Chapter 6: Preemptive War

Thus far, we have ignored the importance of being the first state to declare war. Regardless of whether state A initiates, state B initiates, or the states meet half way, A always wins with probability p_A and B always wins with probability $1 – p_A$.

However, striking first can lead to some inherent advantages. The primary benefit is the element of surprise. If the other side does not foresee the impending attack, it will be caught off guard and vulnerable. China achieved this exact advantage during the Korean War. As United Nations troops crossed the 38^{th} parallel, the United States anticipated China to remain dormant and therefore did not prepare for the ambush waiting for the allied soldiers. We all know how this affected the outcome: the United States failed to unify Korea, leaving us with the North and the South of today. Second, first movers can dictate when and where the battles will take place and can therefore better pick at the opposition's soft spots.

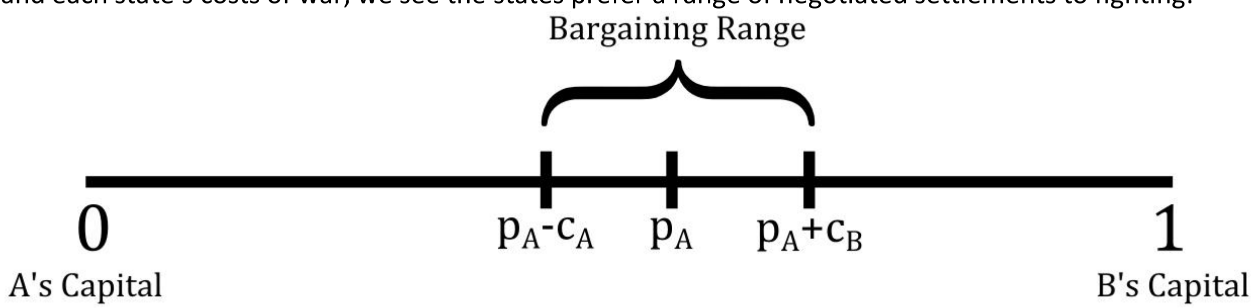
These two advantages indicate that leaving p_A as just p_A is inadequate. How do first strike advantages alter the bargaining environment? Can war be inevitable if striking first provides too much of a benefit? This chapter shows that the risk of surprise attacks can lead to rational preemptive wars—at least in theory. Although there still exist agreements that leave both sides better off than had they fought, neither side is willing to play defensively and suffer the consequences of an offensive onslaught.

However, in practice, first strike advantages are not a compelling explanation for war, as states normally bargain thoroughly before heading into combat. As such, we believe that first strike advantages instead incentivize conflict under the framework of one of the other rationalist explanations for war.

[skip 6.1]

6.2: When Is Preemptive War Inevitable?

Consider the geometric bargaining model from the second chapter. If we plot the balance of power and each state's costs of war, we see the states prefer a range of negotiated settlements to fighting:



This model treats power the same regardless of whether A launches the war or B launches the war. In practice, the initiator of war might have a first strike advantage. We will see that this closes the bargaining range; if the first advantages are sufficiently great, the bargaining range completely disappears, and states must fight.

Consider state A's expected utility for war. For state A to be willing to sit down at the bargaining table, it must expect to earn at least as much by assuming a defensive posture as it would by launching a surprise attack on B. As usual, let p_A be the probability A prevails in a war and c_A be the cost A expects to pay for fighting. For this model, let A's advantage of striking first be $\Delta_A > 0$. Note that since $p_A + \Delta_A$ represents the probability A wins if it strikes first, it must be the case that $p_A + \Delta_A \leq 1$; if $p_A + \Delta_A$ were greater than 1, then A would prevail in a conflict more than 100% of the time, which is impossible. For simplicity, again suppose that each side's cost of war remains unchanged regardless of who strikes first.

Let x be country A's share of the settlement if the states sit down at the bargaining table. Then for A to be willing to negotiate, it must be that:

$EU_A(\text{defend}) \geq EU_A(\text{attack})$
$EU_A(\text{defend}) = x$
$EU_A(\text{attack}) = p_A + \Delta_A - c_A$
$x \geq p_A + \Delta_A - c_A$

We can define B's payoff for a first strike similarly. B normally prevails in a conflict with probability $1 - p_A$. If it strikes first, its probability of victory shifts to $1 - p_A + \Delta_B$, with the restriction that $1 - p_A + \Delta_B \leq 1$. However, it still pays the cost c_B. If the states negotiate, recall that B receives everything A does not, or $1 - x$. Thus, to be willing to sit down at the bargaining table, B's expected outcome must be at least as good as its first strike payoff:

$EU_B(\text{defend}) \geq EU_B(\text{attack})$
$EU_B(\text{defend}) = 1 - x$
$EU_A(\text{attack}) = 1 - p_A + \Delta_B - c_B$
$1 - x \geq 1 - p_A + \Delta_B - c_B$
$-x \geq -p_A + \Delta_B - c_B$
$x \leq p_A - \Delta_B + c_B$

We now have each side's bargaining constraint. The remaining question is whether an x can satisfy both sides and thus allow the actors to avoid war. Such an x fulfills both constraints if:

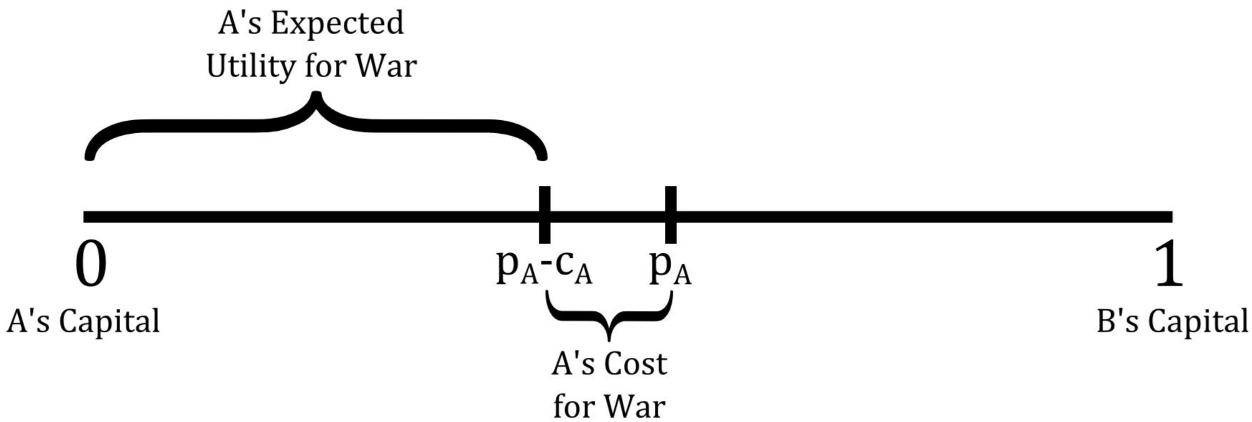
$x \geq p_A + \Delta_A - c_A$
$x \leq p_A - \Delta_B + c_B$
$p_A + \Delta_A - c_A \leq x \leq p_A - \Delta_B + c_B$

To check whether such an x exists, we must see if $p_A + \Delta_A - c_A$ is less than or equal to $p_A - \Delta_B + c_B$. If so, there exists at least one value that the state can agree to and avoid war:

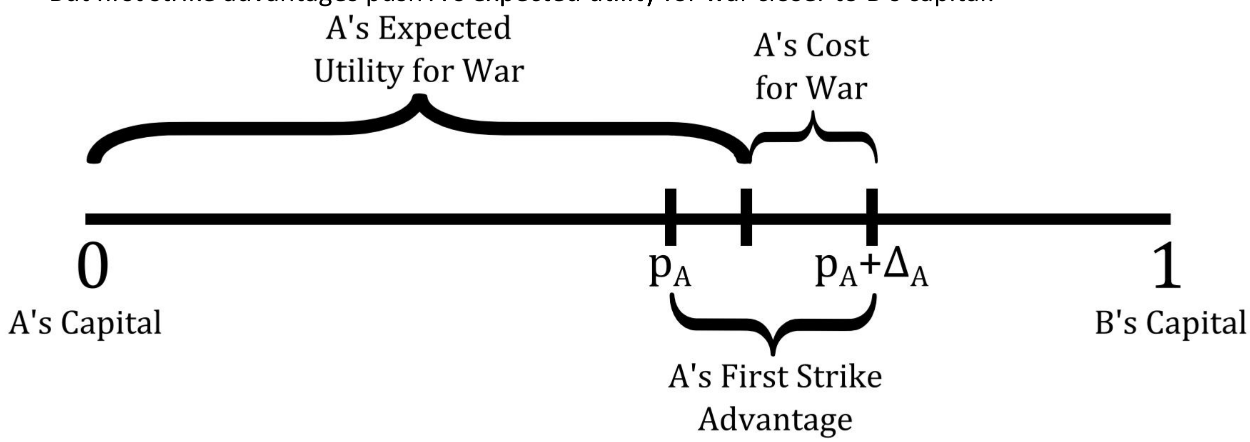
$p_A + \Delta_A - c_A \leq p_A - \Delta_B + c_B$
$\Delta_A - c_A \leq -\Delta_B + c_B$
$\Delta_A + \Delta_B \leq c_A + c_B$

So if the sum of each state's first strike advantage is less than the sum of each state's cost of war, a peaceful solution exists. But if the summed first strike advantages are greater than the cost, then the bargaining range disappears, which forces the states to fight. Thus, preemptive wars are inevitable as long as the first strike advantages are sufficiently great.

We can see this more clearly if we look at the model geometrically. Without first strike advantages, A's expected utility for war looks like this:

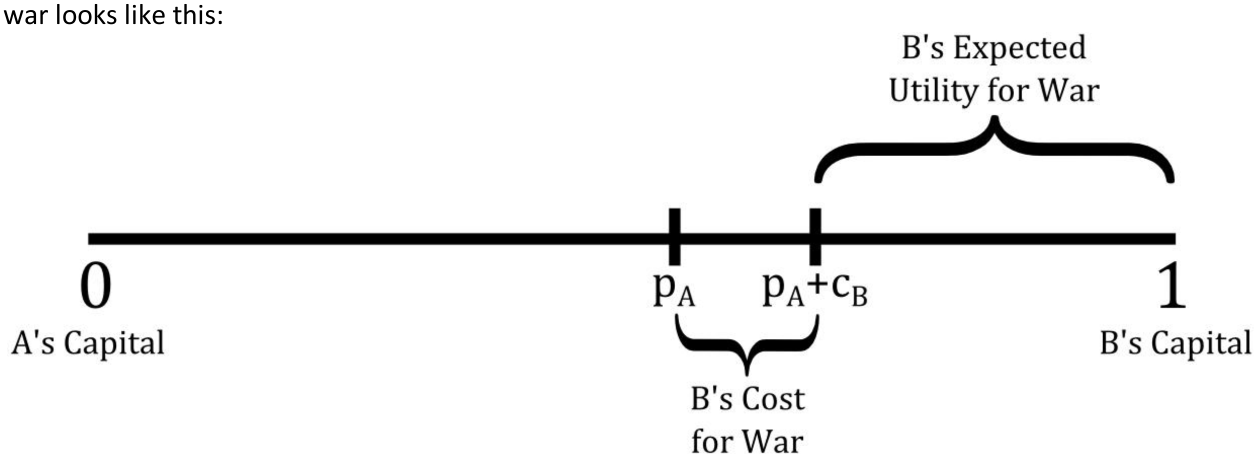


But first strike advantages push A’s expected utility for war closer to B’s capital:

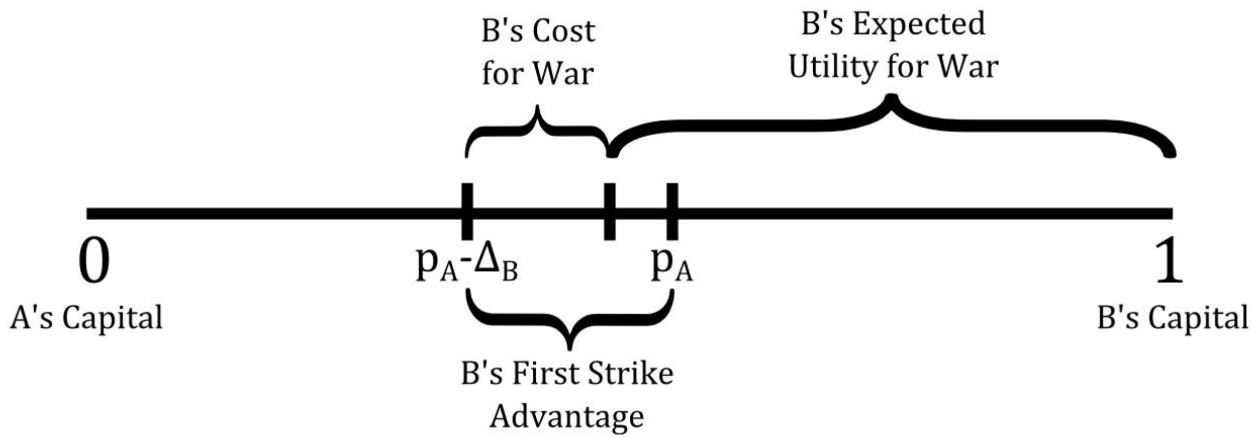


Thus, to satisfy A, the settlement needs to give A at least $p_A + \Delta_A - c_A$ of the good.

Now consider B’s bargaining position. Without the first strike advantages, B’s expected utility for war looks like this:

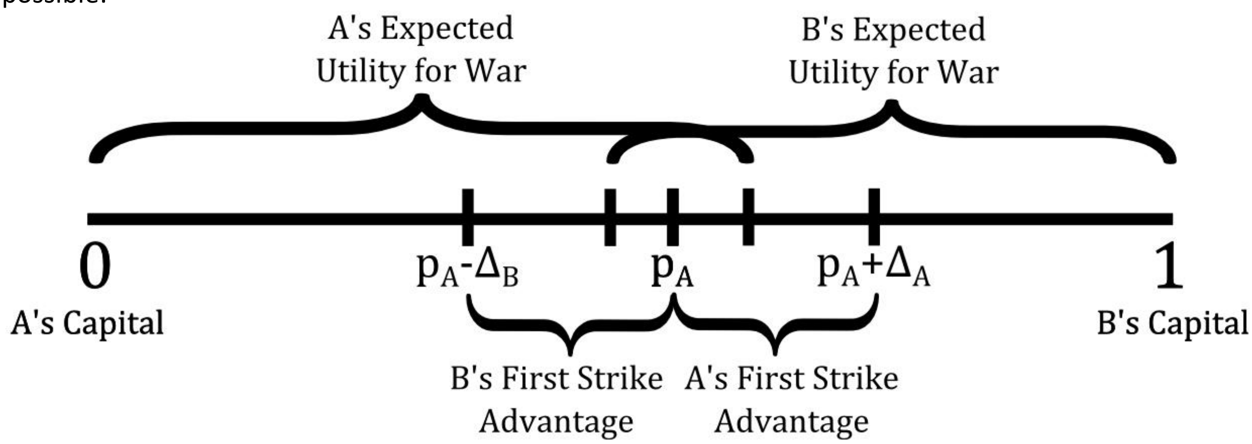


Including B’s first strike advantage moves B’s expected utility closer to A’s capital:



With the first strike, A cannot take more than $p_A - \Delta_B + c_B$ or it will prompt B to declare war.

However, if we overlap each side's expected utility for war, we will see that no peaceful solution is possible:



The sum of A's expected utility for war and B's expected utility for war exceed 1. Thus, no division of that good valued at 1 can simultaneously satisfy both sides' reservation values for war. Consequently, the states are doomed to fight in this scenario.