

Introduction to International Relations

Lecture 7: Causes of War

Professor Branislav L. Slantchev

Department of Political Science, University of California - San Diego

April 29, 2005

Overview. We now seriously apply our bargaining ideas to the study of the causes of war. At the most basic level, we want to know if there exist conditions that would cause fully rational actors to engage in costly conflict. We find that there exist at least four such mechanisms: asymmetric information, time inconsistency, indivisibility, and costs of peace. We apply these ideas to the study of preventive war, preemptive war, power transition, and strategic pre-commitment in bargaining, among others.

OUTLINE OF LECTURE 7: THE CAUSES OF WAR

1. The Puzzle of War

- why rationalist explanations?
- war as a bargaining problem
- bargaining range always exists
- latent use of force to drive bargains

2. Rationalist Explanations for War

1. indivisibility of stakes

- shared control
- side-payments

2. asymmetric information

- private information
- incentives to misrepresent
- risk-return trade-off
- risks vs. mistakes
- war as a bargaining process
- wartime revelation of information
- convergent expectations
- two examples: Seven Weeks War (1866), Serbo-Bulgarian War (1885)

3. credible commitment problem

- time inconsistency
- preemptive war
- preventive war
- armistice and cease-fires
- ethnic conflict and disarmament
- power transition theory?
- two examples: Peloponnesian War (432 BC), Second World War (1939)

4. the costs of peace

- preventing power shifts
- example: China's intervention in the Korean War (1950)

Why do wars occur?

Our conclusion from last time is that the international system is structured in a way that creates permissive conditions for the occurrence of war. In the words of Kenneth Waltz, war occurs because there is nothing to prevent it. However, as we also noted further, even if the system creates such conditions, it cannot really explain why war sometimes breaks out and other times does not. One attempt to address this was to study the distribution of power in the system, but as we noted, such explanations have serious empirical deficiencies. As we also see today, they have even more serious theoretical problems.

Even if the other two images (state composition and human nature) can tell us quite a bit about particular instances (or at least seem that way), in general they cannot be very explanatory. After all, people are different and to argue that it takes the mistakes and stupidity of leaders to start the majority of wars is a bit far-fetched, not to mention unhelpful for analysis. If each war is idiosyncratic (that is, has causes that are unique to it), then we have no hope whatsoever of ever understanding the phenomenon of war.

We shall assume (or fervently hope) that there is something common to wars, if not all, then perhaps among groups of wars. We shall look for a common mechanism that could potentially generate war among competing rational actors.

Why do we insist on actors being rational? Note that we are *not* saying that there is no irrationality in the world. That would be stupid for of course there is. We are not saying that people do not act in anger without thinking through the consequences of their actions, for they do. We are not saying that people do not miscalculate and make grave mistakes, for they do. In other words, we are not saying that every war is a result of a fully rational decision on all sides.

What we are saying is this: *if even fully rational actors have incentives to engage in this enormously destructive behavior that is war, then our explanation of what causes war has passed the most difficult hurdle.* If anyone is to avoid the costly and risky fighting, it would be the fully rational actors that can perfectly analyze the situation. If even such imaginary actors are unable to avoid war, then war is indeed a possibility that will probably never disappear. Mistakes, miscalculations, and anger can all make war more likely (perhaps fear can make it less likely), and in an important way, all these factors make it easier to explain war.

We set for ourselves the harder task. We will not simply say “they fight because they hate each other” or “they fight because they made mistakes during the crisis and now they can’t help it.” Instead, we shall insist that our theoretical mechanism explains who actors who can analyze the situation correctly without being subject to distorting emotions can go to war with their eyes wide open. If we can explain that, then we are truly onto something.

The central puzzle about war is this: Since wars are costly, then why do rational actors engage in them? Clearly, one possible fully rational explanation is that actors fight because they like fighting. That is, they derive benefits from

the act of fighting itself. Although this would be a tenable explanation, it is not going to get us very far because in just about every case that I can think of, people have found war abhorrent. Yes, war used to be depicted in glorious terms, especially before the carnage of WWI. Still, because of the risks and costs involved, people would generally prefer peace to war if they can help it. Also, even if fighting can generate positive benefits for some actors, it is not at all clear why they cannot get these benefits in peace. For the rest of our discussion, we shall assume that actors do not like war for war's sake.

To summarize, we are looking for rational explanations for war because these should be the hardest to get. They will be fundamental in the very important sense that they will be permanently operating factors that no amount of goodwill, communication, or mutual understanding would be able to remedy. Of course, every real war would have some (or all) of these factors operating in addition to the usual plethora of bad judgment, stupidity, and heated feelings. These only make matters worse: wars are more likely to erupt, and more difficult to end.

1 The Puzzle of War

In some sense, a war must be wanted by both sides if it is to occur. It takes two to make war and each side always has the option of capitulating. Hence, at the simplest level, when wars occur, it must have been the case that both sides prefer to fight rather than capitulate. It may not seem like it right now, but this simple statement frames the puzzle entirely. Here's why.

Suppose actor *A* and actor *B* are disputing a valuable piece of territory worth \$10 billion. Each would prefer to get as much as possible, and it is best if the other agrees to a division without fighting because war is very costly for both of them. This is a classic mixed-motive situation: actors want to coordinate on a peaceful division of the territory but conflict over the terms of this division. As we said, this is an occasion for bargaining, so imagine that the actors are involved in negotiations where they are trying to revise the territorial distribution.

No rational actor would agree to give up what he can keep by force. Suppose that if they fight, *A* has 65% chance of defeating *B* and absorbing the entire territory in dispute. This means that *B* has a 35% chance of defeating *A*.¹ However, fighting would cost *A* \$2 billion and would cost *B* \$1 billion.

What can each actor expect to get by fighting? Actor *A* can expect to win the \$10 billion with probability 65%, and lose (that is, gain nothing) with probability 35%, all at a cost of \$2 billion. The expected payoff from war for actor *A* would therefore be:

$$\$10 \times 0.65 + 0 \times 0.35 - 2 = \$4.5,$$

that is *A* can expect to profit \$4.5 billion from a war. Similarly, actor *B* could expect to get

$$\$0 \times 0.65 + \$10 \times 0.35 - 1 = \$2.5,$$

¹We ignore the possibility of a stalemate. The argument goes through if we allow for it.

that is B can expect to profit \$2.5 billion from a war. Note here that war is expected to be profitable for both actors. Further, neither would acquiesce to bargain away more than he can get by fighting. Hence, actor A would not agree to any division of the territory that would give him less than \$4.5 billion, and B would not agree to any division that would give him less than \$2.5 billion.

Now, since the territory itself is worth \$10 billion, it is clearly possible to strike a bargain that would give *both* actors more than each could obtain by fighting. To see this, note that A will demand a chunk worth at least \$4.5, so anything more than that is better than fighting. Analogously, B would demand at least \$2.5, and anything more than that is better than fighting. The sum of the minimal demands is $\$4.5 + \$2.5 = \$7$. This means, that it would take \$7 billion worth of territory to satisfy the minimal demands for both actors which would induce them to accept a peaceful resolution. Because the territory is worth \$10 billion, both demands can be satisfied, and there is an additional surplus of \$3 billion left over that actors can distribute between themselves. Regardless of how they split this surplus, both actors are better off with a negotiated settlement than fighting.

The minimal demands determine the **bargaining range** of possible agreements that are acceptable to both sides. In our example, the bargaining range consists of all agreements that give A at least \$4.5, and B at least \$2.5. Clearly, there are many, many such agreements because one can imagine an infinite number of ways to split the remaining \$3 billion between the actors and keep them both satisfied. So, the bargaining range is quite large here.

As the reading by Fearon shows, this result is generic and does not depend on the numbers that we choose. Since war is costly, there always exist negotiated settlements that are better for both actors than going to war. Intuitively, since the war will eventually end somehow, a new redistribution of the territory will follow (in our example, the winner simply grabbed everything, but this does not have to be the case). However, *if the actors had agreed on the expected distribution prior to fighting, they would have achieved the same outcome and would have saved the costs of war.*² There are many other possible redistributions that are better than fighting for both actors. Generally speaking, the bargaining range always exists under these conditions.

This, then, is the fundamental puzzle of war: Given that war is costly and inefficient, and that superior bargains always exist, why do rational actors sometimes choose to fight? As you should recall, for a war to occur, both sides have to choose to fight it, so the question is quite strong for it asks not why one actor

²This is what people refer to when they say that war is inefficient after it has been fought, or *war is inefficient ex post*. The *ex post* phrase refers to the point in time when we make the evaluation, in this case, *after* the war is fought and over. If we want to talk about how we evaluated war before it is fought, as we did in the text when we calculated the expected payoff, then we would talk about war *ex ante*, that is, prior to it being fought. Since war is uncertain, actors are unsure about its outcomes, which is why the *ex ante* and *ex post* values differ. The first refers to what one expects to get from war, the second refers to what one actually got when the war was over.

would prefer to go to war but why both would do that. Again, the goal is to explain why actors are unable to pick one of the available settlements instead of fighting. In our example, one such settlement would be to split the surplus evenly, which would result in a division that gives A a total of $\$4.5 + \$1.5 = \$6$, and would give B a total of $\$2.5 + \$1.5 = \$4$. That is, A would end with $\$6$ billion worth of territory, and B would end up with $\$4$ billion worth. The facts that A was stronger (it had a higher probability of winning) and that its costs, although higher, did not exceed B 's costs by much, accounts for the somewhat uneven distribution.

Incidentally, this is perfect example of how force comes into play even when it is not used. In this case, each actor was able to “argue” that it would reject any division that would not give him what he would be able to take by force. I put “argue” in quotes because in our stylized setup, actors did not really communicate: since the situation is common knowledge, each could infer what the other's minimal demand is going to be. The size of that demand depends on the value of victory, the value (or lack of it) of defeat, the costs of fighting, and the probability of winning. These components make up the calculation that determines how useful force is.

Once an agreement is struck, it must reflect the relative capabilities of both sides; that is, it would reflect the potential use of force, or the *threat to use force* implicitly delivered by each actor. In our language, *each actor can credibly commit not to accept any deals that would leave him worse off than what he expects to get by fighting*. Force, even when not directly applied, was fundamental in arriving at the bounds of acceptable bargains.

Going back to our puzzle, we want to know how come rational actors do not strike one of the available bargains. Intuitively, it would appear that somehow the bargaining range that we identified must fail to exist. In fact, all three explanations that we are going to give hinge on reasons why this can happen. We shall use our example to investigate the possibilities.

2 Indivisibility of Stakes

The most obvious problem is that the territory may not be neatly and continuously divisible in a way that would satisfy both actors simultaneously. Going back to our example, suppose that the territory was worth $\$10$ billion, and that the wealth comes from an oil well worth $\$8$ and a river with commercial value of $\$2$ billion. Everything else is a desert and is worth noting. If the territory is divided, the actor who gets the portion where the oil is located can exploit it, and the actor who gets the portion with the river can extract the benefit from it.

The expected value of war for both actors is the same as in our original case because victory results in a complete conquest, so the winner gets to enjoy both the oil and the river. So, as before, A would not accept anything less than $\$4.5$ billion, and B would not accept anything less than $\$2.5$ billion. Given the distribution of value in the disputed territory, there is no agreement now that

would satisfy both actors simultaneously. If *A* gets the chunk with the oil, he would be better off than fighting. However, *B* gets only \$2 billion worth of territory, so he is better off going to war. Conversely, if *B* gets the oil, then *A* would be better off going to war.

If the issue at stake is indivisible as in this example, then rational actors could end up fighting even if they are fully cognizant about their respective chances of success and war is costly. For example, if both Israelis and Palestinians agree that Jerusalem is indivisible, then there may not exist a deal that would satisfy both of them, which in turn can render fighting a rational choice for both.

There are many reasons why territory may not be perfectly divisible. For example, as in the case above, it may contain resources that are distributed in a way that makes sharing impossible. Some territories are more valuable than others and the value cannot be split, and so commercial access to rivers, or ports would make territories valuable in an indivisible way. Strategic capabilities will do that too: when Hitler demanded that Czechoslovakia cede the Sudetenland in 1938 under the pretext that the German minority there was mistreated and needed to be brought into the Reich, the real problem was that the mountainous area contained the entire defense system of the Czechs. When it was gone after Munich, the rump country was left with no way of protecting itself against Germany, which is why Hitler found it exceedingly easy to annex it a year later without a fight. The Sudetenland was indivisible but the Czechs were not strong enough to oppose the Germans alone.

Other indivisible issues may be all-or-nothing type of demands. For example, during the negotiations to end the Korean War, the Americans and the North Koreans could not agree what to do with the prisoners of war. The Americans demanded that all North Korean POWs were given a choice to remain in the South or return to the North. The North Koreans disagreed and insisted that all POWs should be forced to return to the North upon release. In a situation like this, there is no way to divide the issue.³

Issue indivisibility, however, may be a less compelling explanation than you might think. That's because in principle it is not clear at all why an issue cannot be divided or, if it truly cannot, then that the parties cannot make side payments for compensation. Going back to our territorial example, one possible solution to the indivisibility problem would be to arrange for a joint exploitation of the oil: e.g. by setting a joint-stock oil company, *Shared Oil*, in which *B* has a share of, say, 10%. This means that when *A* exploits the oil, it would get 90% of the profits, or \$7.2 billion, and *B* would get 10%, or \$0.8 billion. They would then divide the territory with *A* getting the oil and *B* getting the river. With *Shared*

³On ethical and moral grounds "splitting the difference" simply does not work. One cannot agree to let 50% of the POWs choose and then extradite the other half. As it happens, the final agreement was essentially on the American side: all POWs that desired repatriation were to be returned within 60 days, the rest would become civilians. Also, civilians could visit the POWs during that time, and civilians on both sides of the armistice line could move across to settle in the part they liked better. A lot of Chinese and Koreans went South (or to Taiwan), along with significant numbers of civilians.

Oil, *B*'s total benefits are now \$2.8 billion, which is clearly better than fighting, and *A*'s total benefits are \$7.2 billion, which is still better than fighting. Hence, in principle, the actors could work out a deal to turn an indivisible issue into a divisible one.

This scheme may work for some things but won't work for the POWs case, which appears to be truly an intractable all-or-nothing issue. However, most negotiations are not made over a single issue. In principle, one could envision compensation on other issues that would make giving in on the all-or-nothing one more attractive to the other side. Even without other issues under consideration, a player could offer some compensation that would be of value to its opponent. For example, suppose that there was no way to form the *Shared Oil* company. Still, *A* knows that it would get \$8 billion once it gets access to the oil, so it can offer to pay *B* some amount, say \$1 billion in addition to giving up access to the river. This would bring *B*'s total benefit to \$3 billion, which would be better than fighting. And, of course, *A* gets to enjoy the net profit of \$7 billion, still far above its payoff from fighting. An indivisible issue could be made "divisible" by similar side-payments.

There is a further problem with indivisibility: it can be strategically induced. That is, actors may *claim* that something cannot be divided in order to commit themselves to a bargaining position that would ensure a better outcome. This sort of commitment is similar to restricting your freedom of action: you somehow find a way to convince your opponent that it is impossible for you to accept a division of the issue in question. Then, if your opponent wants to avoid fighting, he will have to give in on the issue altogether (perhaps induced by suitable side-payments).

Leaders, therefore, may have extremely strong incentives to pre-commit to particular bargaining positions by claiming that some part of the issue is not subject to negotiation. For example, Israel can claim that Jerusalem is an "eternal and indivisible" part of the country. In effect, such a stance makes Jerusalem an all-or-nothing issue, like the POWs case. It is not that one cannot imagine a way to "divide" the city (e.g. shared jurisdictions or perhaps an international mandate), it is that at least one side claims that any division is impossible. Usually, the reasons invoked could be patriotic (historical), or religious, or matters of national honor, or take your pick. The point is that to be credible, such a commitment has to be irreversible and one way of making it so would be by *tying one's hands*: that is, making it impossible to agree to any division without a special national mandate, like a referendum, which may very well fail.

The obvious problem with such a pre-commitment is that the other side can do the same thing. If both claim that Jerusalem is indivisible *and* simultaneously a part of each respective country, then there will be no way of resolving the dispute short of war. Perhaps side-payments could help in some instances, but perhaps not in such intractable ones, especially if pre-commitment on both sides takes the form of public tying of hands.

The point, however, should be clear: anything is divisible in principle, either

through some arrangements of shared control, or through side-payments. However, making something indivisible is a bargaining tactic, that is, it is a choice that could serve as a useful commitment device if it works. If it does not, it can virtually guarantee bargaining breakdown. The conclusion is that while indivisibility is a potential explanation of war among rational actors, it is not a very persuasive one because one must be able to explain why actors are unable to agree to accept side-payments or make it divisible in some other way.

3 Asymmetric Information

Asymmetric information refers to situations in which one actor has information that is different from what the other actor knows. In these cases, we refer to the **private information** that each actor possesses. Of course, in the real world, each state has secrets, especially of military nature, that it guards jealously. Asymmetric information is ubiquitous. However, imperfect knowledge of the other side is not, by itself, sufficient to cause war because in theory the actors should be able to communicate this information to each other. Why would they want to do that? Because they know that with complete and symmetric information, the bargaining range will always exist, and therefore if they reveal everything, they would be able to agree on some settlement within that range and avoid going to war.

Hence, the real problem here is not that actors have private information, but that they may have incentives not to reveal it. So, let's examine why a rational actor would want to conceal information even though he knows that if it is revealed, war will not occur.

We modify our territorial example as follows. Suppose that everything is common knowledge, as before, except actor *B*'s cost of fighting. Its true value is known only to *B*, while *A* only has a rough idea what it might be. That is, *B*'s cost of fighting is private information for *B*, and *A* has some belief about it.

For simplicity, suppose that *A* believes that the cost is either \$4 or \$1. That is, player *B* is either weak because its cost of fighting is high (\$4 billion) or he is strong because its cost of fighting is low (\$1 billion). From intelligence reports, player *A* estimates that the chances of *B* being strong are about 30%, and the chance that he is weak is about 70%.

This now is a situation of asymmetric information. Suppose first that *B* is strong, in which case his expected payoff from going to war is \$2.5, as before. Clearly, he has a strong incentive to communicate this private information to *A* because doing so would enable them to strike some bargain and avoid fighting. However, *B* cannot credibly reveal this information; that is, there is no way that it can convince *A* of its strength by just talking. Here's why.

From *A*'s perspective, its opponent's reservation value (the minimal demand that it would accept) is \$2.5 if he is strong. This means that the largest demand that *A* can make to a strong opponent is to keep \$7.5 for himself because anything more than that would be rejected for sure. On the other hand, *B*'s minimal

demand is

$$\$0 \times 0.65 + \$10 \times 0.35 - 4 = -\$0.5$$

if he is weak.⁴ What does this negative number mean? It means that for the weak B , war is too costly and, given the low odds of winning, not worth the effort. The weak B can never credibly threaten to go to war against A even if A grabs the entire territory. Hence, if A believes that his opponent is weak, he will demand (and receive) everything. Incidentally, the way we just defined if an actor is strong or weak is what we will usually mean in this class: *an actor is strong if his expected payoff from failing to reach an agreement is high, and he is weak if this payoff is low*. In crisis bargaining, disagreement means war, so a strong actor is one who has a high expected payoff from fighting.

To summarize, A can demand at most \$7.5 if his opponent is strong, and can demand the entire territory if he is weak. If A demands only \$7.5, then peace will occur with certainty because both the strong and the weak types would accept that. Although the strong type is getting the absolute minimum, he can still do no better by going to war, and hence we would not go to war. On the other hand, the weak type is positively getting a great deal for if A knew for sure that he was weak, he'd get nothing instead of \$2.5 billion. Therefore, war can be avoided for sure if A 's demand does not exceed \$7.5 billion.

Unfortunately, A 's demand will exceed that for sure. That's because it is not optimal for A to offer \$2.5 to B , and the reason for that is that A would be overpaying with certainty. The amount of \$2.5 is too high because if B happens to be weak, then A is agreeing to give away something he does not have to. Hence, A will *risk* going to war in order to extract a little higher *return* from a negotiated settlement. That is, it will balance the slightly increased risk of getting an offer rejected and going to war with the slightly improved terms of the negotiated settlement.

A knows that the weak B would accept anything, including no territory for himself. At the same time he knows that the strong B would reject anything smaller than \$2.5 billion. Hence, from A 's perspective any offer lower than \$2.5 billion would have a 30% chance of being rejected and 70% chance of being accepted. These, of course, are just the probabilities that B is strong and weak, respectively. Since it does not pay to offer anything to the weak B , the optimal offer for A would be *to demand the entire territory, even at the risk of war*. Given that this offer will be rejected and war would occur with a 30% chance (in which A expects to get \$4.5 billion) or accepted with 70%, the expected payoff from making this demand is

$$\$4.5 \times 0.3 + \$10 \times 0.7 = \$8.35,$$

which is strictly greater than getting only \$7.5, which is what offering the \$2.5 billion necessary to ensure peace would leave A with. In other words, A has a strong incentive to deliberately risk war in order to obtain a better deal at the

⁴The difference between the two is, of course, the difference between the two costs associated with player B .

bargaining table. In this case, it is willing to run a risk of 30%, which is not a trivial amount.

To see that this risk cannot be eliminated by communication, note that *B* has no way to convince *A* that it is strong. Suppose that *A* believed *B* and were persuaded that he is strong. Then, *A* would clearly offer *B* at least \$2.5 to avoid war and peace would be the outcome. The problem is that if *A* believed *B*'s claim about strength, then the weak *B* would have an incentive to claim that he is strong as well: after all, this would net him a nice profit while admitting to weakness would leave him with nothing.

Thus, *actors with private information have incentives to misrepresent* what they know. Simply put, they have incentives to lie and bluff in order to improve the terms of the settlement they get. No amount of communication is going to help here, which is why we often say that talk is cheap. This conclusion is quite general—in our example we could have assumed that both sides had private information about their costs of fighting, or that they disagreed about who would win the war, or any other number of possible sources of disagreement—the result would hold still (see Fearon's article).

This, then, constitutes a real reason rational actors can end up in war and it is a result of deliberate decisions to run risks in order to obtain better deals. This is called the **risk-return trade-off**: an actor accepts a slightly higher risk of war in order to obtain a slightly higher return from the bargain.

We should carefully distinguish between *mistakes* that any actor can make, and *deliberately risky behavior* that could potentially cause disaster. In our example, if the actors end up fighting, it is not because *A* made a mistake and underestimated the strength of his opponent, but rather because he chose to make a larger demand knowing that there is some chance that things would go to hell. Even though misperception, bias, and miscalculations occur, the mechanism here is quite different.

For this explanation to be complete, we have to explain how fighting can resolve the informational problem. Why must we demand this? Because without it, our explanation will not be satisfying: if we say that private information with incentives to misrepresent can cause war but then fail to explain how war solves the problem that supposedly caused it, we can't really explain how wars end.

Wars, of course, can end in basically two ways: either one side completely eliminates the other, or the two sides reach a settlement. If one side is gone completely, then war does resolve the informational problem quite trivially: there's no opponent left. It turns out, however, that most wars do not end in this way: only about 36% of wars since 1815 have ended in a clear-cut military victory; the remaining 64% have ended in compromise settlements while both sides could continue the fight.⁵ This now means that opponents *decided* to end the fighting, and because they were able to do so, they must have resolved the problem that

⁵The data set for this and subsequent claims is available from the replication package for my paper "How Initiators End Their Wars," published in October 2004 in the *American Journal of Political Science*. The entire data set is on my website.

caused them to go to war in the first place. If they had not resolved it, they would have achieved peace while the problem existed, which is another way of saying that the supposed problem could not have caused the war... after all, they were able to find a peaceful settlement without solving it.

All this now means that a complete explanation for war should not only account for its outbreak but also for how fighting solves the cause. It turns out that the risk-return trade-off can be easily extended to accommodate that. It's a bit technical, so we won't go into it here, but I will give you the logic.⁶ In the model we just used, the start of the war reveals all the private information: after all, at that point *B* has chosen to fight rather than accept the offer *A* made, and only the strong *B* does that. This means that after war breaks out, *A* knows for sure that *B* is strong, and we're back in our original example where they must be able to settle immediately.

War, however, is not an all-or-nothing event, but really a sequence of battles spread through time. One can potentially choose to negotiate after any battle. In other words, diplomacy is (at least in principle) always available during the war itself. Let's assume that a battle costs \$500 million to a weak *B* and only \$100 million to a strong one (that is, \$0.5 and \$0.1 billion respectively.) Going back to our original example, suppose that war breaks out and after a single battle *A* gets convinced that its opponent is strong, and so it offers the minimum such *B* would accept, or \$2.5 billion. Of course, *B* would accept such an offer no matter if it is strong or weak. (I am drastically simplifying here in assuming that after one battle the rest of the war essentially looks the way it did at the beginning. See the paper cited in footnote 6 for a better treatment that requires a lot more math.) Consider now the per-war calculations from the perspective of a weak *B*. Rather than having to pay \$4 billion to wage war to the end, he can pay \$0.5 billion for one battle and then obtain \$2.5 billion from settling on *A*'s terms. Would he do that? Sure: the expected payoff is now $\$2.5 - \$0.5 = \$2$ billion, which is strictly better than getting nothing by accepting *A*'s initial offer! In other words, even the weak *B* would go to war if he expects that doing so would persuade *A* to yield.

We conclude that it cannot be the case that *A* would be convinced of *B*'s strength after only one battle. They have to fight longer. How long? Enough to make war sufficiently unpleasant to a weak *B* and deter him from entering it in the first place. Since a battle costs \$0.5 billion and a settlement will be worth \$2.5 billion when *A* gets convinced that *B* is strong, five battles would obliterate any positive advantage from fighting them: a weak *B*'s payoff will be zero if he expects to fight five battles (total cost of \$2.5) and then settle for \$2.5 from *A*. Would the strong *B* be willing to fight these battles? His expected payoff of doing so will be $\$2.5 - \$0.1 \times 5 = \$2$, which is strictly better than settling for *A*'s initial demand of the entire territory. In other words, it pays the

⁶If you are interested, the formal development is in my 2003 paper "The Principle of Convergence in Wartime Negotiations" published in the December issue of the *American Political Science Review*. A copy is available on my website.

strong *B* to fight five battles in order to convince *A* that he is strong and obtain his minimum demand of \$2.5. Unfortunately, this now means that war will not only occur (if *B* happens to be strong) but that it will last for a while before *A* is persuaded to revise his beliefs. Eventually, however, *A* will acquire enough evidence that *B* really is strong, and at this point he will be willing to settle. For that to happen, fighting must have been unpleasant enough for the weak *B* so that he would have quit earlier.

It is (roughly) in this way that war can reveal information: fighting serves to persuade the opponent that one is stronger than previously thought. The idea is to induce the opponent to revise his expectations and make a better offer than he otherwise would. Unfortunately, this may require a commitment to fight for a while, with all the attendant risks, costs, and pain. Again, there is no way a strong *B* can persuade *A* to yield short of paying these costs: words are cheap, and only the willingness to bear the pain can credibly reveal *B*'s strength. (This is an instance of *costly signaling*.)

This now gives you a pretty good idea about *war as a bargaining process*. We've extended the Clausewitzian logic (war as instrument of policy) and have defined it much more precisely. Our stylized explanation actually treats war as a costly way to influence the expectations of the opponent. The process of fighting (and possibly negotiating while fighting) reveals information and the expectations of the two opponents can begin to converge. Once sufficient information is revealed, they expectations come close enough to open up the bargaining range, and they can settle. In this way, the asymmetric information explanation can account for why war breaks out, why fighting may persist, and how fighting can make the problem go away and allow the two actors to settle.

How good is the informational explanation? Pretty good, actually. Most wars are pretty short: the median duration is under 6 months.⁷ That's another way of saying that information seems to accumulate very quickly, as it should: after all, wars are terribly expensive, and there is no need to fight longer than absolutely necessary. In addition, unexpected events tend to carry a lot of weight in war. For example, if a side that everyone thought was very weak suddenly wins a major battle, it will cause a drastic re-evaluation of its opponent's expectations, perhaps causing a swift end to fighting. I will give you two examples, both from the 19th century (and both from my 2004 paper).

Prior to the Seven Weeks War (1866), Prussia and Austria appear to be quite evenly matched in terms of military capabilities. However, it was generally believed that Prussia was the weaker side and that it would be unable to defeat the Austrian Empire. This belief was shared by the French and the Russians, with the former actually helping the Austrians promote the war. The Austrians were also thought to be in a good defensive position to repel both the Prussians and the Italians.

⁷The median duration of wars that ended with a clear military outcome is 4 months, and for those that ended in some sort of compromise, it is 6.4 months. So it's not the case that all short wars simply involved military victories.

Austria did defeat Italy. However, the Battle of Königgrätz on July 3, 1866 was a stunning surprise. The Prussians routed the Austrians at one-fifth of their losses, a “breath-taking accomplishment that utterly destroyed Austrian morale.” The unexpected victory, which the French called the *surprise de Sadova*, compelled Austria to seek peace, and the war ended after less than a month and a half, with Prussia, the initiator, enjoying significant gains and Austria embarking on its quick decline as a major power.

Another war in which military parity was coupled with optimistic estimates is the Serbo-Bulgarian War of 1885. King Milan of Serbia was concerned with the alteration of the balance on the Balkans that would result from Bulgaria’s unification with Eastern Rumelia in violation of the Berlin Decree. Serbia, goaded by Austria, rejected half-hearted offers of territorial compensation and adopted a hardline course, further emboldened by the Russian recall of all its officers from the Bulgarian army in protest of the independent policies of the young prince Alexander.

When Bulgaria proceeded with the unification, it was a state less than seven years old, with an army denuded of officers, and with no experience. Almost all of its 30,000 troops were concentrated in Eastern Rumelia expecting an Ottoman attack. The border with Serbia was undefended. Under these propitious circumstances, King Milan invaded the country with all currently mobilized forces, or about 35,000 troops. No one thought Bulgaria would last long and the government was expected to sue for peace.

However, in a stunning maneuver utilizing the limited resources of the single railway, and making heavy use of horses, the bulk of the Bulgarian army marched across the country, returning to face the invaders at Slivnitsa, 30 kilometers west of the capital. In the ensuing battle on November 17-19, the Bulgarian army, with contingents going into combat immediately upon arrival, routed the Serbs and pursued them back across the border. This battle, a “remarkable achievement for an untested army shorn of its senior officers,” came as a “surprise to most observers.”

The unexpected defeat of the Serbian army stirred the Austrians into defense of their protege, and the Bulgarians were forced to settle under the threat of invasion from the Austro-Hungarian Empire, ending the war after less than a month of fighting. Serbia, the initiator, was compelled to make concessions and accept the enlarged Bulgarian state.

So, instead of thinking about war as just random (wanton) violence, or (vaguely) as an instrument of policy, we can now think of war as a bargaining process. As Schelling put it, “War [is] a bargaining process—dirty, extortionate, and often quite reluctant as a bargaining process on one side or both—nevertheless a bargaining process.” Now you know at least one way in which this definition can be made precise. But... that’s not the only way we can explain war!

4 Dynamic Commitment Problem

Although the asymmetric information argument provides an internally consistent explanation for war, it does not appear to do a very good job of accounting for certain conflicts. For example, whereas interstate wars are generally short (as we noted before), which is in keeping with the information revelation notion, civil wars are much longer. There were 123 civil wars from 1945 to 1999, of which 70 lasted over five years, and 39 lasted over ten years.⁸ It is hard to believe that combatants can take over a decade to learn enough about each other. There must be something else that is preventing the settlement of these wars.

Another problem is that the informational argument gives a strained reading of certain historical cases. For example, how would we account for the outbreak of the Second World War? Why did Britain choose to resist Germany over Poland when it had acquiesced over Czechoslovakia just a year before? The informational argument holds that had Britain possessed complete information about Hitler's minimum demand, it would have satisfied him and avoided fighting. But this is clearly non-sense. In fact, the exact opposite is true: by 1939, Britain had concluded that Hitler's goals are too expansive and that it could not accommodate him. Appeasement over the remilitarization of the Rhineland, the *Anschluss* with Austria, and the partitioning (and gobbling up) of Czechoslovakia boiled down to attempts to satisfy what Britain perceived as somewhat legitimate, but limited, German goals. The demands Hitler made of Poland clearly dashed British hopes that the Germans would stay put. In other words, Hitler's actions revealed to Britain that Germany's goals placed it squarely against British interests, and Britain resolved to resist by force. Contrary to the informational argument, the previous deals were mostly made possible by British uncertainty about Hitler's goals, and war came when Britain acquired enough information about them.⁹

We conclude that we need an explanation that does not assume that peace automatically follows if actors have complete information about each other. But why actors fight if they know everything that can happen and agree in their expectations about the consequences? The mechanism that can cause war between rational actor is due to time inconsistency: Actions that one cannot commit to performing at a future date even if doing so would help achieve a better outcome today. This is often call the **dynamic commitment problem** or, even more simply, the credible commitment problem.

There are several ways in which time inconsistency can manifest itself. One is when each actor enjoys an *offensive advantage*; that is, if the probability of winning the war is higher for the actor who strikes first. This is explored in Fearon's article and we have already discussed some of the logic behind first-strike ad-

⁸James Fearon, "Why Do Some Civil Wars Last So Much Longer Than Others?" *Journal of Peace Research*, May 2004.

⁹For a detailed development of this argument and the limitations of the informational explanation, see Robert Powell's "War as a Commitment Problem," *International Organization*, 2005.

vantages when we studied the question of crisis stability and proliferation of nuclear weapons. This mechanism is closely related to the idea of **preemptive war**, where attacking first improves one's chances of success immensely.

Intuitively, since each actor could choose to go first, each must be compensated enough with the bargain to offset the temptation to do so. As the first-strike advantage increases, the bargaining range shrinks, and may disappear entirely. At this point, at least one of the actors will be tempted to exercise the first-strike option because both know that there is no deal that could satisfy them simultaneously. The time inconsistency here comes from the fact that neither actor can credibly promise not to use its first-strike advantage. As we have seen, in addition to the commitment problem, the incentive to preempt quickly creates a spiral of escalating expectations of preemption that eventually would cause one of the actors to jump the gun.

We shall discuss another, perhaps a bit more interesting, source of time inconsistency, which in our case also has a ready and intuitive application. Suppose that *B*'s strength is increasing over time, perhaps due to economic growth that is much faster than *A*'s. Initially *A*'s chances of victory in war are 75%, but in 10 years they would drop to 25%. That is, we have an occasion of power transition between the two actors. Suppose further that each actor's costs of fighting amount to \$1 billion.

Consider what would happen today, before the transition has occurred. Both actors can analyze the economic projections and know that in a few years *B* would grow quite strong. Whatever the distribution of territory that they agree on today, *B* can always demand a revision later. So, let's see what would happen if *B* demands a revision of the territorial distribution in ten years. At this point, his chance of victory will be 75%, so going to war then would yield him an expected payoff of:

$$\$0 \times 0.25 + \$10 \times 0.75 - 1 = \$6.5.$$

That is, its minimal demand would be for \$6.5 billion. What would *A* do? His payoff from going to war would be:

$$\$10 \times 0.25 + \$0 \times 0.75 - 1 = \$1.5.$$

Hence, his minimal demand would be for \$1.5 billion. The total minimal demands add up to \$8, and hence there will always be some bargain available for any distribution of the surplus \$2 billion. In ten years, *B* will demand a revision of the territorial status quo, and in the resulting settlement will get at least \$6.5 billion. There will be no war because both actors will prefer to strike such a bargain.

We now return to the present. If war occurs and *A* wins, he would eliminate *B* and enjoy the benefit of the entire territory now and in ten years. Of course, if he loses, he will get nothing now or in ten years. Hence, *A*'s expected payoff from fighting now is:

$$(\$10 + \$10) \times 0.75 + (\$0 + \$0) \times 0.25 - 1 = \$14.$$

That is, if A fights right now, his expected payoff from war is \$14 billion. This takes into account both today's benefits and the ones he will reap in 10 years (if it wins). Similarly, B 's present payoff from war is:

$$(\$0 + \$0) \times 0.75 + (\$10 + \$10) \times 0.25 - 1 = \$4.$$

Notice that the total payoff from the territory over time is \$20, and the minimal demands sum up to $\$14 + \$4 = \$18$. That is, there seems to be enough to meet both players' demands and perhaps resolve the conflict without fighting. Clearly, A 's current payoff of fighting is quite high, and A needs to be compensated properly if he is to remain at peace. The most that B can agree to right now is to transfer the entire territory to A . However, A knows that in ten years, B will not acquiesce to anything less than \$6.5, and so the most A can expect then will be \$3.5 billion. This means that if B agrees to *transfer the entire territory* to A today, A 's best payoff over time will be: $\$10 + \$3.5 = \$13.5$ billion. This includes the \$10 billion now plus the maximum amount of \$3.5 that B will agree to in the future.

And we now have a problem for even if B agrees to give everything to A today, A still prefers to attack immediately because doing so yields a payoff of \$14 billion as opposed to \$13.5. That is, the maximum transfer that B can agree to today is not sufficient to compensate A . The reason, of course, has to do with B 's anticipated behavior in ten years. If only B could commit to offer slightly more to A then, say just \$1 billion more, then A 's expected payoff from peace would be $\$10 + \$4.5 = \$14.5$, and so he would remain at peace.

However, any such promise by B is not credible because both actors know that when the ten years pass, B will have grown so strong that he will not want to agree to pay this additional \$1 billion. B 's promise is incredible and so peace is time inconsistent.

This is truly a problem for B would love to be able to make such a promise! To see this, suppose that he did transfer everything to A now and then gave A the required \$4.5 in ten years. The payoff from remaining at peace would then be $\$0 + \$5.5 = \$5.5$ billion, which includes nothing for the decade, and then the remainder after compensating A after that. This \$5.5 billion is strictly better than fighting today which yields only \$4 billion. Hence, B is strictly better off agreeing to no territory today, making the promise to pay an additional \$1 in ten years, and fulfilling this promise.

But this does not make it possible to do so: B 's inability to credibly commit to an action in the future makes things extraordinarily bad for him today. He ends up fighting in a war today without real chances of success because he cannot promise to keep his commitment tomorrow. Note further that A would also be better off if B could make such a credible commitment: he would get \$14.5, which is better than fighting.

We conclude that two rational actors can end up in a war without misperception, mistakes, or emotions due to time-inconsistency (credible commitment) problems. Such problems can make ending conflicts quite difficult as well. To

end fighting, both sides must agree to a cease-fire. Suppose now that one side uses the time to reorganize instead of negotiate for peace in good faith. This enables it to create a power transition, turn around, and renew the fighting with much better chances of success tomorrow. That is why in general, cease-fires are expected to be used by all actors in such a way (and they are), which is also one of the reasons why they often do not work.

We shall see credible commitment problems again when we discuss ethnic conflicts. The situation there follows a similar logic. To stop the fighting, the rebels have to negotiate some sort of power-sharing agreement with the government. Suppose they reach such an agreement. The rebels are then supposed to disarm, but doing so exposes them to an attack by the government forces, a temptation which the government cannot credibly promise to resist. Hence, rebels may not want to disarm, making a deal very difficult. In such situations, intervention of a third-party that would separate the rebels from the government and guarantee the safety of the disarmament may be quite helpful. But then again, once the rebels have disarmed, it is not clear why the government would want to uphold its end of the power-sharing bargain. After all, now there is no organized rebel force to threaten it!

You may probably see that our logic leads to serious modifications of power transition theory as well. To see this, note that there is no special relationship between the likelihood of war and how close the actors are in capabilities (probability of winning war). The transition causes war in our theory because the declining actor knows that he would have to settle for much less in the future, so it prefers to go to war today while he is still strong. The rising actor's problem is that there isn't enough territory to compensate the declining one and induce him to remain at peace.

Further note that the general description of **preventive war** goes something like this: A declining state fights today while still strong to avoid having to fight tomorrow when it will be weak. This makes sense but it is different from our theory. The declining state never fights at worse terms in the future. The reason, of course, is that there always exist bargains that both can accept instead of going to war. The problem is that even without fighting, it will be forced to accept a worse settlement once it declines. So it fights today to prevent having to settle for less tomorrow, not to avoid a war in the future.

The most famous preventive war caused by this commitment problem is the Peloponnesian War between the democratic commercial maritime Athenian empire and the oligarchic agricultural land-based Spartan alliance.¹⁰ Traditionally, Sparta had been the premier land power in Greece, her professional army was considered unbeatable. The Spartans had developed an extensive network of alliances but pursued largely conservative inward-looking policies designed to maintain their preeminence in the Peloponnesus. Athens, on the other hand,

¹⁰The great Athenian general Thucydides wrote an account that covers the first part of the war. For a modern, eminently readable and engrossing, history, see Donald Kagan's *The Peloponnesian War*, Penguin 2003.

was a dynamic society that had recently gone democratic. The Athenians were expanding their sea-based commercial empire, and were getting very rich and powerful in the process. The Spartans looked on uneasily, but did nothing about it for a long time. Indeed, the two city-states cooperated against the Persian invasion of Xerxes. Sparta had borne the brunt of the first assault in the legendary battle at Thermopylae (and later the decisive battle at Plataea), while Athens had led the Greeks to victory in the crucial sea-battle at Salamis. The shaky alliance, however, did not outlast the war by much.

The break eventually came when the faraway colony of Corcyra got itself into yet another conflict with its mother-city of Corinth, and the two appealed for help from Athens and Sparta, respectively. The Athenians tried to warn the Spartans not to declare war, and argued that had done nothing wrong in their pursuit of wealth and power. They had acquired their empire peacefully and had acted with great moderation and justice than common at the time. In short, Sparta had nothing to fear except a war with Athens because it would find the democracy quite capable already. In other words, the argument was a mixture of reassurance (no need to fight Athens, she will act justly) and warning (if you do fight, you will find a formidable opponent).

It did not work. The Spartans deliberated for a short time and then decided to fight. Thucydides summarizes their reasoning aptly: “The Spartans voted... that war must be declared, not so much because they were persuaded by the arguments of the allies, as because they feared the growth of the power of the Athenians, seeing most of Hellas already subject to them” (1.88). He argued that the immediate cause (the broken treaty) was not as important as the “real case, ... [which he considered] to be the one which was formally most kept out of sight. The growth of the power of Athens, and the alarm which this inspired in Sparta, made war inevitable” (1.23). Whereas Athens did not present an immediate danger, inaction would spell disaster: not just because it will embolden the Athenians but because it will frighten Sparta’s allies who may well defect, leaving Sparta to fend for herself in the future. Thus, Sparta resolved to fight a preventive war rather than face a future from a position of weakness.

We now have to consider the question we did at the end of the previous section: if the commitment problem causes war, then how does fighting resolve it? After all, if war does not solve the problem, then it could now have been caused by it.¹¹ As it turns out, this is an area where research is ongoing. Right now, all I can suggest is a tentative answer from a work I have done with my colleague Bahar Leventoglu. She and I have shown that when resources for war are limited, the destruction of the benefits one can obtain from a bargain (e.g. territory devastated, infrastructure ruined, etc.) reduces the incentives to hold out for more, and at some point it can actually open up the bargaining range that did not exist from the outset.

¹¹Recall the logic: if actors agree to stop fighting and achieve piece while the problem still exists, then why did they have to start the war in the first place? Why didn’t they achieve this peace then? A complete explanation has to account for both the outbreak of war and its termination.

Further, and somewhat ironically, because in principle actors can always sue for peace, the incentive to do it any particular moment is weakened. To see why, consider some point in the war. You can negotiate for peace right now on the basis of your current military and political situation. You also know that you can negotiate for peace in the future, after one or more battles. You may be tempted to fight these battles to improve your position (or, who knows, perhaps your opponent is facing an imminent collapse, so you can actually win outright). Your adversary, of course, faces analogous temptations. In the end, because you both know that you have incentives to seek peace, you undermine the prospects of peace today. If either one could credibly threaten to fight to the bitter end and never seek peace in the future, then the temptation to prolong the war just a little bit will be eliminated: peace acquires this “now or never” quality and both sides may be much more willing to forego the risks of continuing the war. As the Roman Varro put it, if you want peace, prepare for war. In our case, if you want to end the fighting, you have to threaten (credibly) not to offer peace terms in the future. But this, of course, is the problem: both sides know that each of them will have incentives to settle in the future, and so these threats may not be credible. When they are not credible, peace may not be possible. In the paper, we show how fighting can make these threats credible too. Unfortunately, the paper is quite technical and requires quite a bit of math to show that these claims are true.

Even with this explanation, we still do not have a good one that handles the case of differential growth. Research has shown that countries recover fairly quickly after a war, and so war does not seem to hurt their long-term potential very much.¹² In other words, war will not really solve the differential growth problem, or at least we have no idea (for now) how it can. So now you basically know about as much as I do on this issue!

5 The Cost of Peace

One legitimate question is to ask if actors would be willing to pay to avoid war that is caused by the commitment problem. The resounding answer is “it depends.” Let’s consider our credible commitment story again. As before, *A*’s military strength will decline over time, and so his probability of winning will drop from the initial 75% to 25% within a decade. We already know that this will cause *B* to have a commitment problem, and will trigger preventive war by *A* today. What if *A* could slow its decline by investing x dollars in defense? For example, suppose that there exists a technology that costs $\$x$ and that would cause *A* to lose only a fraction of his strength, and so in 10 years, his probability of winning will be 50%. (That is, it’s still worse than today’s but at least not as bad as without having the hardware.) What happens if *A* invests $\$x$ in his

¹²This is known as the Phoenix Factor: defeated powers tend to recuperate quickly and return to their international position within a single generation.

defense? In a decade, his expected payoff from war would be:

$$\$10 \times 0.50 + \$0 \times 0.50 - 1 = \$4,$$

and *B*'s expected payoff from war will be:

$$\$0 \times 0.50 + \$10 \times 0.50 - 1 = \$4.$$

Hence, *B*'s minimal demand would be \$4 billion (rather than \$6.5 if *A* does not build up his defenses). Hence, *A* can expect, at most, \$6 billion in the future from *B* because *B* cannot credibly offer anything more.

We now return to the present. *A*'s expected payoff from war now is, as before \$14, and *B*'s expected payoff from war now is, also as before, \$4 billion. However, the expected payoff from peace has changed. If *B* agrees to transfer the entire territory to *A* for a decade, *A*'s best payoff will be:

$$\$10 + \$6 - x = \$16 - x.$$

That is, \$10 billion from having the entire territory now plus the best deal he can expect in the future minus the costs of building the defense that would ensure that *B* actually makes that offer in a decade. This payoff is better than going to war as long as $x < \$2$, that is, as long as the defense costs less than \$2 billion. To see that, note that $16 - x > 14$ if, and only if, $x < 2$. Hence, if defense is not too expensive, then *A* would invest in it, eliminate the commitment problem, and peace will prevail.

Unfortunately, if the defense technology costs more than \$2 billion, then investing in peace is just not worth it. To see that, suppose that $x = \$3$, and so the best payoff from peace for *A* is now $\$10 + \$6 - \$3 = \$13 < \$14$. That is, the best peace can offer is worse than going to war. If that is the case, peace may be unaffordable: actors cannot avoid the commitment problem even if they really wanted to, and war is likely.

Although this explanation is very closely related to the previous one—after all, it is the commitment problem that causes war here as well—it serves to highlight the notion that actors may not be able to avoid the problem even if they had the means of doing so. The costs of peace are sometimes overlooked in analyses of crises, and this example shows that doing so can be quite misleading.¹³ For example, among the reasons Mao had for entering the Korean War on the side of the North Koreans was that if the U.S. was allowed to conquer the territory up to the Yalu, the new Chinese state would have to be on a permanent war footing in order to deter further encroachments. Maintaining such defenses would be ruinously expensive at a time when the country needed desperately to recover from the civil war. It would also probably jeopardize the new regime

¹³Hint: think about how much the U.S. expected to have to pay to continue the containment of Iraq under Saddam Hussein—in terms of maintaining the no-fly zones, the sanctions, the bases in Saudi Arabia, possible eventual development of WMD, etc.; and in terms of expected duration—Hussein was not that old.

by moving disaffected peasants away from the communist camp, and toward a possible resurgence of Chiang Kai-shek from Taiwan. In other words, Mao believed peace with America in China's backyard would be too expensive.

6 Conclusion

To summarize, we have uncovered four theoretical mechanisms that could lead rational actors into costly conflict. The first, the problem of issue indivisibility, we tend to dismiss because in principle it is not clear that any issue cannot be divided or at least compensated for with side-payments. It is still an intriguing subject of study to see how actors create indivisibility strategically in order to pre-commit and obtain better outcomes.

Second, we found that when actors possess private information that they cannot credibly reveal, there is always a real danger of conflict because rational actors would engage in the risk-return balancing by trying to obtain slightly better deals at the increased risk of conflict. In such situations, the likelihood of conflict remains strictly positive.

Third, we explored several mechanisms that can create dynamic commitment (time-inconsistency) problems. Here, actors can end up fighting because at least one of them cannot credibly promise to uphold its end of a bargain in the future. We have also seen how first-strike advantages and power transitions can create this sort of problem.

Fourth, we noted that even if actors could potentially avoid the commitment problem, they may not be able to do so. Here, the cost of peace is just too high and an actor may prefer to fight (due to a commitment problem of his opponent) rather than spend the resources necessary to secure a peaceful resolution.

These results are very pessimistic. It appears that reason alone would never be sufficient to create peace. In fact, reason can positively lead to war even between actors who find war costly, who do not make mistakes, and who would love to be able to avoid fighting.