U.S. Foreign Policy: Explanations for War

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

Last updated: July 7, 2014

1 The Puzzle of War **2**

2 Mutual Optimism **7**
2.1 How Mutual Optimism Can Lead to War 7
2.2 The Role of Fighting 10
2.3 Communication to the Rescue? 11
2.4 Sources of Optimistic Expectations 14

3 Power Shifts and Commitment Problems **18**
3.1 How Large Power Shifts Can Lead to War 19
3.2 Why Actors Cannot Commit: Anarchy 21
3.3 The Role of Fighting 22
3.4 Sources of Power Shift Anxieties 25

4 The Costly Peace **28**
4.1 How Peace Can Be Worse Than War 29
4.2 The Role of Fighting 32
4.3 Sources of Costly Peace 32

5 Military Advantage, Sacred Land, and Divisibility **33**

It is common knowledge that war is perhaps the costliest and riskiest enterprise that human beings can engage in. This very fact should give polities very powerful incentives to avoid it. And yet, the record of human history in that respect is spectacularly dismal: fighting wars seems to have been more or less a regular activity since the earliest anthropological evidence we can find. This is puzzling. We cannot just say, as we often do, that war is politics with admixture of other means. We must also explain why resorting to this particularly awful type of "admixture" is desirable or at least necessary. In this lecture, we shall take a (very brief) look at possible explanations of this puzzle. That is, we shall collect a set of variables that seem to have been useful in understanding why wars begin and why they end. It is these variables that we shall then use in our analysis of how particular societies fought particular wars, and how these wars in turn helped shape these societies.

1 The Puzzle of War

Although it seems that the nature of the conflict should be an important variable in our explanation of war, there is a powerful argument to be made that the search for causes can abstract away from the issue, at least as a first cut, and instead focus on answering why political communities might be unable to resolve a conflict despite their desire for peace. Now, at a very basic level, one might argue that polities go to war because they like fighting (this is akin to the "expressive" motivation for war which we discarded in favor of the instrumental model). If polities go to war for war's sake, then the question of why they fail to reach a peaceful agreement does not even arise. Here we shall assume that peace is generally desirable, war is generally undesirable, but that it is not the case that polities are ready for peace at any cost. These seem like fairly mild assumptions but they are enough to create a serious puzzle about the occurrence of war altogether. Let us put these assumptions together so you can see what I mean.

Consider a (very abstract) setting in which there are only two polities, which we shall call "actors". We shall label the first one A, and the second one (unimaginatively) B. To keep the exposition clear, I shall refer to actor A as "he" and to actor B as "she." These actors wish to divide some benefit. For the sake of simplicity, let's call this benefit "territory" and assume that each actor desires more territory. To make things even more abstract and simpler, let us represent that territory by a line of length 1. Points on this line represent the share of territory that A controls, from 0 (none) all the way to 1 (all of it). Naturally, for any point x on that line, $1 - x$ represents B's share. One way to think about this to put A's capital at 0 and B's capital is at 1. Any point x on the line represents the distance of the border from A's capital, and $1 - x$ represents the distance of the border from B's capital. Let the location of the existing border be at q (the status quo demarcation). Figure 1 shows this representation.

We shall represent conflict in a very simple way. First, we shall assume that war is *costly* — these costs are from the destruction of life and property that is inevitable in every war, but also from supplying and maintaining the military for battle, from dislocations caused to the economy from the redirection of resources away from civilian to military use and the withdrawal of manpower to the armed forces, and possibly from distortions caused by the government's policies (we shall deal with all of these in some detail later). Let $c_{\mathrm{A}} > 0$ represent the war costs to actor A, and $c_{\mathrm{B}} > 0$ represent the war costs to actor B.

Second, we shall assume that war is *risky* — neither of the participants can be assured of victory. This uncertainty arises from the friction that we talked about, both environmental and strategic. To simplify matters even more so that the logic is crystal clear, we shall assume that war is a lottery with only two possible outcomes: an actor can either win it or lose it, draws are not allowed. With this simplification, we can let $p \in (0, 1)$ represent the probability that A prevails in the war, in which case $1 - p$ is the probability that A loses (and so B wins). This probability depends on many factors such as the relative size and quality of the armed forces, the strength of the supporting economies and ability to finance the fighting, the quality of command, as well as the unpredictable environmental factors. We shall call this probability the **distribution of power** because it summarizes the likely outcome of the war as determined by the relative power of the two polities.

Finally, we shall assume that war is a *winner-take-all* affair: the victorious polity absorbs the entire territory of the defeat opponent. This means that p also represents the expected division of the territory if the actors fight a war. For example, if actor A has $p = 0.45$ chance of winning the war, then he will end up with the whole territory (1) with that probability and will lose everything (0) with $1 - p = 0.55$ probability. The expected division, then, is $(0.45)(1) + (0.55)(0) = 0.45 = p$, as we said. Note that we have not assumed anything in particular about the relationship between the status quo distribution of the territory and the distribution the actors expect will prevail if they fight.

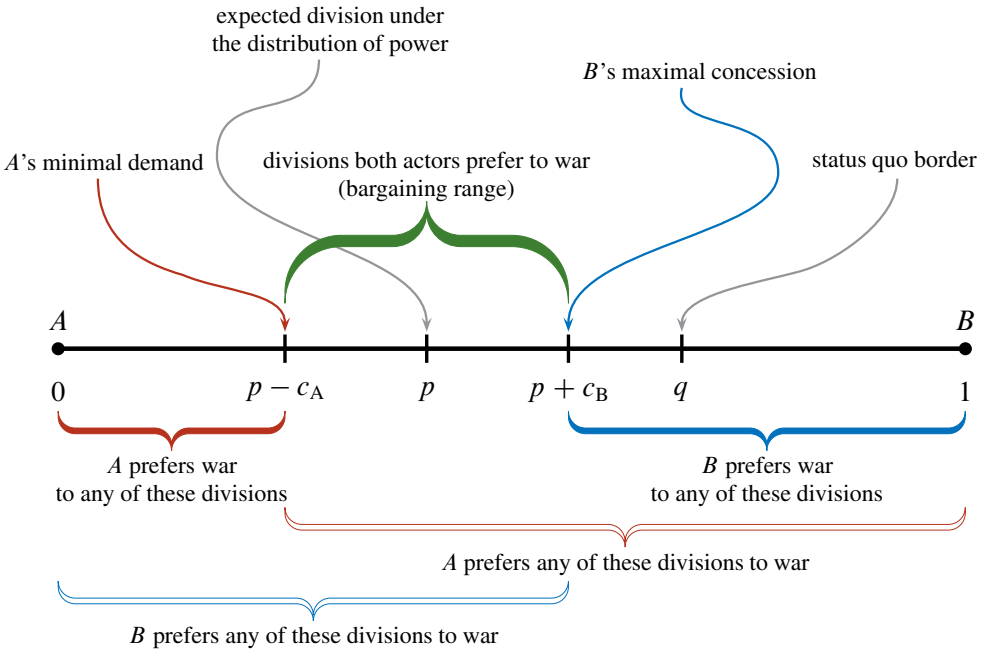

Figure 1: The Puzzle of War. (Points on the line represent A's share.)

We now have all the elements necessary to represent the instrumental value of war in a simple abstract manner. What does actor A expect to happen if war breaks out? With probability p he will win, in which case he will gobble up the entire territory (1). With probability $1 - p$ he will lose, in which case his opponent B will take everything, leaving

polity A with no territory (0). Regardless of the outcome, A must pay the costs of war, c_A. Thus, the **expected value of war** for actor A is

$$W_A = p(1) + (1-p)(0) - c_A = p - c_A.$$

Since this is what A expects to get from war and because he can always choose to fight if he wants to, he will never agree to peaceful concessions that leave him with less territory than this expected share. Thus, W_A represents the **minimal terms** that A would demand in any negotiation with B. Conversely, $1 - W_A$ represents the **maximal concession** that A would be willing to make to B peacefully. In other words, A would agree to any division of the territory that puts the border to the right of his minimal terms. Since the existing distribution of the territory exceeds A's expected value of war, he is **satisfied,** and we would not expect him to fight to overturn the status quo.

Turning now to the other actor, we ask the same question: What does actor B expect to happen if war breaks out. With probability $1 - p$ she will win, in which case she will grab the entire territory, and with probability p she will lose and get nothing. Regardless of the outcome, B must also pay costs of war, c_B. Thus, the expected value of war for actor B is

$$W_B = (1-p)(1) + p(0) - c_B = 1 - p - c_B.$$

Since B's capital is at 1, we can find the maximal concession B will make by marking off a segment of length W_B starting from the end of the line: $1 - WB = p + c_B$, as indicated in Figure 1. Thus, B would agree to any division of the territory that puts the border to the left of this point (her minimal terms). Since the existing distribution of territory is less than B's expected value of war, actor B is **dissatisfied,** and so she would fight to overturn the status quo.

It is worth emphasizing that this bargaining model of war is a representation of the concept of war as an instrument used in pursuit of political objectives. The political objective here is the benefit to be divided (e.g., territory). Victory and defeat are both defined in terms of that political objective. War has no value in itself: it is just a costly and risky way to divide that benefit. We have modeled war as a costly and risky process that culminates in either victory or defeat and we have not allowed either actor to influence the conduct of war or war to influence policy (although we have obviously allowed the threat of war to influence policy in the determination of the minimum terms actor would accept in lieu of fighting).

We now state a simple but perhaps non-obvious fact: *since the costs of war are strictly positive and peace is free, there always exist distributions of territory that simultaneously satisfy the minimal demands of both actors.* Mathematically, we just note that the sum of their minimal terms is strictly smaller than the size of the benefit (territory) to be divided:

$$W_A + W_B = p - c_A + 1 - p - c_B = 1 - (c_A + c_B) < 1.$$

In other words, the simple fact that war is costly engenders the possibility of peace.

We can actually say a bit more than merely asserting the possibility of peace. We can even locate the set of distributions of territory that would be mutually acceptable to both actors. For this we take the intersection of their maximal concessions. Recalling that all

divisions to the right of $p - c_A$ are those that A would agree to without a fight, and that all divisions to the left of $p + c_B$ are those that B would agree to without a fight, we conclude that all divisions between these two boundaries must be agreeable to both. This is called the **bargaining range**, and it is the set of all possible divisions of the territory such that agreeing to such a division leaves both actors with more benefit than their expected values for war. In other words, both actors are better off with any division from this set than going to war. The range comprises divisions that are better than the minimal terms of each actor and less than the maximal concessions they are willing to make.

It is immediately obvious that if the war is costly enough for both actors, the bargaining range can extend to cover the entire territory. Intuitively, if war is that bad, then any peace is preferable to fighting. Thus, for war to occur it has to be the case that fighting is not expected to be exceedingly costly. Not surprising, of course, so we will not dwell on this point except to note that this model might have a hard time accounting for the extreme destruction that many actual wars do entail. We shall return to this point in a bit when we discuss how the *cumulative* costs of war can easily exceed the value of the benefit even when actors are choosing their optimal strategies. We can restate our "simple but perhaps non-obvious fact" as follows: *if war is costlier than peace, then the bargaining range always exists*. It is crucial to realize the importance of this implication. We are saying that the mere supposition that war is costlier than peace means that there always exist deals that can make both actors better off than fighting. But if this is so, then how can we explain war? If there are peace deals that both polities can live with, why would they ever fight?

Does it have something to do with an actor's dissatisfaction with the status quo? Nowhere in this discussion did we make use of the location of the border except to note that B would rather fight than live with it. We have now asserted the possibility of peace, but clearly such a peace must involve a revision of the border in B's favor. Perhaps surprisingly, it does not matter what the status quo distribution of the territory is for the conclusion that peace must prevail. Before we can establish this, observe that at most one actor can be dissatisfied with the status quo. For example, suppose that B is dissatisfied. Because $1-q < W_B$ means that $1 - W_B < q$, we can reduce this to $q > p + c_B$, as depicted in Figure 1. We now prove that when B is dissatisfied, A must necessarily be satisfied. For this, observe that $W_A + W_B < 1$ can be rewritten as $W_A < 1 - W_B < q$, and so A is satisfied because the status quo benefit exceeds its expected value of war. (We can do an analogous calculation by supposing that A is dissatisfied and then showing that in this case B must be satisfied.) Thus, it cannot be the case that both actors are dissatisfied with the status quo: either they are both satisfied, or else only one of them is dissatisfied.

Consider now a simple scenario (not depicted in Figure 1), where the existing distribution is within the bargaining range. Since the benefit of living with this division is strictly higher than the expected values of war for the actors, they are both satisfied, and so neither would fight to overturn the status quo. Moreover, this division is likely to be stable in the sense that it will not be revised through peaceful negotiations. To see this, note that moving the border in either direction must make one of the actors worse off, and this actor would simply refuse to agree to it. Since the other would not fight to force the move, the border will remain at its status quo location.

Perhaps less obviously, peace will prevail even if the status quo is not in the bargaining range (as in Figure 1) although the territorial division will not be stable in that case. In our

example, B is dissatisfied with the existing distribution and would fight unless A agrees to move the border. War, however, would still not occur because A is ready to make enough concessions to satisfy B's minimal demands: any border in the bargaining range represents such a deal. We cannot say where, exactly, the new border would be but we can say that it will lie in the bargaining range. We conclude that when one of the actors is dissatisfied, then the distribution of territory will be revised such that this actor becomes satisfied, and so the border is not stable but peace nevertheless prevails.

Another possibly surprising implication of this model is that even actors who are certain to lose the war might be able to obtain concessions from their opponent. For example, suppose that A is certain to win: $p = 1$. Clearly, B will be willing to give up everything to avoid war since $W_B = -c_B < 0$, and so relinquishing the entire territory is preferable to fighting. Does it follow that A will be able to get everything? Not necessarily. A's expected value for war is $W_A = 1 - c_A < 1$, and so his minimal terms lie to the left of B's capital. The bargaining range comprises all deals that save A the cost of fighting and obtaining sure victory. Thus, it is entirely possible that B can get away with a division of the territory that leaves it with something rather than nothing. Even actors who are certain to be defeated retain some bargaining power because they can still impose the costs of fighting on their opponent. This gives their opponent an incentive to offer a (small) concession and avoid having to pay these costs.[1]

Since we already know that it cannot be that both actors are dissatisfied with the status quo, these two situations exhaust all possible relationship between the status quo distribution of territory and the distribution of power (which determines the satisfaction with the status quo). In all of these, war does not occur. So how can we explain war? The bargaining model of war suggests that we should be looking for reasons that prevent actors from locating a deal in the bargaining range. Broadly speaking, there are three reasons this might happen. First, they might be unsure as to where the bargaining range really is, and so they do not know what concessions are reasonable. Second, they might be afraid of the consequences of not fighting or it might be difficult to commit to upholding the peace deal. This can happen when one actor fears that the other might become much stronger in the future and that it would then force a redistribution of the benefit that is very undesirable. Third, it could be that peace is not free, as the model assumes, but that each actor must incur costs related to maintaining the distribution of power that underpins the territorial division. If that is the case, it might be worth eliminating the threat and reducing the defense burden than living with a costly defense establishment in the long run. In this case the bargaining range might not even exist. Let us now illustrate these possibilities in the basic model of war we have developed so far.[2]

[1] An early statement of this logic can be found in Paul Kecskemeti. 1958. *Strategic Surrender: The Politics of Victory and Defeat.* Santa Monica: Rand Corporation. Available online at `http://www.rand.org/pubs/reports/R308.html`, accessed December 25, 2012. The idea that the losing side can still extract some concessions was called "strategic surrender" but perhaps because of the unfortunate name was badly misunderstood by US Senator Stuart Symington, who apparently thought that RAND was promoting defeatist policies. In an ironic climax of this misconception, US Congress passed a prohibition on using tax dollars to study defeat or surrender of any kind.

[2] This is not to say that these are the only possibilities. For example, if those that decide on war stand to gain disproportionately more from it than society on average and suffer disproportionately lower costs than society on average, then the decision-makers might be biased toward fighting. Under some circumstances, concern

2 Mutual Optimism

One surely heroic assumption we have quietly made is that actors know everything there is to know about the simple world in which they exist. For example, we have assumed that they know the distribution of power, which allows them to figure out their expected values of war, and from there locate the bargaining range. In practice, this assumption is almost certainly violated: polities might be uncertain about the capabilities of their opponents, about the skill of their (or their own) commanders, about the morale of the armed forces, and so on. Since all of these factors affect the probability that one actor will prevail in a war, not knowing any of them means not knowing the distribution of power. In the absence of full information about the distribution of power, actors must rely on their best estimates (or guesses). The problem is that without a commonly agreed to estimate of the distribution of power, actors can end up harboring vastly different views of how war might unfold. This is what Blainey calls "disagreement about relative power." Blainey's argument is that war is caused by such disagreements, but we wish to be more precise in defining what this means. The model will help.

2.1 How Mutual Optimism Can Lead to War

Suppose that there are two possible states of the world. In one, actor A is strong and has a high probability of winning. Label this probability $p_{\mathrm{H}} \in (0, 1)$. In the other state of the world, A is weak, and has a low probability of winning. Label that probability $p_{\mathrm{L}} < p_{\mathrm{H}}$. Neither actor knows the state of the real world, and so neither knows whether A is strong or weak. Figure 2 shows these two possibilities. Observe that if the actors knew which state of the world they actually live in, they would be able to avoid war because the bargaining range exists in each of those states. The only difference is that when A is strong, he can expect to obtain a better deal than if he is weak. The problem, then, is that the actors do not know the actual state.

All actors have are beliefs about the what the state of the world might be. Let $q_{\mathrm{A}} \in (0, 1)$ denote the probability with which A believes that he is strong and so $1-q_{\mathrm{A}}$ is his belief that he is weak. Analogously, let $q_{\mathrm{B}} \in (0, 1)$ denote B's belief that A is strong and so $1 - q_{\mathrm{B}}$ is her belief that A is weak. There is no necessary relationship between these beliefs: q_{A} is merely A's belief that they dwell in the world in which he is strong, and q_{B} is merely his opponent's belief that they dwell in that world.

It is evident from the illustration that if the difference between p_{H} and p_{L} is not very large, the bargaining ranges for the two possible worlds will intersect. If this happens, then peace must prevail regardless of the beliefs that actors have about which world they dwell in. This is so because when the two ranges intersect, B's most demanding minimal terms (in the world where A is weak) are smaller than A's most demanding minimal terms (in the world where A is strong). Any deal between these two is acceptable to both no matter what they believe, and so war would not occur. When the bargaining ranges are disjoint, as in Figure 2, a **zone of maximal disagreement** exists between the most demanding minimal terms of the two actors.

with retaining power domestically can distort the incentives of the ruler who might choose to take the gamble of war instead of facing the unpleasant prospect of being removed from office.

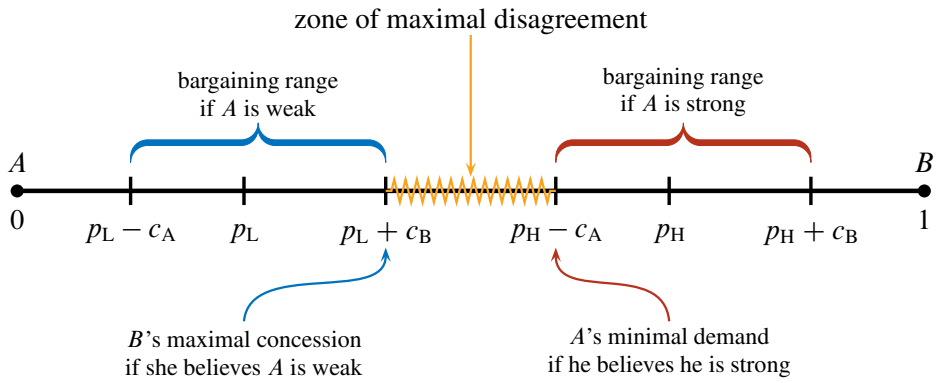

Figure 2: Mutual Optimism and War.

The existence of this zone creates a problem for locating a mutually acceptable bargain when the actors are too optimistic. It should be evident from the illustration that no such deal would exist if A were certain he was strong but B were certain that he is weak. A would not agree to any peace deal that does not locate the border to the right of $p_H - c_A$. Analogously, B would not agree to any peace deal that does not locate the border to the left of $p_L + c_B$. But since $p_L + c_B < p_H - c_A$, it follows that there is no way to position the border such that both demands are simultaneously satisfied. There is no bargain that makes both actors want to avoid war at the same time. As a result, either one can initiate war to resolve the impasse.

We now establish this result for the general case in which actors might be uncertain about the actual state of the world. When A is unsure whether he is strong or weak, his expectation about war must include both possibilities:

$$E(W_A) = \underbrace{q_A}_{\substack{A\text{'s belief} \\ \text{that he is strong}}} \times \underbrace{(p_H - c_A)}_{\substack{A\text{'s war payoff} \\ \text{when strong}}} + \underbrace{(1 - q_A)}_{\substack{A\text{'s belief} \\ \text{that he is weak}}} \times \underbrace{(p_L - c_A)}_{\substack{A\text{'s war payoff} \\ \text{when weak}}},$$

and this defines the minimal terms that A would demand in order to agree not to fight. We can rewrite this for convenience and make the dependence on beliefs very clear:

$$E(W_A) = p_L - c_A + q_A(p_H - p_L).$$

In words, A's war expectation is the payoff he would get if he happens to be weak, $p_L - c_A$ plus the "bonus" if he happens to be strong, $p_H - p_L$, which he expects to obtain with probability q_A (the belief that he is strong). When A is maximally optimistic, $q_A = 1$, we obtain the upper bound on the zone of maximal disagreement. As his optimism decreases (q_A goes down), the upper bound moves to the left, shrinking the zone of disagreement.

Analogous calculations show that B's expectation about war is

$$E(W_B) = 1 - p_L - c_B - q_B(p_H - p_L).$$

In words, B's war expectation is the payoff she would get if she happens to be strong (A's probability of winning is p_L) minus the "penalty" if she happens to be weak, $p_H - p_L$,

which she expects to have to pay with probability q_B (the belief that she is weak). Since B's maximum concession is defined as $1 - E(W_B) = p_L + c_B + q_B(p_H - p_L)$, when she is maximally optimistic, $q_B = 0$, we obtain the lower bound on the zone of maximal disagreement. As her optimism decreases (q_B goes up), the lower bound moves to the right, shrinking the zone of disagreement.

Since the disagreement zone shrinks as the optimism of the actors decreases, it stands to reason that at some point that zone would disappear altogether and a peace would become possible despite the uncertainty. We now derive a condition that *ensures* that a zone of disagreement exists; that is, we derive a condition on the various parameters of the model that is *sufficient* to guarantee that war must occur.

As before, the war expectations determine the minimal terms that actors would demand if they are to agree not to fight. If these terms exceed the benefit that they can divide in peace, then there would be no war to divide that benefit to their mutual satisfaction – at least one of the players would have to receive something strictly worse than his war expectation. Naturally, he would not agree to this, and war would have to follow. This suggests that we can define **mutual optimism** as those beliefs that ensure that these minimal terms cannot be satisfied:

$$E(W_A) + E(W_B) > 1.$$

Substituting the definitions of these expectations that we derived above and rearranging terms yields the **mutual optimism condition**:

$$\underbrace{\underbrace{(q_A - q_B)}_{\substack{\text{extent of}\\ \text{disagreement}}} \quad \underbrace{(p_H - p_L)}_{\substack{\text{size of difference}\\ \text{between the}\\ \text{two worlds}}}}_{\substack{\text{expected total}\\ \text{benefit of war}}} > \underbrace{c_A + c_B}_{\substack{\text{total}\\ \text{costs of war}}} . \qquad \text{(MO)}$$

We can interpret this condition as follows. On the right-hand side are the total costs of war, which are always positive. Moreover, if they are sufficiently large, then this condition cannot be satisfied. In other words, if war is sufficiently costly, then no amount of mutual optimism would cause the actors to fight. The left-hand side comprises two terms. The *extent of disagreement* is the difference between the probabilities that A and B assign to being in the world where A is strong. The *size of difference between the two worlds* is the difference between the expected gain when A is strong and when he is weak. Multiplying the extent of disagreement by the size of the difference gives us the *expected total benefit from war*. Trivially, the equation states that war must occur when its total expected benefit exceeds its total expected costs.

Let us think a bit about the two terms that define the total expected benefit. Consider first the size of the difference, $p_H - p_L$, which is always positive, and which measures how important the consequences of disagreement might be. Intuitively, if the two probabilities are close to each other, then the expectations about war in the two states of the world would have to be fairly close as well. This means that for the disagreement to matter, the extent of disagreement would have to be very large. (If they are too close, then the bargaining ranges of the two worlds will intersect, and war will not occur.) Conversely, if the potential difference between the two worlds is great, then even a relatively modest

extent of disagreement might wipe out any mutually-acceptable bargains. Anything that increases the expected difference between the two worlds (e.g., an untested technology that might give A a decisive advantage in battle) can make war more likely.

Turning now to the extent of disagreement, $q_A - q_B$, it is worth noting that since we put no restrictions on the beliefs actors hold, it is possible for this term to be non-positive: $q_A \leq q_B$. In this situation, A places a smaller probability on the state of the world in which he is strong than B does. It should be immediately obvious that if this is the case, then war would never occur: B, who pessimistically believes that A is strong, would offer a sizeable concession that A, who thinks he is actually weak, would only be too happy to accept. Mathematically, if this term is non-positive, then the mutual optimism condition can not be satisfied, so actors will not fight. If the term is positive, then it captures the extent to which A believes more strongly than B that he is strong. If the extent of disagreement is relatively small, then the condition for war can only be satisfied if the potential size of the difference between the two worlds is very large. Conversely, if the disagreement is relatively large, then even small potential differences between the two worlds can matter. Anything that makes A more optimistic (increases q_A) would make war more likely, just like anything that makes B more optimistic (decreases q_B) does.

Thus, (MO) defines mutual optimism: it specifies the condition that beliefs must satisfy for a disagreement zone to exist. Thus, we can use (MO) to explain war as a consequence of mutual optimism. It is not simply necessary that actors disagree about their relative strength; it must be that they are both "too optimistic" about what they expect war to accomplish. Condition (MO) gives a precise meaning to the phrase "too optimistc" by relating the difference in beliefs to the size of the differences between the two possible worlds and the costs of war. To put it in another way, each actor must be so optimistic that the maximal concession its opponent is prepared to make cannot satisfy that actor's minimal terms. When this happens, war must break out.

How can such a divergence occur? One possibility is in the different doctrines of war that the actors might have. For example, actor A might believe that he has a very strong offensive capability and is doctrinally committed to waging an aggressive war. Actor B, on the other hand, might have developed tactics that it believes will be effective in defense, and so believes that A's chances of success from aggressive war are very small. This can result in a great divergence between the expected consequences from the different distributions of power. Depending on the emphasis the actors place on their doctrines being true - the optimism about being correct (and such "motivated bias" is not uncommon), these differences can close the opportunities for peace.

2.2 The Role of Fighting

It is one thing to say that "mutual optimism causes war" but it is quite another to explain how fighting a war is supposed to resolve that. And resolve it war must because very few wars end with the total obliteration or disarmament of the defeated party. Most wars actually end while both sides have the ability to continue to fight. This suggests that somehow fighting has enabled them to agree to peace even though initially neither was willing to make concessions. To study this, we need to move from a model of *absolute* war to a model of *ideal* war where fighting and bargaining are simultaneous processes. Doing so makes the

analytical work more demanding, so for our purposes I will simply show the mechanism that such a model reveals.[3]

Once war begins and neither actor is immediately defeated, fighting can gradually reveal what the true state of the world is. For example, if the true state is that A is strong, then A would be more likely to prevail in battles, maintain its army, and generally do better in the war than B. Since most of this is actually observable by both actors, they will begin to revise their estimates about the true state of the world. The process can be slow and noisy because chance factors might still intervene and cause A to lose particular engagements (friction!) but on average A would be doing better and both actors would know it. As B becomes more pessimistic, her minimal terms would become progressively more concessionary, shrinking the disagreement zone. With enough fighting, this process would cause this zone to disappear altogether. The bargaining range will reappear in the form of deals that both the optimistic A and the now pessimistic B can agree to, and the fighting will end. War provides the "stinging ice of reality," as Blainey calls it, which cures actors of their optimism. Since performance in the war is necessarily related to the true state of the world, the beliefs about the state (and this about the location of the true bargaining range) revised on the basis of this performance would converge until they reach a point where peace becomes possible. The war terminates when actors agree on their relative strength, and the process of fighting allows them to revise their estimates until they do come to such an agreement. Thus, war occurs *because of* mutually optimistic assessments of what war will be like, and fighting continues *in order to* reconcile these expectations, which can enable the actors to terminate the war before one of them is defeated militarily.

2.3 Communication to the Rescue?

One might wonder why actors would fail to reconcile their contradictory assessments about the distribution of power without resorting to a fight. After all, if excessive optimism is pushing them to a war that would not occur if they agreed on a common estimate of its outcome, then they have a strong mutual interest in sharing information so that their beliefs converge without a war. Unfortunately, the private interest in obtaining more favorable terms can overwhelm the collective desire for peace. To see how this can happen, observe that sharing of information, especially information that is not readily verifiable but still crucial (e.g., about one's resolve or the morale of troops) can be very difficult for strategic reasons. The reason is that ultimately, the willingness of an actor to make larger concessions depends on that actor's estimate of the opponent's value of war: each actor wants to push for a deal that barely satisfies the minimal terms of the opponent. As we have seen, in our scenario, the problem arises from uncertainty about the magnitude of these terms. This is what the opponent is now supposed to rectify by volunteering information about its own expected value of war. But why would that opponent be truthful? There are two reasons to doubt that he will be truthful, one that concerns the incentives of a weak actor to conceal his weakness, and another that concerns the incentives of the strong actor to pretend that he is weak.

[3]The argument in this section uses the extension of the model to *ideal* war in Branislav L. Slantchev. 2005. "The Principle of Convergence in Wartime Negotiations," *American Political Science Review,* 97(4): 621–32. This is one, but by no means the only one, formalization of the process of belief convergence.

Consider the incentives an actor, say A, who believes he is likely to be weak: would he necessarily wish to reveal that belief? Suppose that A is truthful: he tells B that he is weak when he believes himself to be weak, and tells B that he is strong when he believes himself to be strong. Since he is truthful, B will believe these statements and will revise her beliefs, offering a small concession to the self-described weak A and a better deal to the self-described strong A. But if the opponent believes his statements, will A wish to be truthful in his communication? Of course not: when he believes himself weak, he can simply lie and tell B that he is strong – since she expects him to tell the truth, she will offer the more attractive terms, which he will happily accept. Naturally, B is quite aware of this possibility and as a result will not believe any unverifiable pronouncements that A makes. The possibility for truthful communication is undermined by the incentives the actors have to misrepresent their beliefs for bargaining leverage.

In principle, one could overcome this problem by devising signals that A can only send if he is truly more likely to be strong; signals that he cannot fake if he is likely to be weak. For example, military maneuvers can reveal the training of his troops; a public discussion of his budget can reveal the extent of preparedness and popular consensus behind the policies; and revealing one's troop location and equipment can help establish the extent of mobilization and readiness to fight effectively. Badly trained troops will perform miserably in these exercises; a divided polity would voice its disagreements with the policy; and not having troops in sufficient numbers or with adequate equipment would reveal that one cannot expect to fight effectively. It then follows that revealing the exercises, the debates, or the location of troops can credibly inform the opponent about the likelihood of the actor being in a strong military position. Thus, one might think that all one has to do to overcome the communication problem is find such strategies.

Unfortunately, this need not be the case: when it is possible for the opponent to use this information against an actor's interest, then the incentive to reveal it might disappear even though the lack of revelation could lead to war.[4] Suppose A is strong and consider the strategy of revealing his troop dispositions in an effort to impress B. One possible reaction, of course, is that B is duly impressed and offers better terms. The other reaction, unfortunately, is that B uses this information to prepare a more effective assault and in the war that results A's military advantage is neutralized. Revealing the information about being strong can enable the opponent to take counter-measures that undermine that strength and dissipate whatever bargaining leverage A was hoping to obtain. In a situation like this, A might not merely seek to conceal such sensitive information; she could also actively try to mislead B by fostering a sense of *false optimism*. Actor A who believes he is strong pretends to be weak in the hope that this would cause B to indulge in excessive overconfidence and perhaps rush into a war without adequate preparation. While A would have to forego the chances of obtaining a better peace deal (after all, now B is even more optimistic than before), he will have a much better chance of victory against an unprepared and surprised opponent. This sort of problem cannot be corrected through communication even when credible demonstrations are available — actors with incentives to feign weakness would not reveal the information even when they could.

[4]This section summarizes the argument in Branislav L. Slantchev. 2010. "Feigning Weakness," *International Organization,* 64(Summer): 357–88.

One manifestation of this problem occurred in 1950 when the United States was trying to ascertain whether China would intervene in the Korean War. The initial objective of the U.S.-led intervention had been accomplished – the North Korean army was expelled from South Korea and all but destroyed. With no opposing forces between them and the Yalu River, the Americans were tempted to invade North Korea and unify the peninsula under the leadership of the South. The only military power that could potentially stand in the way was China (backed by the Soviet Union), and the U.S. did not want to fight China over Korea. Before making the crucial decision to invade the North, the U.S. leadership made a concerted effort to determine whether China would intervene. The Chinese leadership was claiming, more or less, that they might, but we know how much faith one should place in such statements. Consequently, the Americans tried to infer China's intent by looking for behaviors that China would engage in if it were truly prepared to fight. In our language, China could either be strong (prepared to intervene and willing to do so) or weak (lacking in one or both of these). The U.S. used planes to try to detect troop movements – was China sending troops to North Korea to defend it? The U.S. used intelligence to monitor preparations in Beijing – was the Chinese government ordering citizens to board up their windows as defense the inevitable American air strike in case of war? The U.S. asked all its allies with links to China to estimate the likelihood of intervention. All sources pointed to the same conclusion: there was no evidence of troop movements to Korea, no preparations for war, and no credible communication even privately that China would fight. Having thus decided that the absence of a credible signal of strength is evidence of weakness, the U.S. crossed the 38th parallel and invaded the North...

...only to run headlong into a massive Chinese army of crack troops. The Chinese were not merely in Korea, they were there, hiding, in strength. What had happened? If the Chinese were going to intervene, why had they failed to reveal that to the Americans? It is fairly clear from the evidence that had they done so, the U.S. would not have invaded. Why conceal all preparations and practically guarantee that the U.S. would go North? The key to understanding that episode is in the lack of incentive to reveal the military preparations that would have convinced the U.S. to stay out. Showing the disposition and numbers of their troops would have persuaded the U.S. that the Chinese were serious. Unfortunately, the Chinese had no way of knowing what the consequence of that would be: the U.S. might acquiesce but it might instead choose to use its massive firepower to pummel those very troops that China was using for deterrence. At the time, the U.S. had a clear superiority in air power – the nascent Chinese forces had no air force, and the negotiations with Stalin to provide air cover for the ground operations had stalled. Thus, should the U.S. choose to take advantage of the information the Chinese revealed, it would have little difficulty neutralizing their forces. (With General MacArthur in triumphant after the success of his Inchon landing, such an operation might not have appeared unlikely.)

The Chinese were caught between a rock and a hard place: do not reveal the troops and risk almost certain war in which they would enjoy the advantages of surprise; or reveal the troops and either get a good deal (the U.S. stays south of the 38th Parallel) or end up in a war against a fully prepared United States that pulverizes your defenseless troops. With such an unenviable decision to make, the Chinese leadership opted to preserve the military advantage of surprise – they moved over 300,000 soldiers in complete secrecy using round-about routes and marching only at night (officers had orders to shoot any stragglers

who broke cover during the day precisely in an effort to avoid detection from American overflights), ending up in North Korea and delivering a serious blow to the unprepared Americans.[5]

The bottom line is that the incentives to hide information can make it impossible to reveal it in a way the opponent would believe. But if the opponent's beliefs are not affected by communication, then belief convergence cannot occur without actual fighting. Since fighting is more "truthful" than words (performance in war depends on the actual state of the world, not on the state the actor wishes it to be or claims it to be), war can play a function that diplomacy cannot. Bullets speak louder than words: the belligerents can converge on an expected outcome and end the war.

2.4 Sources of Optimistic Expectations

We conclude that mutual optimism can be a cause of war, and so anything that promotes such optimism can be a contributing factor to both the outbreak of war and its continuation. Conversely, anything that reduces that optimism is a contributing factor to peace and war termination. This suggests several variables we might want to consider in our study of war and society.

Consider **overconfidence** in one's ability to win, or in the reliability of one's allies. This can arise from exaggerated sense of the quality of one's own armed forces, the competence of the military leadership, or the ability to sustain the necessary war effort. It can also arise from a very negative assessment of the opponent's quality, competence, and economic potential. The image of one's own superiority and the opponent's inferiority is important, and it can be arise out of religious differences (God is on our side), racism (we are the superior race), nationalism (our society is more civilized and advanced), faith in technology (our weapons are superior), unifying morale (our cause is just), or military culture (we are better warriors and/or our doctrine is superior). Overconfidence might also be an evolutionary adaptation, so we might all be prone to it for biological reasons.

One famous example of overconfidence that turned out to have been misplaced occurred during the Peloponnesian War in 416 BC. The Athenians landed a powerful army on the neutral island of Melos and demanded that the Melians surrender and pay tribute to Athens.[6] The Melians were hopelessly outnumbered so they tried to reason with the Athenians, arguing that "the fortune of war is sometimes more impartial than the disproportion of numbers might lead one to suppose." (There is always risk involved in war, so there is a chance that Athens would not win.) The Athenians countered that while this was generally true, the imbalance of military power between them and the Melians was too great to give the Melians any meaningful probability of avoiding defeat. To this, the Melians replied as follows:

> You may be sure that we are as well aware as you of the difficulty of contending against your power and fortune, unless the terms be equal. But we trust that

[5] See "Feigning Weakness" for a summary of this argument. I deal with the other aspects of the Chinese intervention in Branislav L. Slantchev 2011. *Military Threats: The Costs of Coercion and the Price of Peace.* Cambridge: Cambridge University Press, Chapter 6: The Expansion of the Korean War, pp. 191–223.

[6] Thucydides. 1996. *The Landmark Thucydides: A Comprehensive Guide to the Peloponnesian War, edited by Robert B. Strassler.* New York: Touchstone, pp. 353–54. All quotes used here are from this edition.

> the gods may grant us fortune as good as yours, since we are just men fighting against unjust, and that what we want in power will be made up by the alliance of the Spartans, who are bound, if only for very shame, to come to the aid of their kindred. Our confidence, therefore, after all is not so utterly irrational (5.104).

In other words, the Melians insisted that they were not as weak as the Athenians believed because the gods were on their side, their cause was just, and their allies would help them. This created a large discrepancy between what the Athenians believed about the likelihood of victory and what the Melians believed. The Athenians based their estimate on their military power; on the very high probability that the Spartans would not be able to arrive in time to intervene even if they were inclined to do so, which they probably were not; and on their belief that all that talk about gods and justice was delusional. Both sides were optimistic in the sense we defined it here: the Melians were unwilling to make the concessions that the Athenians were demanding because they disagreed about what war would entail. Instead of surrender and tribute, they offered a friendly neutrality, which the Athenians found too small of a concession. In the end, the Melians firmly refused to yield to the Athenian demands whereupon the Athenians besieged the city, took it, slaughtered all men, and sold all women and children into slavery. The "stinging ice of reality" had shown that the Athenian assessment had been closer to the true state of the world: neither the gods nor the Spartans had materialized to help the Melians.

There also might be **agency problems in civil-military relations**. The generals might not provide entirely accurate assessments to the politicians. Although this often takes the form of claiming unpreparedness (and demanding larger budgets), it also might be out of personal desire for glory, exaggerated self-confidence, or interest in a policy that might not be entirely in line with what the politicians want. The military leadership can also hide adverse developments in an effort to avoid appearing incompetent and facing censure. All of these activities would leave the political leadership with the mistaken impression that its military position is far stronger than it really is. As H.H. Asquith, the British Prime Minister at the outbreak of the First World War once remarked,

> [The War Office kept three sets of figures,] one to mislead the public, another to mislead the Cabinet, and the third to mislead itself.[7]

Then there's the mobilization of **public opinion**. When the government needs to maintain support for its military policies, it might cultivate the desired public opinion with intense propaganda that conceals the true state of affairs. Whether or not the government believes its own hype (and they often do), once the public is whipped into frenzy, it would be very difficult politically to change course. The effect of this falsely created optimism would be equivalent to the real thing: those who believe would work hard for the aggressive policy they support, and those who do not would be silenced out of fear of appearing out of step with the rest. Incidentally, this might mean that democracies might be more prone to this problem because democratic governments might be more constrained by public opinion. One famous statement of this view is by Walter Lippmann, who condemned public opinion outright:

[7] Quoted in Alistair Horne. 1994. *The Price of Glory: Verdun 1916.* London: Penguin, p. 23.

> The unhappy truth is that the prevailing public opinion has been destructively wrong at the critical junctures. The people have impressed a critical veto upon the judgments of informed and responsible officials. They have compelled the government, which usually knew what would have been wiser, or was necessary, or what was more expedient, to be *too late with too little, or too long with too much,* too pacifist in peace and too bellicose in war, too neutralist or appeasing in negotiations or too intransigent. Mass opinion has acquired mounting power in this country. It has shown itself to be a dangerous master of decision when the stakes are life and death.[8]

One need not take such a drastically pessimistic view of public opinion but one should not ignore the impact it has on politicians concerned with retaining office. Even non-democratic leaders might be very sensitive to this opinion, especially if their legitimation claims rest on some sort of claims to competence in foreign affairs or military matters more generally.

Optimistic assessments can also be linked to **windows of opportunity**: temporary weaknesses of opponents that invite aggression. One very common example is attacking or otherwise trying to exploit an opponent who is engaged in another conflict already. With resources and army already committed to that other conflict, the opponent is (temporarily) weaker or at least expected to be weaker, and so the group can attempt to drive a hard bargain. A polity torn by revolution or civil war can also invite aggression in the belief that it cannot muster the resources for a proper defense. Sudden or extra-constitutional changes in government (e.g., death of a monarch or a coup) can also destabilize the internal cohesion of the group and create doubts about its ability to mobilize for military action. An economic or fiscal crisis, a government near bankruptcy, or the perception of imperial overstretch, all of these can motivate optimistic beliefs in the opponents.

Consider briefly the consequences of three revolutions: the Iranian, the Russian, and the French. When the Iranian Revolution toppled the Shah in 1979, it was not immediately clear that the religious faction would come to dominate politics — there were many competing groups, most of them secular. With the Iranian government in chaos, neighboring Iraq's Saddam Hussein decided that the time was perfect for an invasion, an attempt to annex the oil-rich Khuzestan Province while the country is in turmoil. The consequences of the 1980 invasion are well-known: Hussein's optimism proved unwarranted — the Iranians buried their differences, and since the religious faction was the only seemingly capable of organizing any sort of defense, it emerged dominant from the Revolution. After eight years of gruelling and vicious war, Iran prevailed, Khuzestan was saved, and the government had become theocratic.

Whereas the aggressor failed in this instance, opportunistic interventions can often lead to remarkable results. For instance, when the Soviet coup toppled the Tsarist rule in Russia in 1917, the Germans — who had been fighting the Russian Imperial Army for three years — decided to press home their advantage. The Soviet regime was weak, the Russian army was disintegrating, and the forces of counter-revolution were already on the move. The Soviet government could not hope to deal simultaneously with these threats. This is not to say that it did not try: in their first negotiations with the Germans, the Soviets argued for peace

[8] Walter Lippmann. 1955. *The Public Philosophy.* New York: Atlantic-Little Brown, p. 20 (emphasis added).

without territorial concessions (meaning that they simply wished to withdraw from the war and keep the pre-war boundaries). The Germans disagreed and resumed their advance, threatening to reach the capital and perhaps undo the new regime. In an extremely painful political move, Lenin prevailed and persuaded the Soviet government to make concessions to the Germans in order to disengage from the First World War to focus on its internal situation. The concessions the Germans wrested in this way were immense. In the Treaty of Brest-Litovsk (1918), the Soviets had to relinquish a quarter of the Empire's population and industry, and almost all of its coal mines, among other things. Ukraine — among the most important sources of grain for the Empire — was also lost. Despite the harshness of the peace, the Bolsheviks did gain the breathing room they needed, and managed to consolidate their hold on power after nearly five years of a brutal civil war. With Germany getting defeated in 1919, the treaty was abrogated, and the resulting disputes over territory fueled numerous conflicts between the two world wars.

Thus, a country torn by revolution could be a tempting target because of the optimism this window of opportunity can create. The fact that the "stinging ice of reality" sometimes disabuses the aggressor from that optimism should not obscure the fact that it was this confidence in victory – usually well-founded in such chaotic situations – that propelled him into action.

One potentially important twist might be the ruler's **reputation** or at least the perception that failure to defend one's interests vigorously (especially if this happens under duress) would be interpreted as a sign of weakness and invite future aggression. Under the reputational logic, the group fights now in order to demonstrate that it is strong and thus deter further challenges, possibly by groups not necessarily related to the one it fights now. Fighting is a signal of strength that is supposed to dispel the unwarranted optimism of opponents that is assumed to arise if the group fails to fight (which is why the group expects more serious challenges in the future). Just as in the mutual optimism explanation, fighting is supposed to lead others to form more correct estimates of the distribution of power and of one's resolve. Because of this, they would be more likely to agree to acceptable peace deals instead of indulging in demands that are likely to provoke war because one is unwilling to make the necessary concessions. Note, however, that this line of reasoning does not require that one be particularly optimistic about the war that is being waged out of reputational concerns (although one would be hard-pressed to see how such a war is supposed to impress others if one were to lose it).

One famous statement of this logic is the *Domino Theory* according to which if one of the dominoes is allowed to fall, then all the others must necessarily follow, thereby justifying fighting to prevent the fall of the first one. Although the metaphor is different (one rotten apple infecting the others in the same barrel), the logic used by US Secretary of State Dean Acheson in 1947 to persuade key congressmen to support military measure to contain the spread of Communism was the same:

> Soviet pressure on the Straits, on Iran, and on northern Greece had brought the Balkans to the point where a highly possible Soviet breakthrough might open three continents to Soviet penetration. Like apples in a barrel infected by one rotten one, the corruption of Greece would infect Iran and all to the east. It would also carry infection to Africa through Asia Minor and Egypt,

> and to Europe through Italy and France... The Soviet Union was playing one of the greatest gambles in history at minimal cost... We and we alone are in a position to break up the play.[9]

This line of reasoning has rather a lot of bodies buried in it. There are many difficulties with the concept of reputation to begin with, and with its application to inter-group conflict and war more specifically. Reputation, of course, is something that others confer on one, and as such it is largely beyond one's control. Manipulating the beliefs of others is very difficult because how others interpret one's behavior might have just as much to do with them and with the relations between the two, as with the acts one engages in. For example, too vigorous of a defense of one's interests might easily be interpreted as a sign of aggression and prompt a countervailing response. Appeasing behavior by a friendlier group might be interpreted as an act of generosity rather than weakness. Aggressive behavior in itself might have a detrimental effect if it causes the other to believe that it is being used merely for the sake of maintaining reputation.

At the end of the day, however, whether reputation exists or not and whether it can be successfully manipulated or not might be less relevant than whether groups and their leaders believe that reputation is worth fighting for. If they do, then reputation is just as real as any other factor one might wish to consider.[10]

3 Power Shifts and Commitment Problems

The model we have been discussing is static in the sense that the interaction between the two actors occurs once and for all. Although useful in highlighting some features (like the ones we discussed), the model might obscure others. Relations among groups happen continuously and there are many occasions for conflict, sometimes even over the same issue. We shall now extend our model in a very simple way: the actors will engage in conflict bargaining twice. If they fight on either occasion, the war settles the conflict for good: to the victor go the spoils forever (so, again, we are dealing with absolute war). If they do not fight, they enjoy whatever shares of the benefit they agreed to. Their overall payoff is simply the sum of the payoffs from each negotiated deal.

To introduce some dynamics into the model, let us assume that A is declining in relative power: if war occurs today, A's probability of winning is p_{H} but if war occurs tomorrow, A's probability of winning drops down to p_{L}. Consult Figure 2 for the bargaining ranges that this entails. Instead of depicting two possible states of the world about which actors might be uncertain, the figure can be interpreted as depicting the bargaining problem today (while A is strong) and tomorrow (after A becomes weak). Of course, the whole point now is that these cannot be treated in isolation from each other: what happens today will have implications for the future. In that sense, the problem today is not one of merely locating a

[9] Dean Acheson. 1987. *Present at the Creation: My Years in the State Department.* New York: W.W. Norton & Company, p. 219.

[10] Some have argued that reputation is not worth fighting for. See, for example, Jonathan Mercer. 1996. *Reputation and International Politics.* Ithaca: Cornell University Press. Whether this is so scarcely matters for our purposes — if groups believe it to be so and fight for it, then we need to consider it.

deal in the bargaining range that would result if today is considered in isolation. The problem today is to take into account both the immediate payoff and the future consequences of the choice.

3.1 How Large Power Shifts Can Lead to War

What should we expect to happen now? When deciding on a strategy today, the actors will naturally look at its possible consequences for the interaction tomorrow. This means that we should begin by analyzing what the future holds in store for them. If war occurs now, whoever emerges victorious will enjoy the entire benefit tomorrow. That's the easy case. If, on the other hand, the actors do negotiate some peaceful agreement today, then they get to engage in another round of negotiations tomorrow. Since A will decline after peace today, the future settlement would have to be contained in the bargaining range when A is weak. Regardless of the terms A agrees to today, he can expect no more than $p_{\mathrm{L}} + c_{\mathrm{B}}$ tomorrow. This is because we already know that under the new distribution of power, B cannot be induced to make a larger concession. This, then, is the *maximum future benefit that B can credibly commit to*. Observe in particular that the actors do not expect war tomorrow. What they do expect is a different peace; one that would involve terms that are worse for A. The question now is: Can the actors avoid war today?

Consider the declining group's war payoff if he fights now. With probability p_{H} A will win the war and enjoy the full benefit twice while suffering the costs of war: $W_{\mathrm{A}} = 2p_{\mathrm{H}} - c_{\mathrm{A}}$. This represents the *minimal* terms that A would have to be offered to forego war today. What can B offer?

As we have seen, the largest concession that B can credibly offer tomorrow is $p_{\mathrm{L}} + c_{\mathrm{B}}$. Actor A would not believe any promise of a larger concession because he knows that B would have no incentive to fulfill such a promise. After all, once tomorrow comes, the incentives B will have are going to depend on the actual distribution of power that obtains, not on the past promises she might have made. After the power shifts in her favor, B's incentives are clear: she is better off fighting than agreeing to a share that gives her less than the expected payoff from war under the new distribution of power.

By analogous logic, the largest concession B would be willing to make today to preserve the peace must be such that she is at least as well off as fighting at the present distribution of power. (If she is strictly better off with peace, then she could potentially give A an even larger share without violating her own incentive to remain at peace.) Thus, the *minimal* terms that B would require today are $W_{\mathrm{B}} = 2(1 - p_{\mathrm{H}}) - c_{\mathrm{B}}$. Let x denote A's share in the concession she makes today. As we have seen, the *best* terms she can offer A in addition to this concession would leave her with $1 - p_{\mathrm{L}} - c_{\mathrm{B}}$ tomorrow, so the overall payoff of conceding x now and offering A the most she can credibly commit to becomes $2 - x - p_{\mathrm{L}} - c_{\mathrm{B}}$. Actor B can therefore offer any x that satisfies $2 - x - p_{\mathrm{L}} - c_{\mathrm{B}} \geq W_{\mathrm{B}}$. We conclude that B would have to offer

$$x \leq 2p_{\mathrm{H}} - p_{\mathrm{L}}. \tag{1}$$

This is the *maximal* concession that B will be willing to make today if she will also offer the maximal concession to A tomorrow. It is worth noting that this constraint might not be binding: if $2p_{\mathrm{H}} - p_{\mathrm{L}} \geq 1$, then B would be happy to give A the *entire* benefit today in

order to avoid war while she is still weak. To give a specific numerical example, suppose $p_H = 3/4$ and $p_L = 1/3$. The constraint on the concession is $7/6$, and B can offer $x = 1$.

Turning now to A, we know that he will be induced to eschew war today only if the benefit of peace is at least as good as fighting. With B offering x and no more than $p_L + c_B$ tomorrow, A's payoff from remaining at peace is $x + p_L + c_B$. This is at least as good as war if $x + p_L + c_B \geq W_A$, or if

$$x \geq 2p_H - p_L - (c_A + c_B). \tag{2}$$

This is the *minimal* demand that A will make in order to stay at peace today and receive the best possible deal tomorrow.

We can put the two constraints together and observe that a concession x would satisfy both simultaneously only if

$$x \in [2p_H - p_L - (c_A - c_B), 2p_H - p_L]. \tag{3}$$

The constraint on the maximum concession by B in (1) is strictly larger than the constraint on the minimum terms for A in (2) for any strictly positive costs of war. This implies the interesting (and important) conclusion that in principle there always exists a share of the benefit that B would be willing to offer such that A would prefer to agree to accept it, remain at peace, and allow the power shift to occur. The question now becomes: is such a concession feasible? The most B can concede at any point in time is the entire benefit, so we know that $x \leq 1$ must hold. Although the interval in (3) always exists, there will be no feasible concession that belongs to that interval if its lower bound exceeds the maximum concession that B can physically make. In other words, if it is the case that $2p_H - p_L - (c_A + c_B) > 1$, then there is *nothing* the B can credibly offer that would satisfy A's minimal demands today. Using our numerical example, if the total costs of war, $c_A + c_B$, are less than $1/6$ of the benefit, then the condition will be satisfied and B will be unable to induce A to stay at peace today. A wages **preventive war** in order to avoid the unpleasant consequences of a decline in relative power.

It is important to realize that this mechanism has an important nuance often not appreciated in discussions of preventive war: the declining actor fights today not because he is afraid of fighting after the power shifts in favor of his adversary, but because he is afraid of the unattractive peace he will have to live with after that happens. Because of this, arguments about the likelihood that the opponent will, in fact, fight tomorrow are beside the point. One cannot argue against this type of preventive war by asserting that the shift is irrelevant since B will not fight after it occurs. As we have seen, whether B fights or not depends on the terms A is willing to offer. What we have seen, however, is that the maximal concession B can make tomorrow is just not going to be good enough for A from today's vantage point even though it would be if peace prevails and that tomorrow comes.

To understand the condition that leads to the breakdown of peace today better, we can rearrange the terms as follows:

$$\underbrace{\underbrace{p_H}_{\substack{\text{chance of victory}\\ \text{while strong}}} + \underbrace{(p_H - p_L)}_{\substack{\text{net gain from fighting}\\ \text{while strong}}}}_{\text{total benefit from fighting while strong}} > \underbrace{\underbrace{1}_{\substack{\text{entire benefit}\\ \text{today}}} + \underbrace{(c_A + c_B)}_{\substack{\text{entire surplus}\\ \text{tomorrow}}}}_{\text{largest total benefit from peace}}. \tag{CCP}$$

This simply states that if the total benefit from war exceeds the total benefit that A can credibly expect from peace, then war must occur. The one term that might need an explanation is $c_A + c_B$. Recall that for any given distribution of power p, the bargaining range is always defined as $[p - c_A, p + c_B]$. The size of that range is simply $p + c_B - (p - c_A) = c_B + c_A$. We call this the **bargaining surplus** because it is the size of the benefit that the actors can collectively save from destruction by choosing not to fight; it is the peace "surplus" over the total collective benefit from war. By remaining at peace, the actors will collectively save their combined costs of fighting, which is exactly what the term represents. The largest concession one actor can make to the opponent is to offer the opponent his minimal terms plus the entire surplus: doing so would leave the conceding actor with a share equivalent to her minimal terms. This is why B's giving the entire future surplus to A is the best credible promise she can make.

The most important feature of condition (CCP) is the *size of the power shift,* $p_H - p_L$: the difference between the distribution of power today and the one that will obtain tomorrow after the power shifts in favor of B. If the change is small, then the condition will not be satisfied, and war will not occur. To see this in the clearest case, suppose that power did not shift at all, so $p_H - p_L = 0$. Obviously, the condition for war cannot be satisfied because it reduces to $p_H > 1 + (c_A + c_B)$ but we know that $p_H < 1$. This is why the more precise definition of this mechanism is that **large shifts in power** can cause war, and (CCP) specifies exactly how large the shift has to be for that to happen.

This now tells us that if the actors were to behave *as if* the shift had not occurred, then war would be avoidable. To see what I mean, suppose that B promises to ignore the power shift in her negotiations tomorrow; that is, she commits to bargaining with A as if the distribution of power is still p_H instead of p_L. With this commitment, the size of the power shift goes to zero, and (CCP) fails; war today is avoided. Making such a commitment is certainly in B's interest today because it enables her to avoid a very unattractive war. The problem is that she cannot credibly commit to making good on such a promise — the incentives she will have tomorrow are such that she will not want to abide by any such promise, not matter how much she wishes she could do so today. The issue is not that B might be lying — she quite sincerely *wants* to make such a commitment — the issue is that she will not have the incentive to follow through, and both actors know it. This is why the mechanism that explains war as caused by large power shifts is often called the **credible commitment problem** explanation. This is why we use that acronym for condition (CCP).

3.2 Why Actors Cannot Commit: Anarchy

Why is it that groups cannot generally be trusted to abide by their promises in the sense that the only credible promise (or threat) is assumed to be the one that the actor would be willing to follow through on *given his incentives at the time he must act on it*. Why it is so often that promises in international relations do not seem to be worth the paper they are written on? One possible answer is that because there often does not exist an overarching authority that can enforce such implicit or explicit contracts. This is why many scholars argue that international relations occur in a context of **anarchy**. This does not mean chaos. It means that the international environment does not have an entity to force actors to fulfill their promises (or threats) when it is not in their interest to do so. There is no world government

to provide enforcement analogous to what we have at the domestic level where the police enforces the rulings of courts. In these case, the only enforcement must be provided by the actors themselves. When the incentive to fulfill the terms of one's promise exist, then it will be in the actor's interest to follow through, providing the **endogenous enforcement** of the terms, and rendering the promise credible.

There are some ways of making some promises stick. For instance, one group can "force" the other to cooperate by threatening to withhold cooperation in the future — this can work when both actors do care about their future interactions and so the weight of the cooperative behavior that would be foregone by the failure to cooperate today can exceed any temporary gain from exploiting the other. This sort of self-enforcement will be much less useful, however, when it comes to the types of issues where the use of military force, and with it the hope for a permanent settlement, becomes a possibility. In these contexts, the only way for one actor to make another abide by his promises is by threatening to fight if he fails to do so. In the context of anarchy, the use of force is always an option in disputes because nobody has a monopoly on the legitimate use of force the way a government (usually) does. But this is precisely the problem we have explored so far: our model assumes that the only thing actors can guarantee themselves is what they expect to secure by the force of arms.

It might be interesting to compare briefly the credible commitment problem with the mutual optimism explanation. The obvious difference is that the commitment problem arises when actors are fully informed about the state of the world, and thus, strictly speaking, uncertainty is not a necessary cause of war. A more subtle difference is that under mutual optimism, actors fight because they hope that their opponent is weaker than the opponent seems to believe. Fighting corrects the mistaken expectations as groups learn from the developments on the battlefield, and once they become sufficiently convinced that the opponent is strong, they make peace. In contrast, the commitment problem arises when the declining actor fears that its opponent will become very strong: if the power shift is small, then a bargain can still be struck. This suggests that uncertainty about the size of the power shift might promote peace because an actor who faces the prospect of either a serious (war-inducing) decline or a mild (war-avoiding) one might well take its chances with peace provided it places enough weight on the latter possibility. Thus, the more optimistic the declining actor is, the less likely is war to occur. This highlights an interesting tension between the two explanations. Under mutual optimism, war occurs because both sides have unreasonably high expectations about what war holds in store for them. Under the commitment problem, war occurs because one actor is very pessimistic about its military prospects in the future.

3.3 The Role of Fighting

You might have noticed that this mechanism also models war as *absolute*: once it begins, it is fought to the end without any possibility to re-negotiate terms. We can modify the model to accommodate the notion of *ideal* war as follows. Suppose that when actors negotiate and choose whether to fight, fighting does not result in an *absolute* war but in a military engagement that might end in stalemate. If that stalemate occurs, they get to negotiate and choose whether to fight again. If they choose not to fight, they enjoy the peaceful distribution of the benefit for one encounter, which also allows them to consolidate whatever gains

they have made. The actors can then renegotiate the terms in their next encounter or fight another battle. To introduce a commitment problem, suppose that peaceful consolidation permanently shifts power in favor of one of the actors but that it takes some time to achieve. This means that when actors decide whether to allow peace to happen, one of them has to worry that the opponent would use the peace to gain a military advantage, which will then be used to extract more concessions. The process continues indefinitely: bargaining can be interrupted by occasional bouts of fighting.[11]

Fighting a battle has two effects: (a) it gives both actors an opportunity to reach a decisive military victory and enjoy the benefit unmolested; (b) it slows the pace of consolidation, which means it affects the rate with which power shifts. When power shifts more slowly, actors have opportunities to accommodate the changes in the expected benefits from fighting without resorting to force. The logic can be illustrated by referring to our original model of credible commitment. Recall that $p_{\mathrm{H}} - p_{\mathrm{L}}$, the size of the power shift, was the key component of the mechanism that lead to bargaining breakdown. Suppose now that instead of occurring all at once, it happened more slowly and actors could negotiate after each change. For example, suppose they while the power shifts they have $n > 1$ opportunities to bargain and these are equally distributed in time. The rate of the power shift is then simply the amount with which the distribution power changes with each bargaining encounter: $\triangle = (p_{\mathrm{H}} - p_{\mathrm{L}})/n$. Since $n > 1$, the amount power shifts after each encounter is smaller than the completed power shift, which means that condition (CCP) will not hold when $\triangle$ is small enough. In other words, power shifts slowly, actors would be able to accommodate the resulting smaller changes and avoid fighting. This is why scholars usually say that the commitment problem arises when there are **large and rapid power shifts**.

The logic now readily extends to the *ideal* war model if one assumes that fighting can slow down the process of consolidation: power shifts at a declining rate. If peace would allow power to shift too quickly, the declining actor would fight. If the battle does not end with a decisive victory, some consolidation will occur but since the rate is assumed to be declining, at some point the remaining size of the power shift will be small enough to allow the actors to negotiate a peace even though such a peace would result in the complete consolidation of one of the opponent. Thus, actors fight in order to forestall adverse shifts in power. When the shift slows down enough, peace becomes possible.

Another approach to modeling *ideal* war is to assume that actors have finite military resources and those are depleted while fighting continues. War then does two things — it reduces the size of the benefit (as before) but it also limits the effort that actors can dedicate to fighting. Consider now a situation in which actors can negotiate a distribution of the benefit but that after the settlement is implemented (or if no agreement is reached), they get to decide whether to engage in a military contest (a battle, or an engagement). If an actor attacks the other, then a battle is fought, and it can either end in a decisive military victory for one of the actors (like our simple war) or it can end in stalemate. A decisive military victory ends the war in the usual way: to the winner go the remaining spoils. A stalemate enables the actors to negotiate again, and then decide whether to fight another battle, and so

[11] This section is somewhat loosely based on two articles. Robert Powell. 2006. "War as a Commitment Problem," *International Organization,* 60(Winter): 169–203. Robert Powell. 2012. "Persistent Fighting and Shifting Power," *American Journal of Political Science,* 53(3): 620–37.

on. The process of bargaining and fighting continues until one the following happens: (a) one of the actors defeats the other with a decisive battle; (b) one of the actors collapses from attrition of his resources; or (c) neither chooses to attack once a settlement is reached.[12]

We next introduce the potential for power shifts in this *ideal* war model by assuming that if an actor surprises the opponent — that is, attacks when the opponent does not expect him to — he is more likely to win that particular battle (not the war). This means that a sneak attack creates a power shift in favor of the attacker. Even though this advantage is temporary, it does offer a hope for a permanent resolution if the battle turns out to be decisive. Peace requires not merely negotiating a distribution of benefits that both actors prefer to war that they expect to occur, but one that both prefer to launching a surprise attack on the opponent.

How can actors confidently expect peace in such a scenario? If the sneak attack succeeds, the winner gets to enjoy the entire of the surviving benefit unmolested, which gives the incentive to violate the peace. The disincentive to doing so must therefore arise from the consequences of failing to win that battle. The worse the consequences, the greater the disincentive to launch a surprise attack. Since war in this model is a sequence of decisions to fight battles, the *greatest disincentive* to violate the peace would be to threaten to fight an *absolute* war if that happens. In other words, if an actor violates the peace by launching a surprise attack, the fighting will continue until either one of them is decisively defeated in battle or collapses from attrition; there will be no further negotiations. *Mutual deterrence is best sustained with threats to fight absolute war.*

Early in the war, the total size of the benefit not yet destroyed is very large, and so the advantage of the power shift from a sneak attack is fairly substantial, making early peace very unlikely. Even threats to fight to the finish will not be able to deter sneak attacks. However, as the damage accumulates, the expected benefit from a surprise attack declines, and so the threat of absolute war can eventually provide a sufficient disincentive. This makes peace with threats of absolute war possible, and the fighting can end. The very destructiveness of war reduces the size of the benefit from the power shift, and opens up the road to war termination. In this way fighting can resolve the commitment problem that caused the war in the first place.

So far, so good, but we need to ask: are threats of absolute war credible? No, they are not, at least not in general, and here's why. Suppose that an actor launches a surprise attack but neither wins a decisive victory. According to the threats that were supposed to sustain the peace, they must now continue the war until the very end; they must fight an *absolute* war. Both know that such a war is going to be very long and very costly, so both would prefer to settle it sooner. Of course, they have to face the commitment problem yet again in the new negotiations, but in general they will have opportunities in which they can negotiate sustainable terms before one of them is disarmed. Since they have a mutual interest in avoiding total war, this means that when such an opportunity presents itself, the actors will take it, and the fighting will end short of one of the actors getting disarmed. But if this is true, neither actor should expect to fight an *absolute* war — they should expect only to fight until the first opportunity to negotiate peace; that is, fight an *ideal* war. Thus, the only

[12]This section summarizes the article by Bahar Leventoğlu and Branislav L. Slantchev. 2007. "The Armed Peace: A Punctuated Equilibrium Theory of War," *American Journal of Political Science,* 51(4): 755–71.

credible threats they can make involve fighting until such an eventuality, not until one of them is disarmed. In other words, *threats of absolute war are generally not credible, and the credible threats of ideal war are weaker.*

But if threats of absolute war are not credible, they cannot be expected to produce mutual deterrence; i.e., they cannot sustain the peace. Since threats of ideal war are credible, they can sustain mutual deterrence, but because they are weaker, they cannot do so in all circumstances where threats of absolute war would have worked. In other words, the actors would have to fight longer before they can reach a window of opportunity for peace supported by threats of ideal war. This point is worth emphasizing: *if actors could commit to punish violations of an agreement with an absolute war that can only end in disarming, peace would be easier to achieve than if they commit to fight an ideal war that can end with a peace settlement.* The irony is that despite the desirability of making such absolute war threats, they actors cannot credibly do so precisely because peace is so desirable. This is the **paradox of wanting peace**: the very desirability of peace makes war more likely.

These models also imply that commitment problems might be very difficult to resolve, and as such might even be a more persistent cause of war than mutual optimism. War occurs because of the expectations that a large and rapid power shift can create, and fighting continues in order to change the environment so that either the size of the effect or the speed of that shift are minimized.

3.4 Sources of Power Shift Anxieties

The commitment problem can arise from various factors. One reason that appears quite often in both historical works and in justification for military action by politicians is the fear of decline relative to the power of an opponent. This is said to trigger a preventive war. Historian A.J.P. Taylor has claimed, with some exaggeration, that

> Every war between Great Powers [between 1848 and 1918] started as a preventive war.[13]

This type of war arises from an attempt to forestall an adverse shift in power, and so important variables to consider would be the perceptions of relative decline that might be due to technological and economic development of the opponent that one cannot hope to match, along with perceptions of hostility that support the expectation that the opponent will, in fact, use its newly acquired powers against one's interests. This perception can arise out of one's view of the opponent's political system or society if that view attributes to them militarism or aggressive intent.

Perhaps the most famous (and maybe the earliest) statement of this logic can be found in Thucydides' explanation of why Sparta decided on war with Athens in 431 B.C. Using a squabble between Athens and its ally Corinth as pretext, Sparta declared that the Athenians had violated the Thirty Years' Peace (which had only lasted thirteen years), and effectively declared war. As Thucydides puts it,

[13] A.J.P. Taylor. 1980. *The Struggle for Mastery in Europe, 1848–1918.* Oxford: Oxford University Press, p. 166. For a useful discussion of preventive war and different "better-now-than-later" reasons to fight, see Jack S. Levy. 2008. "Preventive War and Democratic Politics," *International Studies Quarterly,* 52: 1–24.

> The Spartans voted that the treaty had been broken, and that war must be declared, not so much because they were persuaded by the arguments of their allies, as because they feared the growth of the power of the Athenians, seeing most of Hellas already subject to them.[14]

The growth in question was happening because the commercial city of Athens had been free to trade after the expulsion of the Persians from Greece. The Athenians had also rebuild the *Long Walls* that made the city impregnable to a land assault and ensured its supply from sea through their enclosure of the connection to the port of Piraeus. This action vexed the Spartans because it effectively neutralized their formidable land army, and they had no navy with which to threaten Athens from the sea.[15] The Athenians cleverly used their wealth and the interminable inter-city fights among the Greeks to bully and attract many, increasing the number of tribute-paying members of the Delian League. This tribute also swelled the flow of money into Athenian coffers, funding further expansion and public works in the city. All of this caused the reclusive Spartans to lose influence in Greece. The long-term trend was unmistakable: should Athens be permitted to continue its policies unchecked, the distribution of power would shift away from Sparta, probably decisively and permanently. This led Thucydides to interpret the second Peloponnesian War as essentially a preventive war by Sparta against Athens.

One concept closely related to preventive war is that of **preemptive war**, which differs from preventive war merely in that the opponent is perceived as poised for an imminent attack and there is some advantage to be had in striking first. The commitment problem arises from the opponent's inability to promise credibly not to use the advantage of striking first. Since doing so creates an instantaneous power shift in the opponent's favor, one might be tempted to attack in order to prevent that from happening. Thus, the underlying the logic is exactly the same as the one we have been exploring, and the difference from prevention is only one of timing.

For this trigger, one might look at military technology. One possibility is a technology that gives a decisive advantage to striking first. For example, consider the nuclear balance between the United States and the Soviet Union from 1949 until the mid 1960s. During that period both sides possessed nuclear weapons but had neither the numbers nor the defenses to ensure their protection from a surprise first strike. Whoever struck first could, in principle, completely disable the nuclear capability of the other.[16] Thus, if striking first would give the attacker probability p_{H} of prevailing, allowing the opponent to strike first would create an instantaneous power shift in favor of the opponent, so one's probability of winning would immediately drop to p_{L}. If the advantage of striking first is large enough, then this shift would create a commitment problem and cause war.

---

[14] Thucydides. 1996. *The Landmark Thucydides: A Comprehensive Guide to the Peloponnesian War, edited by Robert B. Strassler.* New York: Touchstone, p. 47 (1.88).

[15] After nearly thirty years of fighting, the Spartans finally managed to assemble a fleet with which they successfully blockaded Piraeus, forcing Athens to surrender in 404 B.C. After their victory, the Spartans promptly destroyed the walls.

[16] Only with the advent of nuclear submarines did survivability become feasible. Eventually both sides developed *second-strike capability:* they could absorb a surprise first strike but would still have enough surviving nuclear weapons to launch a devastating retaliatory strike. The era of *mutually assured destruction* had arrived.

A less apocalyptic scenario involves the ability to achieve rapid concentration of one's military force and defeat the opponent before he is fully prepared to engage. For example, in the 19th century, Prussia's military organization was local — reservists lived only at most a few hours away from where they were supposed to go when called to arms, which meant that Prussia could mobilize its army very rapidly. In contrast, Austria's military organization was national — reservists were deliberately assigned to depots in different parts of the country in order to minimize the probability that they would join local rebellions. This meant that Austria's mobilization would be much slower once the call to arms was given. By 1866, Prussia's railway system was far better developed than the Austrian and had far more links to the territories where fighting could occur. This meant that once mobilized, Prussia could move its armies faster and supply them more reliably than the Austrians could theirs. Even though the Austrians had more allies (almost all German states sided with them against Prussia), and even though their combined resources were greater, Prussia's superior ability to concentrate its forces created a power shift in Prussia's favor. Since this advantage would be eroded if Austria were given the time to mobilize properly, Prussia had to strike preemptively. Even though Austria declared it, it was the Prussians who won the Seven Weeks War of 1866.[17]

I should note that whereas it is perhaps easier to justify preemption on the ground that war is inevitable anyway, it is much harder to justify prevention, which, after all promises the certainty of war today in response to a possibly vague threat in the future. As the German Chancellor Otto von Bismarck, who had presided over the Prussian victories in 1866 (over Austria) and in 1871 (over France), once told the Reichstag,

> [Waging preventive war because] it is possible that in some years we might be attacked... [is like committing] suicide from fear of death.[18]

This is not to say that such a war is impossible. As we shall see when we discuss the First World War, there is a well-established tradition of arguing that Germany's attack was predicated on a preemptive logic: the failure to beat Britain in the arms race or the rising power of Russia are said to have contributed to the decision to fight before it was too late. Ironically, some have also blamed Germany's own rise after the Unification as a contributing factor.[19]

The anxiety implicit in the notion of a preemptive war can arise from an escalating sense of **mutual alarm** caused by military movements or arming decisions made for essentially defensive purposes by the actor. The problem with arming is that often the type of weapons

[17] France supported Prussia in this war out of desire to reduce the influence of the Habsburgs in Europe but it soon came to regret it. In 1870, France also declared war on Prussia but was just as unprepared as Austria had been. The Prussians mobilized and utilized their superior railways before the French were even half-ready to meet them. Having achieved rapid concentration, Prussia invaded France and defeated the Emperor at Sedan.

[18] "Es ist möglich, dass wir in einigen Jahren einmal angegriffen werden, damit wir dem nun zuvorkommen, fallen wir rasch über unsere Nachbarn her und hauen sie zusammen, ehe sie sich vollständig erholen — gewissermaßen Selbstmord aus Besorgnißvor dem Tode." Reichstag prorocols, 1875/76, 2, pp. 1329-30 (February 9, 1876). Online in German at `http://www.reichstagsprotokolle.de/Blatt3_k2_bsb00018381_00571.html`, accessed December 28, 2012.

[19] Niall Ferguson. 1999. *The Pity of War: Explaining World War I.* London: Penguin Books. James Joll. 1992. *The Origins of the First World War.* London: Longman.

one acquires can be just as useful in attacking as they are in defending.[20] If their purpose is not neatly delineated — and it usually is not — then each actor must rely on its "guesstimate" of the opponent's intentions. When one group feels threatened by another, it can arm itself to maintain a more favorable distribution of power. In doing so, however, it might make the other group feel threatened in turn (after all, there is no way of knowing whether one is arming for defense or in preparation of attack), and it might respond by increasing its own arming. This action can in turn feed back into the perceptions of the first group, possibly solidifying its view of the opponent as hostile and increasing their suspicions and fears. The first group responds by increasing its own arming, triggering another feedback into the perceptions of its opponent, and so on. This arms race is accompanied by increasing anxiety and might lead to a preemptive strike if there are advantages of moving first.[21]

Finally, commitment problems arising from rapid power shifts can also be caused when the disputed benefit cannot be easily divided without affecting its value. For example, consider a mountain with a single pass. Whoever controls the pass can have a large military advantage. If he is attacked, the mountain is easier to defend. If he attacks, he will not have to get through a heavily defended mountain. The benefit cannot be shared because the moment one actor takes possession of the pass, the power shifts in his favor, creating the potential for a commitment problem. We shall return to this issue in the last section, where we shall also consider other sources of "indivisibility" that might make the benefit more difficult to divide in ways that can satisfy the war expectations of both sides.

4 The Costly Peace

We now arrive at the final explanation for war that we shall consider.[22] One fundamental assumption of the bargaining approach to war is that fighting is a very risky and costly way of resolving the dispute relative to peace. However, we have also seen that peaceful resolution depends on the implied threat of war, which determines the bargaining range and so delimits the set of mutually acceptable peace deals. This implies that force plays an implicit role in the maintenance of peace — bargaining takes place in the **shadow of power**. What the model neglects is that military power is not free — maintaining sufficient forces to ensure a relatively attractive distribution of power and thus a preferable distribution of the benefit entails costs that must be paid regardless of whether the military is ever put to actual

[20] Even military installations that appear to be solely defensive — like a castle — might be perceived as offensive if they enable the opponent to secure a line of communication, a base from which to launch an attack, or provide for the defense of one territory so that it can free its hands to attack another.

[21] The act of making a military move designed to enhance one's security but which, in fact, might well end up worsening it because it triggers a counter-move by the opponent is called the **Security Dilemma**. The escalation logic of mutually reinforcing anxieties is sometimes called the **Spiral Model of War**. See Robert Jervis. 1976. *Perception and Misperception in International Politics.* Princeton: Princeton University Press, Chapter 3.

[22] There are others, many others, actually. Unfortunately, most of them tend to be of limited value because they rest on undeveloped foundations and the arguments are often internally contradictory. Exploring these issues is well beyond the scope of this course. Interested students are encouraged to take the course "Causes of War" or at least read a useful overview like the one provided by Jack S. Levy and William R. Thompson. 2010. *Causes of War.* Chichester: Wiley-Blackwell. For the (disturbingly modest) statistical correlation between many of the supposed causes and war, see the study by D. Scott Bennett and Allan C. Stam. 2004. *The Behavioral Origins of War.* Ann Arbor: The University of Michigan Press.

use. The implicit use of force in peace requires that actors pay the costs of the upkeep of the military that underpins the distribution of power on which the distribution of the benefit relies. In other words, peace is costly too.

These costs can be various: taxes raised or debt incurred to pay for the military, income diverted from other uses to pay for the military, inflationary debasement of the currency to facilitate payment for the military, wealth exported in the form of subsidies to allies, withdrawal of manpower from the economy especially during mobilizations, and economic dislocations resulting from the favored treatment by the government of some sectors of the economy at the expense of others, or the social and political implications of direct government intervention in the economy. All this expenditure of resources means that the group must forego other desirable goals (e.g., investing in economic development, civil infrastructure, social security, health care, and so forth), and the long-term cumulative effect of maintaining one's formidable military power might be quite devastating to the overall well-being of the group.[23]

These costs also help explain why conflicts over "indivisible" issues might be prone to escalating to war. With claims of indivisibility keeping the hostility alive, any sort of shared arrangement must be maintained by the implicit force of arms: the groups must essentially mutually deter each other from attempting to seize full control of the issue (e.g., a sacred place). The long-term costs of maintaining sufficient deterrent capability might outweigh the short-term costs of a war that might secure the place for a long time.

4.1 How Peace Can Be Worse Than War

To see now how the costs of peace might cause war, consider a slightly modified version of our original model. As in the commitment problem, the actors interact twice and they have full information about everything. We now assume that before each interaction they simultaneously decide whether to arm or not. Arming is costly: A pays $k_A > 0$ if he chooses to arm, and B pays $k_B > 0$ if she chooses to arm. Arming confers an advantage in the distribution of power when the opponent does not arm. We shall assume that if both players arm or if neither one arms, the distribution of power is such that each has an equal chance of winning. (This assumption is immaterial but makes exposition cleaner.) If A arms but B does not, then the distribution of power, p_H, favors A. If, on the other hand, B arms but A does not, then the distribution of power, p_L, favors B. Since we wish arming to confer an advantage to the player that arms, assume that $p_L < 1/2 < p_H$. To complete our assumptions, we shall specify that when the bargaining range exists, the actors divide the bargaining surplus evenly; that is, each actor obtains its minimal terms, and they split the rest 50-50.[24]

As before, when actors decide what to do in the first interaction they have to take into account the consequences of their actions for the second interaction. If they fight, the winner

[23] One reason for the collapse of the Soviet Union was that its economic system could not withstand the heavy defense burden imposed by the arms race with the United States. With inefficient production and shortfalls in agriculture, the USSR was increasingly reliant on borrowing from the West to pay for imports of foodstuffs. This directly curtailed its ability to act internationally but was also unsustainable in the long run. The attempt to reform the economy, however, unleashed forces that unseated the political system. See Yegor Gaidar. 2007. *Collapse of an Empire: Lessons for Modern Russia.* Washington: Brookings Institution Press.

[24] For those interested in these things, this division is called the Nash Bargaining Solution.

		B: *Arm*	B: *Do not arm*
A	*Arm*	$\frac{1}{2} + z - k_A$ $\frac{1}{2} - z - k_B$	$p_H + z - k_A$ $1 - p_H - z$
	Do not arm	$p_L + z$ $1 - p_L - z - k_B$	$\frac{1}{2} + z$ $\frac{1}{2} - z$

Figure 3: Payoffs from arming and bargaining in the second encounter.

locks in the possession of the entire benefit and there is no more bargaining since there is no more opponent to contest it. If they initially negotiate a peaceful division, they must negotiate again in the future.

Consider now that bargaining in the future. We need to consider four possibilities, depending on which actor arms and which actor does not:

1. Neither actor arms. The probability that A wins a war is $1/2$ and nobody pays any additional costs. The bargaining range is the set of deals in $[1/2 - c_A, 1/2 + c_B]$, and the mid-point (and thus A's share) is $1/2 + z$, where $z = (c_B - c_A)/2$. Actor B gets the remainder: $1/2 - z$.

2. Both actors arm. The distribution of power remains the same but the costs must be paid regardless of whether they negotiate a peace deal or fight a war. Thus, when they both arm, the bargaining range is the same as in the case where neither armed but their payoffs are reduced by the cost of arming.[25] The payoff for A is the share he obtains net the arming costs: $1/2 + z - k_A$. Analogously, B's payoff is her share net her arming costs: $1/2 - z - k_B$.

3. A arms but B does not. The bargaining range is $[p_H - c_A, p_H + c_B]$, whose midpoint (and thus A's share) is $p_H + z$. Actor A's payoff is this share net the cost of arming: $p_H + z - k_A$, whereas B simply obtains the remainder of the benefit without paying additional costs: $1 - p_H - z$.

4. B arms but A does not. The bargaining range is $[p_L - c_A, p_L + c_B]$. Actor A's payoff is simply his share, $p_L + z$, whereas B's payoff is the remainder of the benefit net her arming costs: $1 - p_L - z - k_B$.

Figure 3 shows these payoffs in a convenient table form with A as the row actor and B as the column actor. Within each cell, A's payoff is in the north-west corner, and B's payoff is in the south-east corner.

To simplify the analysis and illustrate the basic point, we shall assume that the costs of arming are not very large, at least not relative to the benefit. In particular, we shall

[25] You can verify this by taking one of the actors, say A, and noting that the expected war payoff is now $W_A - k_A$, whereas the peace payoff from some deal x is $x - k_A$. Actor A will only agree to deals that are no worse than war, or $x \geq W_A$, which yield the same minimal terms as in the case where neither actor arms.

assume that these costs are smaller than the extra share the player can secure by arming irrespective of what the opponent does. Mathematically, for A's arming costs we assume that $k_A < p_H - 1/2$ and $k_A < 1/2 - p_L$; and the assumption for B's costs is analogous. This assumption means that each actor is always strictly better off arming regardless of what the opponent does. The intuition is that if the opponent does not arm, then the actor prefers to pay the cost of arming in order to extract a larger share of the benefit in the bargain. If the opponent does arm, the actor prefers to arm to avoid conceding a larger share to the opponent.[26]

The arming and bargaining interaction can now easily be analyzed. Since each actor is better off arming irrespective of the arming choice of its opponent, both actors will arm.[27] The actors will then bargain peacefully and conclude a negotiated deal. The payoffs are listed in the highlighted (north-west) cell of the table in Figure 3. It is worth noting that the individual incentive to take advantage of the opponent's failure to arm results in an outcome that is worse for both actors. When they both arm, their shares of the benefit are exactly the same as when neither arms (the payoffs in the south-east cell of the table) but they both pay the costs of arming. They cannot disarm, however, because neither can trust the other not to renege on any such agreement: after all, when an actor expects the opponent to disarm, he would rather arm and obtain a much better peace deal. The other point worth remembering is that the outcome of the interaction is nevertheless peaceful: the actors do bargain in the shadow of power but this does not cause war in the second encounter. We conclude that if the first encounter were to conclude peacefully, the actors will arm and negotiate a peace deal in the second encounter as well.

We are now ready to analyze the first encounter. Since the second interaction always ends the same way regardless of what happens today as long as the outcome is peaceful, the actors can simply focus on obtaining the best possible deals today. But since the context is exactly the same as in the second encounter, we know that whenever peace is the outcome, the actors would both arm and negotiate the deal that splits the bargaining range between them in equal shares. In other words, the peace outcome of the first encounter involves the same payoffs as the peace outcome in the second. Since the total payoff is simply the sum of the payoffs from each interaction, we conclude that if the actors were to negotiate peacefully in the first period, the payoff for A would be $1 + 2(z - k_A)$, and the payoff for B would be $1 - 2(z + k_B)$. Would the actors accept such a peace or would they fight?

Just like arming is preferable when the actors expect peace to prevail, so it is when they expect war to occur. This means that the only alternative we need to consider is when they both arm but instead of negotiating peacefully, they go to war. Since victory eliminates the opponent, the winner can enjoy the entire benefit twice but only pay for arming once. When both arm, each actor expects to win with probability $1/2$, so the expected war payoff for A is:

$$W_A = \left(\frac{1}{2}\right)(1+1) - c_A - k_A = 1 - c_A - k_A,$$

whereas the expected war payoff for B is $W_B = 1 - c_B - k_B$. Actor A strictly prefers to fight a war when $W_A > 1 + 2(z - k_A)$, which reduces to $k_A > c_B$. Analogously, B strictly prefers

[26] For those interested in these things, the assumptions make this game a Prisoner's Dilemma.

[27] For those interested in these things, this is the unique Nash Equilibrium.

war when $k_B > c_A$. In other words, the groups strictly prefer to fight if their arming costs are large enough.[28] The upshot of this (somewhat involved) analysis is clear: sometimes paying to maintain a distribution of power that underpins an attractive distribution of the benefit in peace simply does not pay. *Peace by mutual deterrence might be too expensive to maintain relative to the possibility of a permanent settlement offered by war.*

4.2 The Role of Fighting

Although this is also a theory that uses *absolute* war as the alternative to peace, one can readily see how the argument would extend to an *ideal* war. The comparison between the burdens of peace and the benefits of continuing the war could happen at any point during the war when actors are considering possible termination. If actions taken during the war increase the costs of peace, then termination will become less likely. One such possibility is the effort to finance the war — a topic that we shall consider at great length in this course — through borrowing under limited liability. For example, if it is the case that an actor is more likely to repudiate debts when defeated in war than when either victorious or in peace (a reasonable assumption given how costly defeat could be), then the expected burden of repaying the debt is lower in war than in peace, which places heavier demands on the terms the actor would seek to secure in order to terminate the war. When the other side is unwilling to grant these concessions (this could happen for various reasons, one of which could be that it is also heavily indebted), then war termination will be unlikely. When peace is costlier than war, the bargaining range might not even exist, making these wars very difficult to end; after all, the problem is not one of locating a peaceful deal, the problem is that there are no such deals.[29]

4.3 Sources of Costly Peace

Some of the factors we would have to consider for this explanation are fairly obvious, like the expenditures on one's own military forces (usually in proximity to the opponent) or subsidies to allies paid to distract that opponent. This **armed peace** might also involve sanctions, conflicts by proxy, and incidental mobilizations to discourage the opponent's probes. Beyond that, there are the economic costs arising from lost trade. One might also figure the domestic costs of dealing with any possible interference by the opponent who might attempt to undermine one's rule by supporting rival claimants or encouraging secessionist movements or terrorist activities. The state of permanent military readiness could also entail societal costs as the government expands its reach into the economic and social structures, usurps political rights, and turns the polity into a "garrison state." Finally, the measures a group takes to deter another might endanger its relations with others by damaging their economic and security interests. In short, there is a long list of costs that

[28] You might recall that we initially assumed that the arming costs are not too large relative to the additional share of the benefit that arming can bring in. We are now saying that if they are sufficiently large, the actors would fight. This is not a contradiction but it does require attention to the configurations of the various parameters. For example, for war to occur, we require that $c_B < \min(p_H - 1/2, 1/2 - p_L)$. When this is satisfied, there always exist values for k_A that ensure that A prefers to fight. We can derive a similar requirement for B.

[29] Branislav L. Slantchev. 2012. "Borrowed Power: Debt Finance and the Resort to Arms," *American Political Science Review,* 106(4):787–809.

peace might entail, and their cumulative weight might well push a polity into an attempt to settle its differences with another by force and then enjoy peace unmolested.

5 Military Advantage, Sacred Land, and Divisibility

We have now developed a relatively sophisticated understanding of the factors that should be associated with using war as the instrument to achieve political objectives. The common thread to the explanations we have considered is that bargaining between competing groups takes place in the shadow of power, and so the terms each is prepared to concede or willing to demand are determined by their estimates of what war might hold in store, and what the consequences of peace would be.

An implicit, but very important, element of this analysis can be summarized colloquially as "it takes two to tango." This is the idea that for war to occur, both opponents have to "agree" to fight. This agreement might be quite unpleasant — if one is invaded, it does not appear that there are many choices left — but it is agreement nevertheless since there is always the option of conceding the opponent's terms without fighting. Thus, in a very important sense, wars are always *voluntary* — both actors must prefer to fight than to concede what the opponent demands. This makes wars a matter of choice rather than some apocalyptic inevitability of the human condition. Moreover, it points to the serious deficiency of explanations that rely on the aggressiveness of one actor to explain what is essentially a mutual act. In this view, it does not matter how evil Adolf Hitler or Saddam Hussein were — saying that the Second World War or the Iran-Iraq War were caused by their aggressive politics cannot amount to an explanation of these wars since it only considers why they demanded so much from their opponents. Without an explanation why their opponents preferred to fight rather than grant these demands, we cannot be said to have understood these wars.

Although we listed, and examined, the causes separately, in reality most conflicts would contain elements from several simultaneously. For example, here's a "perfect storm" scenario: a conflict in which a window of temporary vulnerability (optimism) of a usually powerful opponent who is not only expensive to deter (costly peace) but is perceived to be generally on the rise (commitment problem). I am not saying that these factors need to be objectively present: it is often enough that a strong perception that they do exists. I am also not saying that the actors that are claiming these perceptions are even sincere — they might well have other reasons to want a war and are using these arguments to buttress the cause of fighting. (This can easily happen when the people who stand to profit from the war are not the ones who are likely to bear its costs. For them, war might well be a profitable enterprise and the bargaining puzzle would not even arise.)

Going back to our original model, recall that any negotiated peace deal is supposed to allocate the benefit in shares that satisfy at least the minimal terms of both actors. One (unstated) assumption is that it is always possible to do so. Thinking of the benefit in terms of territory might appear intuitive because we can essentially draw borders wherever desired, but in practice things are not quite that convenient. Some shares of this territory might be more valuable than others in a way that cannot be shared by splitting the territory.

Suppose that the reason the territory is valuable to the polities is an oil field it covers, or access to a sea port that is especially desirable for commerce, or because it contains a

defense installation whose possession might be of important strategic value. Since the full value of the benefit is 100% (we represented this by assigning it a value of 1), any deal in the bargaining range must allocate some percentage of that. Suppose, for the sake of argument, that the bargaining range comprises any deal that gives A at least 45% and B at least 35% of the benefit. Were the benefit divisible, they could agree on some distribution that meets these requirements and avoid war. But what if the benefit is not divisible (or the possible divisions do not represent shares that fall in the bargaining range)? For instance, of what value is the control of half a defense line if your opponent controls the other half? How, exactly, would one split access to the port? How do you share access to a sacred site if your religion demands exclusive access? Some scholars have argued that this type of indivisibility can be a cause of war. The problem is the *physical or psychological impossibility of dividing the benefit in a way that satisfies the minimal terms of both sides.*

Now, some of the issues that are said to cause indivisibility can, at least in principle, be dealt with. Take the oil field and port examples. Although they cannot be split to satisfy the actors' terms directly, one can easily imagine an agreement that allocates the entire benefit to one of the actors who then transfers a portion of the income derived from it to the other. In the case of the oil field, the actors can even set up a joint stock company with mutual ownership in proportion to their minimal terms, and thus both would be better off even if the land nominally belongs to one of them. These sorts of agreements can be made to stick by the threat of war that the landless actor can make if the partner fails to live up to the terms. The treat is credible because in the absence of a transfer, the status quo benefit would be worse than fighting. In principle, at least, *some types of physical indivisibility can be overcome with cooperative arrangements or side-payments.*

This is not to say, however, that all physical indivisibility can be dealt with in such a manner. Consider the case of a defense installation whose military value is compromised unless it is used in its entirety. Alternatively, it could be that possession of a key piece of strategic territory is of considerable military advantage (e.g., a mountain would be of significant military value to the side controlling it either if one has to defend it or if one does not have to fight its way through it when attacking). This type of asset cannot really by shared because once one side is in possession the military advantage accrues to it immediately. The only way the other can hold it to any prior agreement is with the threat to fight but, by the definition of the properties of the asset, it must now do so under considerably less favorable circumstances.

One should recognize the logic of the commitment problem causing war here. Transfer of the valuable military asset creates a power shift in favor of the owning group, and if that shift is large enough, the inability to promise not to use it against the opponent might cause the opponent to fight instead of relinquishing it. In this case it is not really indivisibility that is the source of the problem but the combined effect of a large power shift and an inability to commit to promises. From this vantage point, *indivisibility is merely a manifestation of the commitment problem,* so we would not need to treat it as a separate cause of war.

This might lead one to conclude that when it comes to physical indivisibility, we have no new mechanism to explain war: it is either the case that there exist ways of providing compensation to the relevant actor without a physical transfer of the property rights, or the case that when such an arrangement is not feasible, the mechanism leading to war is already specified by the large power shift creating a credible commitment problem.

This conclusion might not yet be warranted, however, because some issues might be perceived as indivisible because of psychological factors even if they are physically divisible. That is, in order to provide benefits to the group, the issue must remain in the group's control in its entirety; whole or nothing! Consider **sacred spaces** — that is, landmarks or plots of land with clearly defined boundaries that are of spiritual importance to some group.[30] These spaces could be sacred for religious or for secular (e.g., nationalist) reasons but they all share in common the notion that they are of unique importance for the well-being of the group. To make things worse, some religions effectively prohibit sharing of such spaces by not simply requiring their exclusive use by the believers but sometimes even mandating the destruction of competing sacred spaces and their replacement with those that adhere to the faith. This clearly poses a problem for any sharing scheme along the lines we explored above for mundane income-generating assets like oil fields. It might appear that a conflict over a sacred space must either end in one group perpetually excluding the other by force when necessary or degenerate into a fight that eliminates one of the claimants.

The difficulty with this line of reasoning is that it does not look like it is historically impossible to share even supposedly indivisible sacred spaces. When neither group can reasonably hope to eliminate or exclude the other, the only possible solution is some type of accommodation. Part of the sacred space can be reserved for one group while another part for the other, or they could have access to the space on alternate days. At any rate, when the alternative is perpetual hostilities and the potential destruction of the site itself, even the most indivisible issue seems to become shareable, if not divisible.

This points to a fundamental problem with the indivisibility approach: the difficulty of separating expressions of genuinely indivisible preferences (sincere belief that the sacred space would be desecrated if the other group is not excluded from accessing it) from strategic expressions of such preferences designed to induce the other side to give up its claims. Given the wealth of historical precedent for sharing sacred spaces, presumably the most difficult to divide, one might well doubt whether genuinely indivisible issues really exist. In light of this, claims of indivisibility are to be taken with a huge grain of salt: they might be ruses that motivated groups use to mobilize support for their cause or to demand larger concessions from their opponents. This obfuscates the straightforward inference one is supposed to make from such claims and can cause actors to ignore such pronouncements much like they would ignore claims of strength in the asymmetric information problem we discussed above.

Despite all these caveats, claims to indivisibility should be taken seriously in our study of war and society. Even if the political leadership cynically maintains such claims for the sake of boosting its popular support or mobilizing support for its policies, the fact that they can do so implies that there is some resonance in the public when it comes to the issue and that the leadership can then behave as if there exists a de facto indivisibility. In other words, it might not be important whether they are sincere or not in their professed beliefs in indivisibility - it matters whether others believe it enough to enable their policies.

When one moves from absolute to idea war, it becomes very difficult to see how a conflict over a truly indivisible issue can possibly be resolved through fighting short of disarming the opponent or eliminating it altogether. One possibility is that one of the sides becomes

[30] Ron E. Hassner. 2009. *War on Sacred Grounds.* Ithaca: Cornell University Press.

convinced that the issue is not worth continuing to fight and give it up entirely. This would imply that the underlying cause was not so much the indivisibility as the inflated expectation of what can be achieved by war, which puts us back in the mutual optimism camp — indivisibility might exacerbate the fighting because now one side has to become very pessimistic about its chances in order to agree to give up the entire benefit; however, it does not appear to function on its own as a separate cause of war. Thus, we can say that in this model, war occurs because of either mutual optimism or a commitment problem, and fighting continues in order to resolve the cause except that indivisibility of stakes can prolong the process, making war termination harder to achieve.