting to the end at the existing resource distribution. Finally, if the peace distribution does not obtain, fighting in the current period is optimal because the opponent will attack too. Since σ_f is SPE itself, the threat to fight to the end off the equilibrium path is subgame-perfect as well. □

Proof of Lemma 6. We first show that, without loss of generality, players will not reallocate in equilibrium. If (nc, nc) achieves peace without reallocation, then players will not agree to reallocate because doing so necessarily leaves one of them worse off, whether or not such reallocation is peaceful. Hence, in equilibrium (nc, nc) may involve reallocation only when players cannot achieve peace unless they reallocate. So, suppose that (nc, nc) results in a fight in the continuation game. Choose the smallest such n; that is, for all $k_i < nc$, they will not reallocate in equilibrium (we shall later see that $n > 2$). So, if they reallocate in equilibrium when $k_i = nc$, players will fight $t \geq 1$ periods and stop fighting at (m_1c, m_2c). Without loss of generality, assume $m_1 \geq m_2$. If $t = n$, then players must equalize resources at some point (otherwise one of them would collapse sooner). But the player who accepts the larger share initially would not agree to this. Therefore, $t < n$ (that is they will not fight until they both collapse). If they fight until one of them collapses, i.e. $m_2 < 1$, then $t < n$ implies player 2 would not agree to the initial reallocation because she can fight longer without reallocating. Therefore, $m_2 \geq 1$ and players achieve peace when they are both still alive and have resources $k_i = m_ic$ each. If $m_1 = m_2$, then they are indifferent between initial reallocation and no reallocation at all, hence our claim of no reallocation in equilibrium holds without loss of generality. So, assume that $m_1 > m_2$. Because each player can guarantee $m_i c$ in that peaceful period by accepting an initial reallocation of $(m_i + t)c$, rejecting all offers, and fighting for t periods, it is a weakly dominant strategy for each player not to agree to anything less than $(m_i + t)c$ in the initial reallocation. Hence, without loss of generality assume they reallocate initially such that each player has $(m_i + t)c$.

If player 2 disagrees with the initial reallocation, then our choice of n implies that they will fight for at least one period. If they achieve peace in the following period, then 2's share will be $(n-1)c$. But since $t \geq 1$ and $m_1 > m_2$, it follows that $n - 1 > m_2$, and therefore player 2 is better off disagreeing with that initial reallocation. This means that they must be fighting at least two periods. Because by our choice of n there is no reallocation on the equilibrium path in the continuation game, we have

$$V_i((n-1)c, (n-1)c) > V_i((m_1 + t - 1)c, (m_2 + t - 1)c) \tag{8}$$

for some i. Otherwise, players would agree to reallocate and continue with $((m_1 + t - 1)c, (m_2 + t - 1)c)$, contradicting our choice of n. But if i disagrees with the initial reallocation, his payoff is: $p_i(n-1)2c + (1-p)V_i((n-1)c, (n-1)c) > p_i(n-1)2c + (1-p)V_i((m_1+t-1)c, (m_2+t-1)c) = V_i((m_1+t)c, (m_2+t)c)$, where the inequality follows from (8) and the equality follows from the fact that players fight for $t \geq 1$ periods after the reallocation $(m_i + t)c$. In other words, player i does better by rejecting the reallocation, which contradicts our assumption that players reallocate in equilibrium. Hence, without loss of generality, players do not reallocate in equilibrium when they start out with equal resources.

Part (A). Because there is no reallocation in equilibrium, we can find the SPE by simple backward induction. First note that when $n = 1$, fighting destroys all resources and both players get a payoff of zero, which is worse than just agreeing to keep the current resources. Therefore, players achieve peace in this period. If players achieve peace in $k_i = nc$, then player i would sneak attack at $k_i = (n+1)c$ if, and only if, $p(2(n+1)c - 2c) + (1-p)nc = (p+1)nc > (n+1)c$, that is, $pn > 1$. This is necessary and sufficient when players do not redistribute, and since they do not redistribute in equilibrium, this establishes our claim.

To prove part (B), label the period where we want to see if players would fight as $t = 1$. If they fight in this period, by assumption they fight T battles, and then settle in $t = T + 2$ on $k_i = nc$. We now have $S = 2(n + T + 1)c$, and so $S - C = 2(n + T)c$. Further, $F_i(T + 1|T + 1) = p_i(2(n + 1)c - 2c) + (1 - p)nc = nc$. Player i will sneak attack at $t = 1$ if, and only if, $2p(n+T)c + (1-p)F_i(2|T+1) = 2p(n+T)c + 2cp_i \sum_{s=1}^{T-1}(1-p)^s(n+T-s) + (1-p)^T nc > (n+T+1)c = k_i$, or, dividing through by c, noting that $2p_i = p$, and adding and subtracting $p(1-p)^T n$, $2p(n+T) + p\sum_{s=1}^{T}(1-p)^s(n+T-s) + (1-p)^{T+1}n > n + T + 1$. Noting now that:

$$\sum_{s=1}^{T}(1-p)^s(n+T-s) = (1-p)\left[\frac{n\left(1-(1-p)^T\right)}{p} + \frac{pT - 1 + (1-p)^T}{p^2}\right],$$

we can further simplify this to condition (3) stated in the lemma. □

References

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

Bueno de Mesquita, Bruce. 1981. *The War Trap*. New Haven: Yale University Press.

Deutsch, Morton. 1973. *The Resolution of Conflict: Constructive and Destructive Processes*. New Haven: Yale University Press.

Farrell, Joseph, and Eric Maskin. 1989. "Renegotiation in Repeated Games." *Games and Economic Behavior* 1 (December): 327–60.

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

Fearon, James D. 1996. "Bargaining Over Objects that Influence Future Bargaining Power." Paper presented at the annual meeting of the American Political Science Association, Washington, D.C., August 28-31.

Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer Than Others?" *Journal of Peace Research* 41 (May): 275–301.

Filson, Darren, and Suzanne Werner. 2004. "Bargaining and Fighting: The Impact of Regime Type on War Onset, Duration, and Outcomes." *American Journal of Political Science* 48 (April): 296–313.

Gartzke, Erik. 1999. "War Is in the Error Term." *International Organization* 53 (Summer): 567–87.

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

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

Langlois, Catherine C., and Jean-Pierre P. Langlois. 2005. "Fully Informed and On the Road to Ruin: The Perfect Failure of Asymmetric Deterrence." *International Studies Quarterly* 49 (September): 503–527.

Leininger, Wolfgang. 1989. "Escalation and Cooperation in Conflict Situations: The Dollar Auction Revisited." *Journal of Conflict Resolution* 33 (June): 231–254.

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

O'Neill, Barry. 1986. "International Escalation and the Dollar Auction." *Journal of Conflict Resolution* 30 (March): 33–50.

Organski, A. F. K. 1968. *World Politics*. 2nd ed. New York: Knopf.

Orme, John David. 2004. *Paradox of Peace: Leaders, Decisions, and Conflict Resolution*. New York: Palgrave Macmillan.

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

Poundstone, William. 1992. *Prisoner's Dilemma*. New York: Doubleday.

Powell, Robert. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87 (March): 115–32.

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

Powell, Robert. 2002. "Bargaining Theory and International Conflict." *Annual Review of Political Science* 5: 1–30.

Powell, Robert. 2004*a*. "Bargaining and Learning While Fighting." *American Journal of Political Science* 48 (April): 344–361.

Powell, Robert. 2004*b*. "The Inefficient Use of Power: Costly Conflict with Complete Information." *American Political Science Review* 98 (May): 231–241.

Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 60 (Winter): 169–203.

Shubik, Martin. 1971. "The Dollar Auction Game: A Paradox in Noncooperative Behavior and Escalation." *Journal of Conflict Resolution* 15 (March): 109–111.

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

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

Slantchev, Branislav L. 2004. "How Initiators End Their Wars: The Duration of Warfare and the Terms of Peace." *American Journal of Political Science* 48 (October): 813–829.

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

Smith, Alastair, and Allan C. Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48 (December): 783–813.

Teger, Allan I. 1980. *Too Much Invested to Quit*. New York: Pergamon Press.

Zartman, I. William. 1985. *Ripe for Resolution*. New York: Oxford University Press.

Zartman, I. William, and Maureen R. Berman, eds. 1982. *The Practical Negotiator*. New Haven: Yale University Press.