War as a Commitment Problem*

Robert Powell

December 2004

Department of Political Science
210 Barrows Hall, 1950
University of California
Berkeley, CA 94720-1950
RPowell@Socrates.Berkeley.edu

*Forthcoming in *International Organization*. I thank James Fearon, Lisa Martin, Sebastian Mazzuca, and Branislav Slantchev for their helpful comments and criticisms. I gratefully acknowledge the support of the National Science Foundation (SES-0315037).

War as a Commitment Problem

Abstract

Although formal work on war generally sees war as a kind of bargaining breakdown resulting from asymmetric information, bargaining indivisibilities, or commitment problems, most analyses have focused on informational issues. But informational explanations and the models underlying them have at least two major limitations: They often provide a poor account of prolonged conflict as well as odd readings of the history of some cases. This paper describes these limitations and argues that bargaining indivisibilities should really be seen as commitment problems. The present analysis also shows that a common mechanism links three important kinds of commitment problem, i.e., in preventive war, preemptive attacks arising from first-strike or offensive advantages, and in conflicts resulting from bargaining over issues that affect future bargaining power. In each case, large, rapid shifts in the distribution of power can lead to war. Finally, the analysis elaborates a distinctly different mechanism based on a comparison of the cost of deterring an attack with the expected cost of trying to eliminate the threat.

War as a Commitment Problem

Formal work on the causes and conduct of war generally sees war as a kind of bargaining process.[1] As such, a central puzzle is explaining why bargaining ever breaks down in costly fighting. Because fighting typically destroys resources, the "pie" to be divided after the fighting begins is smaller than it was before the war started. This means that there usually are divisions of the larger pie that would have given each belligerent more than it will have after fighting. Fighting, in other words, leads to a Pareto inferior or inefficient outcome. Why, then, do states sometimes fail to reach a Pareto superior agreement prior to any fighting and thereby avoid war? This is the inefficiency puzzle of war.

In an important article, James Fearon (1995) described three broad rationalist approaches to resolving this puzzle: informational problems, bargaining indivisibilities, and commitment issues. The first arises when (i) the bargainers have private information about, for example, their payoffs to prevailing or about their military capabilities and (ii) the bargainers have incentives to misrepresent their private information. Informational problems typically confront states with a risk-return trade off. The more a state offers, the more likely the other state is to accept and the more likely the states are to avert war. But offering more also means having less if the other accepts. The optimal solution to this trade off usually entails making an offer that carries some risk of rejection and war.

Bargaining indivisibilities occur if the pie to be divided can only be allocated or "cut up" in a few ways. If none of these allocations simultaneously satisfy all of the belligerents, at least one of the states will prefer fighting to settling and there will be war.

The crucial issue in commitment problems is that in the anarchy of international politics states may be unable to commit themselves to following through on an agreement and may also have incentives to renege on it. If these incentives undermine the outcomes

---

[1] This formal approach can be traced at least as far back as Schelling who observed that "most conflict situations are essentially *bargaining* situations" (1960, 5). Wittman (1979) provides a pathbreaking analysis. Powell (2002) surveys this work.

that are Pareto superior to fighting, the states may find themselves in a situation in which at least one of them prefers war to peace.

Informational problems abound in international politics, and most of the formal work done in the last decade on the causes of war has pursued an informational approach to the inefficiency puzzle (e.g., Fearon 1995; Filson and Werner 2002; Kydd 2001; Powell 1996a, 1996b, 1999, 2004a; Slantchev 2003b; Wagner 2000; Werner 2000). This perspective has contributed fundamental insights, highlighted both the theoretical and empirical significance of selection effects, and yielded testable hypotheses. But, informational explanations and the models underlying them have at least two major limitations. They often provide a poor account of prolonged conflict, and they give a bizarre reading of the history of some cases.

The present analysis begins by describing these limitations and then outlines a complete-information approach to overcoming them. The basic idea behind this approach is to study war and the inefficiency puzzle in the context of complete-information games where there are no informational problems. This approach, it is important to emphasize, should not be seen as discounting the role of informational accounts in explaining key aspects of war. As just noted, informational arguments have made fundamental contributions. Rather a complete-information approach simply lets us abstract away from informational problems in order to focus more directly on other possible solutions to the inefficiency puzzle.

Appealing to bargaining indivisibilities to explain war is consistent with this complete-information approach and may seem to offer a way around the limitations of informational accounts. But they do not. The analysis below shows that bargaining indivisibilities do not offer a distinct solution to the inefficiency puzzle and should really be seen as commitment problems.

Commitment problems may help to overcome the limitations of informational accounts, either as a complement to an underlying informational problem or as the primary cause of conflict. But the concept of a commitment problem will be of little analytic value if the inability to commit leads to conflict in a different way in each empirical case. If the only thing different cases have in common is that the states are in an anarchic realm, i.e., the states are unable to commit themselves, then the concept of a commitment problem is really not doing any theoretical work and is largely serving as a catch-all label. If, therefore, the notion of a commitment problem is to provide a useful way of organizing research, it will be important to establish that a handful of general commitment problems or mechanisms illuminate a significant number of empirical cases.

To this end, the present analysis shows that the three kinds of commitment problem Fearon (1995, 401-09) describes are quite closely related. The same basic mechanism can be seen to be at work in preventive war, preemptive attacks arising from first-strike or offensive advantages, and conflicts resulting from bargaining over issues that affect future bargaining power (e.g., the fate of Czechoslovakia during the Munich Crisis or the Golan Heights). In each of these commitment problems, large, rapid shifts in the distribution of power may lead to bargaining breakdowns and war.

These results build on and extend Fearon (1995) and Powell (2004b) in two major ways. Fearon argues that bargaining indivisibilities provide a coherent rationalist explanation for war because they may eliminate the bargaining range, i.e., "the 'wedge' of bargained solutions that risk-neutral or risk-averse states will prefer to the gamble of conflict" (1995, 388). The present analysis shows that the bargaining range is not empty even if the dispute is over an indivisible issue. Indeed, the fact that fighting is costly also ensures that a bargaining range always exists even the states are risk acceptant or there are large first-strike or offensive advantages. In all three of these cases there are agreements that all of the belligerents prefer to deciding the issue through costly conflict. The problem is that the states cannot commit themselves to abiding by these agreements.

Powell (2004b) shows that a common mechanism is at work in a wide range of substantively diverse studies, namely, in Acemoglu and Robinson's (2000, 2001) model of costly coups and democratic transitions, Fearon's (2004) account of prolonged civil wars, de Figueiredo's (2002) examination of inefficient policy insulation, and Fearon's (1995) and Powell's (1999) models of interstate bargaining in the shadow of shifting power. In all of those studies, inefficient conflict results from large, rapid shifts in the distribution of power. The present analysis shows that this mechanism also explains why bargaining breaks down in war in Fearon's (1996) model of bargaining over issues that affect future bargaining power.[2] Surprisingly, the same mechanism is at work whether the distribution of military power shifts endogenously as in Fearon (1996) or exogenously (because, for example, differential rates of economic growth or sociopolitical development) as in Fearon (1995) or Powell (1999). Moreover, this mechanism also accounts for bargaining breakdowns in some models where there are large first-strike or offensive advantages.

In addition to these extensions, the present analysis describes another related mechanism which may operate at the domestic level. Here fighting results from shifts in the distribution of power between domestic factions who cannot commit to distributions of the domestic pie. Interestingly, there would be no fighting in this case if the states were unitary actors.

Finally, the discussion highlights a distinctly different mechanism based on a resource-allocation problem. Many models of war do not include the cost of securing the means of military power. There is no guns-versus-butter trade off. When these costs are included in the analysis, states may prefer fighting if the long-run cost of continually procuring the forces needed to perpetually deter an attack on the status quo on is higher than the expected cost of trying to eliminate the threat.

The next section elaborates two major limitations of informational explanations. Section three describes the complete-information approach. The fourth section shows that bargaining indivisibilities do not solve the inefficiency puzzle and that the real issue is commitment. Section five takes up commitment problems.

The Limitations of Informational Explanations

Most informational explanations of war begin with a bargaining model in which there would be no fighting if there were complete information. The analysis then adds asymmet-

[2] To the best of my knowledge we only have two models that focus on this kind of bargaining problem, Fearon (1996) and Schwarz and Sonin's (2003) closely related analysis.

ric information and shows that there is a positive probability of fighting in equilibrium.[3] But using models in which there would be no fighting if the states had complete information tends to create an analytic blind spot. This blind spot leads to strained or even bizarre historical readings of some cases.

Consider first prolonged international and intrastate war and the ultimate ability of asymmetric-information bargaining models to provide a compelling explanation of this outcome.[4] An informational approach would generally argue that prolonged fighting results from rival factions' efforts to secure better terms by demonstrating their "toughness" or resolve. Moreover, we also ought to find that significant informational asymmetries exist throughout the conflict as these are a pre-requisite for continued fighting. But these asymmetries sometimes appear to be lacking. Based on his study of civil wars, Fearon observes that while asymmetric information may explain the early phases of some conflicts, it does not provide a very convincing account of prolonged conflict. "[A]fter a few years of war, fighters on both sides of an insurgency typically develop accurate understandings of the other side's capabilities, tactics, and resolve" (2004, 290). Asymmetric information does not appear to explain these conflicts.

A second limitation of the information approach follows from an under-appreciated implication of the assumption that there would be no fighting if there were complete information. A satisfied state always prefers appeasing a dissatisfied adversary state to fighting *no matter how large a concession it takes to satisfied the dissatisfied state.* To

[3] See Fearon (1995) and Powell (1996a,b; 1999, 86-97) for typical formulations. These informational efforts to explain inefficient fighting parallel earlier efforts in economics to explain inefficient delay in bargaining. Rubinstein's (1982) seminal analysis found that a very natural bargaining game had a unique subgame perfect equilibrium when there was complete information. The equilibrium outcome was also efficient: the first offer was accepted and agreement was reached without delay. Economists initially believed that adding asymmetric information would provide a straightforward explanation of delay. But explaining delay in this way proved far from straightforward. See for example the discussion in Gul and Sonnenshein (1988,).

[4] Protracted interstate conflict turns out to be relatively rare. Only six out of the seventy-eight wars fought during 1816-1985 lasted five or more years (Bennett and Stam 1996). By contrast, civil wars are much more likely to last a long time. Seventy of the 123 civil wars started between 1945 and 1999 lasted at least five years and thirty-nine lasted at least ten years (Fearon 2004).

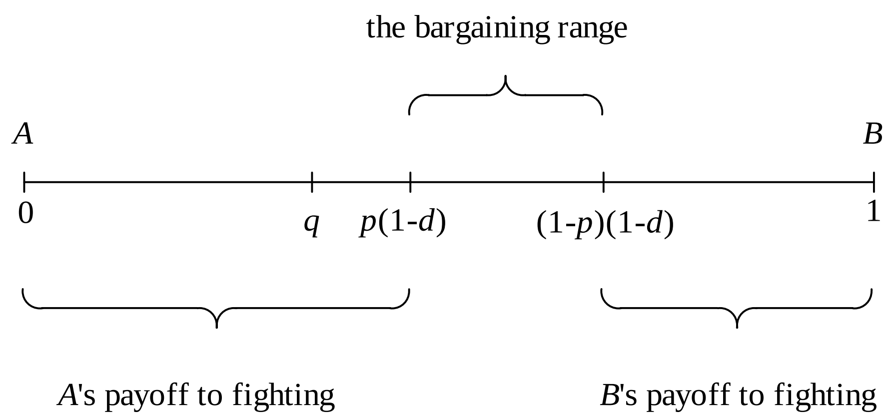

Figure 1: The Bargaining Range

illustrate this implication, consider a simple take-it-or-leave-it offer game in which two states, say A and B, are bargaining about revising the territorial status quo, q.[5] As depicted in Figure 1, A controls all of the territory to the left of q at the start of the game, and B controls all of the territory to the right of $q \in [0, 1]$.

B begins the game by making an offer, $x \in [0, 1]$, to A who can accept the offer, reject it, or go to war to change the territorial status quo. If A accepts, the territory is divided as agreed. If A fights, the game ends in a costly lottery in which one state or the other is eliminated. More precisely, either A wins all of the territory and eliminates B with probability p, or B eliminates A and thereby obtains all of the territory with probability $1 - p$. Fighting also destroys a fraction $d > 0$ of the value of the territory. If A rejects B's offer, then B can attack or pass. Attacking again ends the game in a lottery. Passing ends with the status quo unchanged.

A's payoff if the status quo is unchanged is q, its payoff to agreeing to x is just x, and its payoff to fighting is $p(1 - d) + (1 - p)(0) = p(1 - d)$. B's payoffs are defined analogously. A state is *dissatisfied* if it prefers fighting to the status quo. Thus, A is dissatisfied if $p(1 - d) > q$, and B is dissatisfied if $(1 - p)(1 - d) > 1 - q$.

Suppose then that A is dissatisfied as depicted in Figure 1 and that there is complete information. In these circumstances, B knows the minimum amount it must offer A

[5] See Fearon (1995) and Powell (1999, 2002) for elaborations of this basic setup.

in order to induce A not to fight. To wit, B must offer A its certainty equivalent of fighting $x^* = p(1-d)$. This offer makes A indifferent between fighting and accepting, and, consequently, A would strictly prefer to fight if offered less than x^*.[6]

Thus, B faces a clear choice when there is complete information. It can appease A by conceding x^*, which leaves B with a payoff of $1 - x^* = 1 - p(1-d)$, or B can fight which gives it an expected payoff of $(1-p)(1-d) = 1 - p(1-d) - d$. B clearly prefers the former as long as fighting is costly (i.e., as long as $d > 0$) and *regardless of how much it has to concede* (i.e., regardless of how much larger x^* is than q). Hence, B always prefers to accommodate A whenever A is dissatisfied, fighting is costly, and there is complete information.

A simple intuition underlies this result. If fighting is costly, the pie to be divided is larger if the states avert war because they save d. But B's offer of A's certainty equivalent $x^* = p(1-d)$ means that A's payoff is the same whether it accepts x^* or fights. Thus, whatever is saved by not fighting must be going to B, and this is what leads B to prefer appeasing A.[7]

B's choice is less clear when there is asymmetric information. Suppose A has private information about its military capabilities, e.g., about the effectiveness of its military forces. As a result, B is unsure of A's probability of prevailing but believes that it lies in a range from $\underline{p}$ to $\overline{p}$. This uncertainty confronts B with a risk-return trade off. The more it offers A, the more likely A is to accept but the less well off B will be if A accepts. The optimal offer that resolves this trade off generally entails some risk of rejection, and this is the way that asymmetric information can lead to war.

The implicit assumption that the states do not fight when there is complete information can produce strange historical accounts. Consider, for example, the run up to the Second World War in Europe. It is impossible to tell the story of the 1930s without asymmetric information. There was profound uncertainty surrounding Hitler's ambitions and much

6 Although A is indifferent between fighting and accepting x^*, it can be shown that A is sure to accept x^* in equilibrium.

7 To put the point formally, the difference between B's payoff to satisfying A and fighting is just the amount that fighting would have destroyed: $(1 - x^*) - (1-p)(1-d) = d$.

that we see is consistent with an informational account. Throughout the 1930s, Britain and France made a series of ever larger, "screening" concessions which they hoped would appease Germany.[8]

But, war did not come, as an informational account would have it, because Britain and France would have been willing to satisfy Hitler's demands *if only they had complete information about what those demands were and offered too little because of that uncertainty.* To the contrary, Britain became increasingly confident after Hitler occupied the rump of Czechoslovakia that it was dealing with an adversary it was unwilling to satisfy. Of Hitler's demand for a "free hand in the East," Foreign Secretary Halifax wrote to Chamberlain a few days before the war began, "if he [Hitler] really wants to annex land in the East..., I confess that I don't see any way of accommodating him."[9] Uncertainty still existed on the eve of war, but Britain and Germany appear to have been "types" – to use the language of game theory – that would have fought each other even if there were no uncertainty. The maximum Britain was willing to concede (at least over the long run) was less than what was required to satisfy Hitler.

This of course is a judgement about a counterfactual, about what would have happened had the states had complete information about each other's ambitions. The historiography of the origins of the Second World War is vast, and a detailed treatment is clearly beyond the scope of the present analysis. But a key issue is that even if Britain were unwilling to give Hitler a free hand in the East as proved to be the case, did war result because Germany was uncertain about Britain's determination to stand firm and not give way as it had during the Munich crisis. Three related aspects of this issue are: If Hitler had concluded in late August 1939 that Britain would stand firm, would he still

[8] In a screening equilibrium, an actor with incomplete information makes a series of ever more attractive offers that screen the other actor according to the latter's willingness to settle. Suppose *1* is uncertain of *2*'s degree of dissatisfaction, e.g., *2*'s payoff to fighting. Then, *1* makes a series of increasingly favorable concessions to *2* in the hope of buying *2* off at the cheapest possible price. The less dissatisfied *2* is, i.e., the lower its payoff to fighting, the earlier it settles. These offers therefore "screen" *2* according to its willingness to settle. For analyses of these dynamics, see Powell (2004a) and Slantchev (2003b, 2004a).

[9] Quoted in Parker (1993, 268). Also see Aster (1973, 328) and Weinberg (1980, 654).

have attacked Poland? Had he known a few months earlier that Britain would not back down, would he still have pressed the Poles and instigated the crisis in the spring of 1939? Most importantly, would knowing that Britain would fight have deterred Hitler over the long run from using force to pursue his ends in the East? The answers to the first two questions pertain most directly to the timing of the war. The third centers on whether there would have been a war if there were complete information.

Hitler's early plan or ideas about acquiring Lebenstraum was to annex Austria and Czechoslovakia first, then secure Germany's western front by defeating France, and finally turn east and on the Soviet Union. He did not see a fundamental conflict between Britain and Germany, and sometimes thought of an Anglo-German alliance. But by the fall of 1937, Hitler had come to see Britain as a potential obstacle to his eastern ambitions. Outlining his views at the "Hossbach" conference on November 3, 1937, Hitler declared that the only remedy for Germany' security problems "lay in the acquisition of greater living space [in Europe]" and that Germany had to reckon with Britain and France who "were opposed to any further strengthening of Germany's position either in Europe or overseas" (DGFP 1949-64, 1:31-32).[10]

Almost immediately after occupying what remained of Czechoslovakia in March 1939, Germany began to press Poland. Two months into the crisis, Hitler told his military commanders that the issue had gone beyond Danzig.

> It is not Danzig that is at stake. For us it is a matter of expanding our living space in the East... We cannot expect a repetition of Czechia. There will be war. Our task is to isolate Poland... It must not come to a simultaneous showdown with the West (France and England) (DGFP 1949-64, 6:576.).[11]

Isolating Poland by trying to ensure that Britain would not fight remained a critical

[10] Friedrich Hossbach was Hitler's adjutant at the time, and the record of the meeting is based on his notes. A. Taylor (1962, 131-35) challenged the significance and conventional interpretation of the meeting, but his views have generally been discounted and rebutted. Rich (1973, 287-88) and T. Taylor (1979, 302-7) discuss the controversy and Martel (1986) offers a broader reconsideration of Taylor's analysis of the origins of the war. For overviews of Hitler's thinking about Britain, see Rich (1974, 394-96), Waddington (1996, 22-29), and Weinberg (1970, 1-24).

[11] On the crisis over Danzig and Poland, see Aster (1973, 188-215), Sontag (1971, 356-57), Thorne (1967, 113-52), Watt (1989), Weinberg (1980, 535-627).

part of Hitler's strategy in the spring and summer of 1939.[12] This was a primary motivation for the Nazi-Soviet Pact, and Hitler believed as late as mid-August that Britain might back down as it had over Czechoslovakia. On August 22, three days before the planned attack, Hitler told his Commanders-in-Chief, "Now the probability is still great that the West will not intervene. We must take the risk... England and France have undertaken obligations which neither is in a position to fulfil" (DGFP 1949-64, 7:201-02).

The attack on Poland was originally scheduled for August 26 and Hitler gave the order to carry it out around 3 p.m. the day before. He canceled it less than four hours later. In between, he learned that Britain had ratified the Anglo-Polish alliance formalizing the guarantee of Poland's security that Chamberlain had announced in Parliament at the end of March. Hitler also learned that Italy was not ready to fight.[13]

When cancelling the attack, he told General Keitel, Chief of the High Command of the Wehrmacht, "I need more time for negotiations." When Göring asked if the attack was being put off permanently or temporarily, Hitler said that it was temporary in order to "see whether we can eliminate British intervention."[14] There followed another week of negotiations during which Hitler hoped to drive a wedge between the British, French, and Poles.[15] On August 31, the eve of the German attack, Hitler, according to Chief of staff of the Army Halder, "expects France and Britain not to strike" (Halder 1988, 44).

After Poland fell in September, Hitler renewed his efforts to secure British acquiesce. Britain rebuffed these attempts, leading Hitler on October 9 to direct:

> Should it become evident in the near future that England and, under her influence, France also, are not disposed to bring the war to an end, I have decided without further loss time, to go over to the offensive (Trevor-Roper 1964, 13).

Hitler initially intended that the attack on France take place a few weeks later in November 1939. But opposition from his generals and the weather convinced him to postpone

[12] Hitler generally assumed that France would not fight without Britain (Weinberg 1974, 629).

[13] See Parker (1933, 215 and 331-36), Watt (489-94), and Weinberg (1980, 633-39).

[14] See Thorne (1987, 185) and Aster (1973, 338) on Hitler's comment to Keitel and Rich (1973, 129) on Hitler's comment to Goring.

[15] On these goals, see Rich (1973, 130), Sontag (1971, 380), Watt (1989, 508).

the attack until the next spring (Lukacs 1976, 59-60; Rich 1974, 148-50).

After France fell in June 1940, Hitler tried yet again to convince Britain to acquiesce. But Britain still refused to negotiate, and Hitler came to believe by the end of the month that "Britain probably still needs one more demonstration of our military might before she gives in and leaves us a free hand in the east" (Halder 1988, 219). On July 16, Hilter issued Directive 16:

> Since England, in spite of her hopeless military situation, shows no signs of being ready to come to an understanding, I have decided to prepare a landing operation against England and, if necessary, carry it out (Trevor-Roper 1964, 34).

On August 1, Hitler ordered the Luftwaffe to begin an air offensive against Britain in order to gain control of the skies over the invasion routes and pave the way for a landing. "In order to establish the necessary conditions required for the final conquest of England ... The German Air Force is to overpower the English Air Force with all of the forces at its command in the shortest possible time."[16]

Would there have been war had there been complete information? Incomplete information was clearly present throughout the confrontation. Hitler's up and down hopes that Britain would acquiesce to Germany's continental ambitions demonstrates that he was unsure of Britain's determination.[17] But trying to keep Britain out of the war and induce it to settle without directly attacking it does not show that incomplete information caused the conflict. Trying to compel Britain to come to terms without having to pay the cost of defeating her would have made sense whether or not Hitler would have been willing to fight had he known that Britain would stand firm.

Would Hitler have attacked had he known that Britain would not give him a free

[16] Trevor-Roper (1964, 37). Wheatley (1958, 57) discusses this order and context.

[17] Indeed, Hitler's thinking in the summer of 1940 about the future invasion of the Soviet Union was based at least in part on the idea that defeating the Soviet Union would, at last, compel Britain to come to terms. Some have argued on this basis that Germany never intended to invade Britain. Rich (1973, 160-64, 208-10) examines and rejects this claim. "There can be no doubt of the seriousness of his [Hitler's] invasion plans plans from July to September of that year [1940]. During this period the entire German economy and transport system were disrupted by invasion preparations" (Rich 1973, 160). Lukacs (1976, 102-11) and Wheatley (1958, 44-5, 133-37) also reach the same conclusion.

hand in the East? Whether Hitler would have gone ahead with the attack on Poland had concluded in late August that Britain would fight is uncertain. Gerhard Weinberg, drawing on his extensive study of Hitler's foreign policy, believes that Hitler would have attacked (1980, 654-65). And the evidence is consistent with what we would expect to observe had Hitler preferred to fight an isolated Poland but was still willing to fight even if he knew that Britain would intervene. But the evidence, at least prior to the summer of 1940, is also consistent with an informational account in which Hitler was willing to run some risk of war but would not have pushed ahead if he were certain that Britain would stand firm. Hitler's numerous statements that Britain would not fight, his postponing the original attack in order to try to drive a wedge between Britain and Poland, and his comment to Halder that on the eve of the attack that he still expected Britain and France not to take action all indicate that Hitler thought there was some chance Britain would not fight and this leaves open the question of what would he would have done had he been sure that Britain would fight.

More certain is that Hitler would not have pressed Poland in the spring of 1939 had he known Britain would not give in as it over the Czechoslovakia. At the Hossbach conference in November 1937, Hitler explained that Germany's relative military strength would peak in 1943-45, and "it was his unalterable resolve to solve Germany's problem of space at the latest by 1943-45" (DGFP 1949-64, 1:34-35). But, he foresaw the possibility of acting sooner should a favorable opportunity arise.[18] These same basic ideas were also present in his May 1939 meeting. He planned to "attack Poland at the first suitable opportunity" which meant isolating Poland so that the crisis did not "come to a simultaneous showdown with the West" (DGFP 1949-64, 6:576). But Hitler doubted that in the long run a "peaceful settlement with England is possible. It is necessary to be prepared for a showdown" and the armament program preparing for this "will be complete by 1943 or

[18] The specific context of this discussion is moving against Austria and Czechoslovakia, not Poland, and the favorable opportunity was that France would be unable to intervene against Germany because of internal striff in France or because France was already embroiled in another war (e.g., an Anglo-French-Italian war growing out of tensions in the Mediterranean). See GDFP (1949-64, 1:35-38).

1944" (DGFP 1949-64, 6:576, 580). This suggests that Hitler preferred avoiding a direct confrontation with the West until Germany was ready for it, but that he would also take advantage of favorable conditions. Those conditions arose in the spring of 1939. On the one hand, Polish mobilization and unwillingness to bow to German demands convinced Hitler that he could not count on Poland staying out of a confrontation between France and Germany. And, on the other hand, Hilter also concluded after Britain and France acquiesced in the Germany's occupation of what remained of Czechoslovakia in March 1939 that they would also back down over Poland. Had Hitler known that Britain would stand firm, that a favorable opportunity had not yet arrived, it seems doubtful that he would have forced a confrontation before Germany was ready.[19]

These considerations have to do with the timing of the war, whether Hitler could have been dissuaded from instigating the crisis over Danzig and invading Poland in the summer of 1939. The broader, more important issue is whether Hitler was willing to fight Britain at some point in the future in order to secure his ends in the east. It seems clear that he was. In the summer of 1940, he *did* attack Britain in an effort to defeat her and free himself to pursue his eastern ambitions. The air assault ultimately failed to destroy the British air force and secure command of the skies, and the planned invasion was eventually canceled as Hitler's attention shifted to the Soviet Union. But had any of Hitler's "military plans given promise of a quick and decisive victory over Britain, he would almost certainly have moved in for the knock-out blow before doing anything about Russia" (Rich 1974, 209). Hitler attacked Britain in the summer of 1940 because she was standing firm, not because he was uncertain whether she would stand firm. Germany attacked inspite of, not because of, some residual uncertainty about Britain's determination to stand firm.

In sum, the informational approach has developed in the context of models in which there would be no fighting if states had complete information about each other. These models and the accounts based on them explain important aspects of many cases. But

[19] Sontag reaches a similar conclusion. Hitler "would have have settled temporarily for less than intended, as he did at Munich..." in order to avert or postpone war with Britain or France (1971, 332).

these accounts also provide an odd reading of other equally important aspects of some cases. In important cases, fighting does not seem to result from some residual uncertainty about an adversary that has yet to be resolved. *Fighting ensues when the resolution of uncertainty reveals that a state is facing an adversary it would rather fight than accommodate.* Such cases are not well modeled by the standard informational account in which bargaining invariably leads to efficient outcomes when there is complete information.

A Complete-Information Approach to Costly Conflict

Situations in which war breaks out when a state becomes increasingly confident that it is facing an opponent it would rather fight than accommodate combine two problems. The first is an informational problem created by the state's initial uncertainty about its adversary's capabilities or resolve. This uncertainty played a critical role in the 1930's, and, as Fearon observes, it may also play an important part in the early phase of many civil wars. The second problem is the possibility that there are "types" that would fight each other even if there were no uncertainty. If such types actually are facing each other, then war will come to be seen as more rather than less likely as the states learn more about each other. At some point, one of the states becomes sufficiently confident it is facing a type it is unwilling to accommodate and attacks.

By focusing almost exclusively on models in which there would be no fighting if the states had complete information, recent formal work on war has treated it as a purely informational problem.[20] This focus limits this work's ability to explain cases in which the fundamental cause of war is not incomplete information but something else, like a commitment problem, which would lead to war even if the states had complete information. How, then, can we study these other causes?

Although actual cases may combine informational problems with these other potential causes, we can separate the informational problem and set it aside analytically. Models that incorporate asymmetric information, which is needed to study the informational problem, tend to be complex. This complexity typically forces the modeler to simplify

[20] Three recent exceptions are Fearon (2004), Powell (2004b), and Slantchev (2003a).

other aspects of the states' strategic environment in order to keep the model tractable. We can, however, abstract away from the information problem by working with complete-information games. These models in effect posit that the states already know or have learned whom they are facing. As a result, this complete-information approach focuses directly on trying to illuminate the key features of a strategic environment that may lead to costly, inefficient fighting even if the states have no private information.

Bargaining Indivisibility as a Commitment Problem

Bargaining indivisibilities appear to provide a simple, straightforward solution to the inefficiency puzzle of explaining why states fight even though there are peaceful agreements they all simultaneously prefer to war. If the disputed issue is indivisible or can only be divided in a limited number of ways, one state or the other may prefer fighting to each of these divisions. There are no Pareto superior peaceful settlements, and the question of why the states fail to agree on one is moot. Moreover, an appeal to bargaining indivisibilities would also seem to be part of a complete-information approach to the inefficiency puzzle because incomplete information plays no role in this argument.

This reasoning is flawed. Even if the disputed issue is indivisible, there are still agreements both sides prefer to resolving the issue through costly fighting. Bargaining indivisibilities, therefore, do not solve the inefficiency puzzle. The problem is, rather, that the states cannot commit to these agreements. More generally, the fact that fighting is costly implies that a bargaining range always exists even if the states are risk-acceptant, the issue is indivisible, or there are first-strike or offensive advantages. While Fearon (1995, 388-90, 402-04) only describes the latter as a commitment problem, all three are fundamentally commitment problems.

That bargaining indivisibilities do not offer a distinct rationalist explanation for war runs contrary to the growing literature on bargaining indivisibilities. Although he discounted their empirical significance, Fearon argued that bargaining indivisibilities offered a conceptual solution to the inefficiency puzzle. If the issue were indivisible, there might not be any agreements that all of the states simultaneously preferred to fighting (1995, 386-90). "[I]n principle, the indivisibility of the issues that are the subject of international bargaining can provide a coherent rationalist explanation for war. However, the real question in such cases is what prevents leaders from creating intermediate settlements... Both the intrinsic complexity and richness of most matters over which states negotiate and the availability of linkages and side-payments suggest that intermediate bargains typically will exist" (1995, 390).

Others have begun to assert more recently that bargaining indivisibilities are more common and play a more important role in international disputes. Hassner (2003) believes that sacred places are often seen as inherently indivisible and that this perception impedes efforts to resolve disputes over them. Goddard (2003) and Hassner (2004) endogenize indivisibility. For Goddard, indivisibility is "constructed by the actors during the bargaining process" through the actors' efforts to justify or legitimate their claims (nd, 3). Whether an issue comes to be seen as indivisible depends on the legitimation strategies the parties use while bargaining. Hassner(2004) links indivisibility to entrenched territorial disputes, arguing that as territorial disputes persist the disputed territory comes to be viewed as indivisible. And, Toft (2002/3) explains ethnic violence in terms of territorial indivisibility.

Bargaining indivisibilities, however, do not explain war. Even if a disputed issue is physically indivisible, we should not think of bargaining indivisibilities as a conceptually distinct solution to the inefficiency puzzle. There are still outcomes (or more accurately mechanisms) that give both states higher expected payoffs than they would obtain by fighting over the issue. The real impediment to agreement is the inability to commit.

To see that this is the case, suppose that the territory over which A and B are bargaining in the example above cannot be divided.[21] Either A will control all of the territory or B will. War can be seen as a costly way of allocating this territory. More specifically, A obtains the territory with probability p, B gets the territory with probability $1-p$, and fighting destroys a fraction d of its value. The states' payoffs to allocating the territory

[21] This analysis draws on Fearon (1995, 389) who briefly discusses the possibility of resolving bargaining indivisibilities through "some sort of random allocation" and on Wagner's (2004) insightful discussion.

this way are $p(1-d)$ and $(1-p)(1-d)$ for A and B, respectively. But now suppose that the states simply agree to award the territory to A with probability p and to B with probability $1-p$. This agreement gives the states expected payoffs of p and $1-p$. Both states clearly prefer allocating the territory this way to allocating it through costly fighting. *Thus, there exist agreements that Pareto dominate fighting even if the issue is indivisible.* The the inefficiency puzzle is not moot, and the question remains: Why do the states fail to secure a Pareto efficient outcome?

The example above is based on a take-it-or-leave-it bargaining protocol. But the basic point is much more general. Abstractly, we can think of fighting over an indivisible object as a costly way of allocating it: each state gets the object with a certain probability and at some cost. It follows that both states would prefer an agreement that gives the object to them with the same probabilities but does so without their having to pay the cost of fighting. The problem is not that there are no agreements that are Pareto superior to fighting; the fact that fighting is costly ensures that there are. The problem is that states may not be able to commit themselves to abiding by these agreements.

Somewhat more formally, suppose that the possibly very complicated way of settling a dispute can be represented by a complete-information game, say Γ. In the example above, Γ was a take-it-or-leave-it-offer game. If the states play Γ, then we can characterize an equilibrium outcome in terms of the probability π_A that the issue is resolved in A's favor, the probability π_B that the issue is resolved in B's favor, and the expected fractions of the value destroyed if A prevails and if B prevails, d_A and d_B.[22] The states' equilibrium payoffs can then be written as $\pi_A(1-d_A)$ for A and $\pi_B(1-d_B)$ for B.[23]

[22] The complete-information assumption comes in here. This assumption implies that the states share the same probability distribution over terminal nodes of the tree. Hence, the probability A attaches to B's prevailing is the same as the probability that B gives it, and similarly for the probability that A prevails. This means that π_A, π_B, d_A, and d_B are well defined.

[23] Let $\mathcal{A}$ denote the outcomes or teminal nodes of Γ at which A prevails; take A's payoff at $j \in \mathcal{A}$ to be $1 - d_A^j$; and let the equilibrium probability of reaching outcome j be π_A^j. A's payoffs at all other outcomes is zero. Hence, A's expected equilibrium payoff is $\sum_{j \in \mathcal{A}} \pi_A^j(1-d_A^j) = \pi_A(1-d_A)$ where $\pi_A = \sum_{j \in \mathcal{A}} \pi_A^j$ is the probability that A prevails and $d_A = \sum_{j \in \mathcal{A}}(\pi_A^j/\pi_A)d_A^j$ is the expected cost of fighting conditional on A's prevailing.

As long as this way of settling the dispute is costly (i.e., $d_A > 0$ and $d_B > 0$), then there is *always* a strictly Pareto superior settlement even if the issue is indivisible. Namely, the issue is costlessly settled in A's favor with probability π_A and in B's favor with probability π_B. Settling the issue in this way avoids the cost of fighting and gives the states the higher payoffs of π_A and π_B. Thus, there are agreements both states strictly prefer to resolving the dispute through the costly mechanism Γ. The difficulty is not the absence of Pareto improving agreements but the inability of the states to commit to them.

An analogy may help make this more concrete. In order to avoid the high cost of litigation, the parties involved in a contractual dispute will prefer to settle the matter through binding arbitration as long as the chances of prevailing are roughly the same as they would be if the dispute went to court. In these circumstances, arbitration reduces the cost of resolving the dispute and both parties are better off. Of course, this requires that the arbitration be truly binding. If the party that loses in arbitration can then go to court, arbitration has little to offer. By analogy, the problem with bargaining indivisibilities is not the absence of agreements that states prefer to fighting. The problem is that the states may not be able to commit to following through on them.

The argument above goes beyond bargaining indivisibilities, and the fundamental similarity between bargaining indivisibilities, risk acceptance, and large first-strike or offensive advantages is worth emphasizing. Fearon suggests that the states must be risk averse or risk neutral in order to guarantee that a bargaining range exists. Risk acceptance may eliminate the bargaining range (1995, 388). It is also easy to misread Fearon as saying that sufficiently large first-strike or offensive advantages can close the bargaining range (1995, 403).[24] However, the fact that fighting is costly implies that there are always agreements the states prefer to fighting even if there are indivisibilities, the states are risk acceptant, or there are large first-strike or offensive advantages. If we think of war

[24] Fearon distinguishes between the bargaining range and what he calls the "de facto bargaining range," which is the difference between each state's reservation value for fighting given that it strikes first (1995, 403). Large first-strike or offensive advantages eliminate the de facto bargaining range. The commitment problem created by these advantages is discussed in more detail below.

as a costly lottery, all of the states would do better by agreeing to the equivalent costless lottery, i.e., a lottery in which the states' chances of winning are the same and there are no costs.[25] The problem with these agreements in each case is the same: the states have incentives to renege on them.

In sum, bargaining indivisibilities do not solve the inefficiency puzzle by rendering it moot. The bargaining range is not empty; there are always agreements that all of the states simultaneously prefer to war. Bargaining indivisibilities, risk-acceptant states, and first-strike or offensive advantages should all been seen as commitment issues. Broadly speaking, there are two, not three rationalist approaches to the inefficiency puzzle of war: informational problems and commitment problems.[26]

Commitment Problems

If the notion of a "commitment problem" is to provide a useful explanation of some aspects of war, this concept must be more than a catch-all label. If a different mechanism seems to be at work in each historical case, then the broader notion of a commitment problem will not be of much analytic value. Formalizing these mechanisms may still be useful, but grouping them together under the label "commitment problems" is not really doing any additional theoretical work. The potential value-added of the broader notion of a commitment problem lies in the possibility that a few basic mechanisms will turn out to illuminate a significant number of cases.

Fearon (1995, 401-09) offered a start in this direction by identifying three kinds of

[25] More generally, the equivalent costless lottery induces the same probability distribution over possible outcomes as does the costly lottery. I am grateful to Fearon (private correspondence) for pointing out that the basic argument developed here in the context of bargaining indivisiblities also extends to risk-acceptant states.

[26] Fearon (private correspondence) suggests another approach based on coordination problems. These arise in the context of complete-information games in where there are multiple equilibria, some of which entail costly fighting. The stag hunt is a simple example (see Jervis 1978) and Slantchev (2003a) provides a more recent example. In these models the equilibria in which there is fighting are strictly Pareto inferior to equilibria in which there is no fighting. Absent a compelling theory of equilibrium selection, inefficient equilibria that are dominated by efficient ones provide at best a weak resolution of the inefficiency puzzle.

commitment problem: preventive war triggered by an anticipated shift in the distribution of power, preemptive war caused by first-strike or offensive advantages, and war resulting from a situation in which concessions also shift the military balance and thereby lead to the need to make still more concessions. This section shows that these problems are closely related. They can be seen more generally as different manifestations of the same more basic mechanism. The section also describes an analogous domestic-level mechanism. Here the inability of domestic factions to commit to divisions of the domestic pie leads to international conflict. Finally, the analysis discusses a very different mechanism based on a comparison of the cost of defending the status quo to the expected cost of trying to eliminate the threat to the status quo.

A General Inefficiency Condition: To see the connection between Fearon's three commitment problems, we need to take a step back. Recent work in American, comparative, and, to some extent, international politics has tried to explain inefficient outcomes in a complete-information setting.[27] Powell (2004b) shows that a common mechanism is at work in several of these studies, namely, in Acemolgu and Robinson's (2000, 2001) study of political transitions, Fearon's (2004) analysis of prolonged civil war, de Figueiredo's (2002) account of costly policy insulation, and Fearon's (1995) and Powell's (1999) examination of preventive war.

Although the substantive contexts differ widely, the bargainers in each of these cases face the same fundamental strategic problem. The bargainers are in effect trying to divide a flow of benefits or "pies" in a setting in which (i) the bargainers cannot commit to future divisions of the benefits (possibly because of anarchy, the absence of the rule of law, or the inability of one Congress to bind another); (ii) each actor has the option of using some form of power – mounting a coup, starting a civil war, or launching a preventive attack – to lock in a share of the flow; (iii) the use of power is inefficient in that it destroys some

[27] See, for example, Acemoglu and Robinson (2000, 2001, 2004) on democratic transitions, costly coups, and revolutions; Fearon (1998, 2004) on ethnic conflict and civil war; Alesina and Tabellini (1990) and Persson and Svensson (1989) on inefficient levels of public debt; Besley and Coate (1998) on democratic decision making; Busch and Muthoo (2002) on sequencing; de Figueiredo (2002) on policy insulation; and Fearon (1995, 404-08), Powell (1999, 128-32), and Slantchev (2003a) on war.

of the flow; and (iv) the distribution of power, i.e., the amounts the actors can lock in, shifts over time.

Complete-information bargaining breaks down in this setting if the shift in the distribution of power is sufficiently large and rapid. To see why, consider the situation confronting a temporarily weak bargainer who expects to be stronger in the future (i.e., the amount that this bargainer can lock in will increase). In order to avoid the inefficient use of power, this bargainer must buy off its temporarily strong adversary. To do this, the weaker party must promise the stronger at least as much of the flow as the latter can lock in. But when the once-weak bargainer becomes stronger, it may want to exploit its better bargaining position and renege on its promised transfer. Indeed, if the shift in the distribution of power is sufficiently large and rapid, the once-weak bargainer is certain to want to renege. Foreseeing this, the temporarily strong adversary uses it power to lock in a higher payoff while it still has the chance.

To sketch the idea more formally, suppose that two actors, *1* and *2*, are trying to divide a flow of pies where the size of the pie in each period is one. The present value of this flow is $B = \sum_{n=0}^{\infty} \delta^n = 1/(1-\delta)$ where δ is the bargainer's common discount factor. At time t player $j =$ *1* or *2* can lock in a payoff of $M_j(t)$ but doing so is inefficient because it destroys some of the flow. More concretely, $M_1(t)$ might be *1*'s expected payoff to going to war as in Fearon (1995) and Powell (1999), deposing the faction in power as in Acemoglu and Robinson (2000, 2001), fighting a civil war as in Fearon (2004), or bureaucratically insulating a policy from one's political adversaries as in de Figueiredo (2002). If, for example, *1* locks in its payoff by fighting, then

$$M_1(t) = p_t\left(\frac{1-d}{1-\delta}\right) + (1-p_t)\left(\frac{0}{1-\delta}\right) = \frac{p_t(1-d)}{1-\delta}$$

where p_t is the probability that *1* wins the entire flow less the fraction d destroyed by fighting. More generally, $M_j(t)$ is j's minmax payoff in the continuation game starting at time t. Because j can always assure itself a payoff of at least $M_j(t)$ starting from time t, j's payoff (starting from t) must be at least as large as $M_j(t)$ in any equilibrium.

Now consider the states' decisions at time t if they expect the distribution of power to shift in *1*'s favor. That is, the payoff *1* can lock in increases from $M_1(t)$ to $M_1(t+1)$ in the next period. If the temporarily weak *1* is to induce *2* not to exploit its temporary advantage, *1* must promise *2* at least as much as *2* can lock in, i.e., *1* must offer at least $M_2(t)$. To this end, the most that *1* can give its adversary in the current period is the entire pie. As for the future, the most that *1* can *credibly* promise to give to *2* is the (discounted) difference between what there is to be divided and what *1* can assure itself, namely, $B - M_1(t+1)$. Were *1* to offer *2* more than this, then *1* would be promising implicitly to accept less that $M_1(t+1)$ for itself. But such a proposal is inherently incredible because *1* can always lock in $M_1(t+1)$ and therefore would never accept less than this. Hence, the most that *1* can credibly offer *2* at time t is $1 + \delta[B - M_1(t+1)]$. If this amount is less than what *2* can lock in, i.e., if $M_2(t) > 1 + \delta[B - M_1(t+1)]$, then *2* prefers fighting. In these circumstances *1*'s inability to commit to giving *2* a larger share results in the inefficient use of power.

Rearranging terms and subtracting $M_1(t)$ from both sides of the previous inequality give the inefficiency condition:[28]

$$\delta M_1(t+1) - M_1(t) > B - [M_1(t) + M_2(t)]\,. \tag{1}$$

This condition has a natural substantive interpretation. The left side is a measure of the size of the shift in the distribution of power between times t and $t+1$ (and, therefore, of the rate at which the distribution of power is shifting). The right side is the bargaining surplus, i.e., the difference between what there is to be divided and what each player can assure itself on its own. Thus, the inability to commit leads to inefficient outcomes when

[28] Powell (2004b) shows formally that all of the equilibria of a stochastic game are inefficient whenever this condition holds somewhere along every efficient path.

the per-period shift in the distribution of power is larger than the bargaining surplus.[29]

Shifting Power between States: Shifts in this distribution of power are at the heart of Fearon's three kinds of commitment problems. As Powell (2004b) shows, condition (1) explains the breakdown in Fearon's model of war (1995, 404-8). Following Fearon, suppose that the territorial bargaining game described above lasts infinitely many rounds rather than just one and that *2* makes an offer to *1* in each round. Assume further that the distribution of power is expected to shift in *1*'s favor. Formally, the probability that *1* prevails in the first round, p, increases to $p+\Delta$ in the second round, and remains constant thereafter.

State *2* prefers fighting in equilibrium to appeasing *1* if the adverse shift in power Δ is sufficiently large. To establish this, observe that *2*'s payoff to fighting in the first round is $(1-p)(1-d)/(1-\delta)$. If, by contrast, *2* does not fight, its payoff in round one is certainly no more than one which it would get if it controlled all of the territory. Once the distribution of power shifts, state *2* must offer *1* its certainty equivalent to fighting $x^* = (p+\Delta)(1-d)$ in order to induce *1* not to fight. This means that the best that *2* can do if it decides not to fight at the outset of the game is $1+\delta(1-x^*)/(1-\delta)$. State *2*, therefore, prefers fighting to accommodating if $(1-p)(1-d)/(1-\delta) > 1+\delta(1-x^*)/(1-\delta)$. This relation in turn is sure to hold if *2*'s gain from fighting now rather than later is larger than the cost of fighting, i.e., if $\Delta(1-d) > d$, and the discount factor is sufficiently large.

Condition (1) yields the same result. At the outset of the game ($t = 0$), the players' minmax payoffs are $M_1(0) = p(1-d)/(1-\delta)$ and $M_2(0) = (1-p)(1-d)/(1-\delta)$ which the states get if they fight. State *1*'s minmax payoff rises to $M_1(1) = (p+\Delta)(1-d)/(1-\delta)$ when its probability of prevailing rises to $p+\Delta$. Substituting these into (1) and letting the discount factor go to one give $\Delta(1-d) > d$. Thus, the mechanism formalized in

[29] To simplify the exposition, the distribution of power is assumed to shift deterministically in condition (1). This is not the case in Acemoglu and Robinson (2000, 2001), Fearon (2004), and de Figueiredo (2002) as well as in the general condition described in Powell (2004b) where the the actors are unsure whether the distribution of power will shift (although complete-information assumption means that the probabilities of future shifts are common knowledge). See Powell (2004b, 236) for a generalization of (1) which allows for stochastic shifts in the distribution of power.

the inefficiency condition explains why bargaining breaks down in fighting in Fearon's (as well as Powell's (1999, 128-32)) model of preventive war.

Inefficiency condition (1) also helps explain the commitment problem posed by first-strike or offensive advantages. Fearon (1995, 402-4) shows that what he calls the "de facto" bargaining range disappears if first-strike or offensive advantages are large enough. Suppose that *1* prevails with probability $p + f$ if it attacks and $p - f$ if it is attacked. Then the difference between these probabilities, $2f$, measures the size of the first-strike or offensive advantage. Taking these advantages into account, *1* prefers living with a territorial division x to attacking if $x \geq (p + f)(1 - d)$; *2* prefers x to attacking if $1 - x \geq (1 - p + f)(1 - d)$; and the de facto bargaining range are the x which satisfy these two conditions. This interval is empty and there are no divisions which both states simultaneously prefer to fighting whenever $(p-f)(1-d)+d < (p+f)(1-d)$ or, equivalently, $2f(1 - d) > d$.

Now consider more precisely *how* first-strike advantages undermine potential agreements. A key way is by creating shifts in the distribution of power. When a state decides to bargain rather than attack, it is also deciding not to exploit the advantages to striking first. This decision effectively shifts the distribution of power in the adversary's favor by giving it the opportunity to exploit the advantage to striking first, and this shift can lead to war through the mechanism formalized in condition (1).

The game in Figure 2 helps illustrates this. State *1* begins by deciding whether to attack or bargain by proposing a settlement. If *1* does make an offer, *2* can either accept or reject. If *2* accepts the game ends with the agreed division. If *2* rejects, it has to decide whether to fight or continue bargaining with *1*, and so on.

In order for *1* to be willing to make a proposal x that *2* might be willing to accept, *1*'s payoff to living with the agreement must be at least as large as what it could get by fighting: $(1 - x)/(1 - \delta) \geq (p + f)(1 - d)/(1 - \delta)$. And, *2* would only agree to an offer that gave it at least as much as it could get by rejecting it and then fighting: $x/(1-\delta) \geq 1-q+\delta(1-p+f)(1-d)/(1-\delta)$ where q is the status quo division. No such offers exist (again in the limit) and the bargaining is sure to break down in war whenever

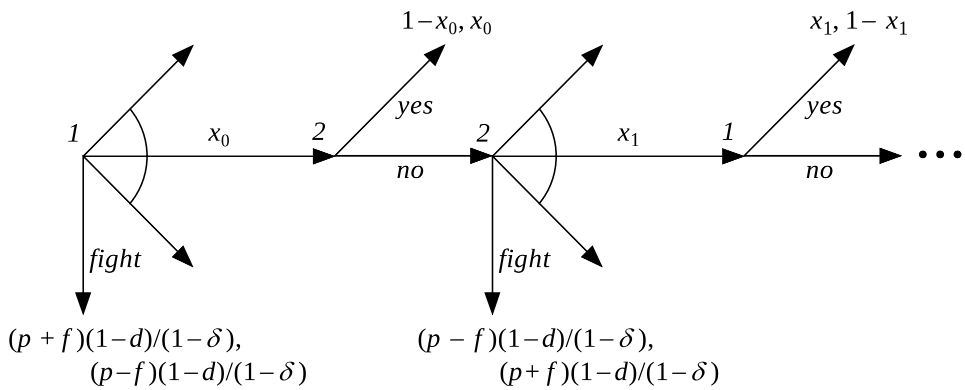

Figure 2: Shifting Power and First-Strike Advantages.

the de facto bargaining range is empty, i.e., when $2f(1-d) > d$.

This is just what condition (1) says. The states' minmax payoffs at time t when *2* is choosing between attacking and bargaining are their payoffs to fighting: $M_1(t) = (p-f)(1-d)/(1-\delta)$ and $M_2(t) = (1-p+f)(1-d)/(1-\delta)$. If *2* decides not to attack, the distribution of power shifts in favor of *1* whose minmax payoff rises to $M_1(t+1) = (p+f)(1-d)/(1-\delta)$. Condition (1) then becomes

$$\frac{\delta(p+f)(1-d)}{1-\delta} - \frac{(p-f)(1-d)}{1-\delta} > \frac{1}{1-\delta} - \left[\frac{(p-f)(1-d)}{1-\delta} + \frac{(1-p+f)(1-d)}{1-\delta}\right]$$

or, more simply, $[(1+\delta)f - (1-\delta)p](1-d) > d$. This relation is sure to hold if the states are sufficiently patient and if $2f(1-d) > d$. Thus first strike or offensive advantages can lead to war through large shifts in the distribution of power

A third kind of commitment problem can arise when states are bargaining about things that are themselves sources of military power, e.g., Czechoslovakia during the Munich Crisis or the Golan Heights (Fearon 1995, 408-09). Making a concession today weakens one's bargaining position tomorrow and necessitates additional concessions. Thus, a single concession may trigger a succession of subsequent concessions. Intuitively, a state might find itself in a situation in which it was willing to make a limited number of concessions but only if its adversary could commit to not exploiting its enhanced bargaining position to extract still more concessions. The inability to commit in these circumstances would lead to war.

Fearon (1996) shows that this intuition is not completely correct and that the commitment problem is more subtle. Suppose states *1* and *2* are bargaining over territory as in the examples above. In each round t, *1* can propose a territorial division $x_t \in [0,1]$ which *2* can accept or resist by going to war. If *2* accepts, x_t becomes the new territorial status quo, *1* and *2* respectively receive payoffs x_t and $1 - x_t$ in that period, and play moves on to the next round with *1*'s making another proposal.[30] If *2* decides to fight at time t, the probability that *1* prevails depends on the territory it controlled at time $t-1$. More specifically, *1* wins with probability $p(x_{t-1})$ where $p(x)$ is continuous, non-decreasing, and $p(0) = 0$ and $p(1) = 1$. Fighting also imposes costs c_1 and c_2 on the states. Consequently, *1*'s payoff to fighting at t is $p(x_{t-1})[\sum_{j=0}^{\infty} \delta^j(1) - c_1] + (1 - p(x_{t-1}))\,[\sum_{j=0}^{\infty} \delta^j(0) - c_1] = p(x_{t-1})/(1-\delta) - c_1$. State *2*'s payoff is defined analogously.

Fearon establishes the surprising result that the states never fight in the unique subgame perfect equilibrium as long as p is continuous. Rather *1* makes a series of proposals that always leave *2* just indifferent between fighting and acquiescing to the current proposal. More specifically, *1*'s offer at time t leaves *2* indifferent between fighting or accepting x_t and moving on to the next round where *1*'s offer will once again leave *2* indifferent between fighting and accepting.

To specify x_t more precisely, note that *2*'s payoff to fighting when *1* proposes x_t is $(1 - p(x_{t-1}))/(1-\delta) - c_2$. If *2* accepts x_t, it obtains $1 - x_t$ in round t and the states move on to round $t+1$ where *1*'s proposal x_{t+1} will leave *2* indifferent between fighting and continuing on. Hence, x_t satisfies

$$\frac{1 - p(x_{t-1})}{1-\delta} - c_2 = 1 - x_t + \delta\left(\frac{1 - p(x_t)}{1-\delta} - c_2\right) \tag{2}$$

where the expression in parentheses on the right is *2*'s payoff to fighting or, equivalently,

[30] The states are assumed to be risk neutral here in order to focus on the inefficiency due to fighting. Fearon's analysis allows the states to be risk averse as well as risk neutral. But risk aversion means that any territorial allocation that varies over time will be inefficient even if the states avoid fighting.

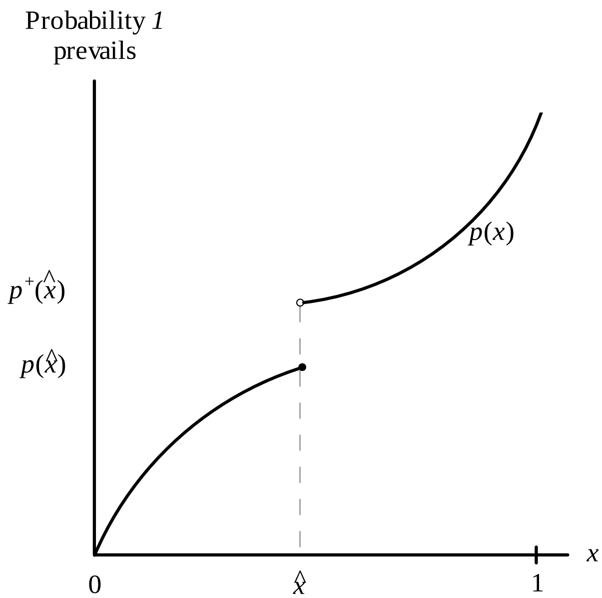

Figure 3: A Discontinuous Probability of Prevailing.

to accepting x_{t+1} and moving on. Equation (2) recursively defines a series of equilibrium demands $x_0^*, x_1^*, x_3^*, ...$[31]

That bargaining does not break down in inefficient fighting turns out to be crucially dependent on the continuity of p, i.e., on the fact that small changes in x only lead to small changes in p. Suppose, instead, that the probability that *1* prevails jumps discontinuously at $\widehat{x}$ as illustrated in Figure 3. Substantively, $\widehat{x}$ might be a strategically important geographic feature like a mountain pass, ridge, or river the control of which gives a state a military advantage. Formally, $p(\widehat{x})$ is strictly less than $p^+(\widehat{x})$ which is the limit of $p(x)$ as x approaches $\widehat{x}$ from the right. Then, bargaining breaks down in war if *1* is dissatisfied at $\widehat{x}$, the equilibrium sequence of offers includes $\widehat{x}$, and the discount factor is close enough to one.

To see why, suppose that *1* is dissatisfied at time t and that the distribution of territory is $\widehat{x}$. This distribution implies that *1*'s probability of prevailing if the states fight in the

[31] Fearon also shows that the absense of fighting is quite robust and does not depend on the simple take-it-or-leave-it bargaining protocal described here.

current round is $p(\widehat{x})$. Because *1* is dissatisfied at $\widehat{x}$, *2* must be willing to make some concession if the states are to avoid fighting. That is, *2* must agree to some $x_t > \widehat{x}$ in the current round. But *1* will then exploit its stronger bargaining position in the next period by making a demand that leaves *2* indifferent between accepting that offer and fighting when its probability of prevailing will have dropped from $1 - p(\widehat{x})$ to $1 - p(x_t)$. Consequently, *2* prefers fighting to agreeing to x_t if:

$$\frac{1 - p(\widehat{x})}{1 - \delta} - c_2 > 1 - x_t + \delta \left(\frac{1 - p(x_t)}{1 - \delta} - c_2 \right) \tag{3}$$

or, equivalently, if $\delta p(x_t) - p(\widehat{x}) > (1 - \delta)^2 c_2 - (1 - \delta) x_t$.

The discontinuity of p at $\widehat{x}$ ensures that this inequality holds if the discount factor is close enough to one. That is, the previous inequality goes to $p^+(\widehat{x}) - p(\widehat{x}) > 0$ as δ goes to one. Thus, both states prefer fighting at $\widehat{x}$ even though there are Pareto superior efficient divisions of the flow of benefits.

Inefficiency condition (1) once again accounts for this breakdown and in so doing helps provide some intuition for the effects of this discontinuity. According to this condition, state *2* prefers to fight at time t rather than accept x_t if accepting this offer would lead to an increase in *1*'s power (measured in terms of minmax payoffs) larger than the bargaining surplus. In symbols, there will be fighting if:

$$\delta \left(\frac{p(x_t)}{1 - \delta} - c_1 \right) - \left(\frac{p(\widehat{x})}{1 - \delta} - c_1 \right) > \frac{1}{1 - \delta} - \left(\frac{p(\widehat{x})}{1 - \delta} - c_1 + \frac{1 - p(\widehat{x})}{1 - \delta} - c_2 \right).$$

This reduces to $\delta p(x_{t+1}) - p(\widehat{x}) > (1 - \delta) c_2$ which goes to $p(x_{t+1}) - p(\widehat{x}) > 0$ in the limit. The discontinuity at $\widehat{x}$, therefore, ensures that the inefficiency condition holds if the states are sufficiently patient. Large, rapid shifts in the distribution of power again lead to costly fighting.[32]

In sum, the three seemingly different kinds of commitment problems share a fundamental similarity. In each of them, large shifts in the distribution of power can lead to bargaining breakdowns and war. These shifts may at some at some point confront a temporarily weak bargainer with the choice between fighting or suffering a large, adverse shift in the distribution of power if it continues to bargain. In the case of preventive war, this shift results from underlying changes in the states' military capabilities due, for example, to differential rates of economic growth or political development. In the case of preemption, a decision to continue bargaining may mean foregoing the advantages of striking first or being on the offensive. And, finally, when the distribution of power depends (discontinuously) on previous agreements, small concessions may bring dramatic changes in the distribution of power.

In order to induce an adversary not to fight in the face of these adverse shifts, a temporarily weak state must offer its adversary at least as much as it could get by fighting. And, the temporarily weak state would rather do this than fight because fighting is costly. But buying its adversary off may require the weak state to make a series of concessions that stretch across several periods during which the distribution of power will shift in its favor. If the once-weak state becomes sufficiently strong, it will renege on the remaining concessions. This prospect effectively limits the amount the temporarily-weak state can credibly promise to concede to its adversary. If this is less than the adversary can obtain

[32] Note that a discountinuous jump of any size can lead to figthing in Fearon's formulation. This result, if robust, would be substantively signficant as the prevalence of natural barriers like rivers and moutain ranges make discontinutity rathen than continuity an emprically more plausible assumption. This result, however, depends on the fact that fighting in Fearon's setup effectively becomes costless has the discount factor goes to zero. That is, the cost of fighting as a fraction of the total flow of benefits goes to zero as δ goes to one: $\lim_{\delta \to 1}(c_1 + c_2)/[1/(1-\delta)] = 0$. If we make the perhaps more natural assumption that fighting destroys a fraction d of the benefits as in the examples above, then fighting remains costly even as the discount factor goes to one. The appendix shows that the inefficiency condition still explains when fighting occurs in this modified model and that a relatively large jump of at least $d/(1-d)$ is needed to ensure fighting.

by fighting, the strong state will attack before the distribution of power shifts against it. *Shifting Power between Domestic Factions:* An analogous mechanism may operate at the domestic level. Here rapid shifts in the distribution of power between domestic factions may lead to international conflict if these factions are unable to commit themselves to divisions of the "domestic pie."[33] The basic idea is that if fighting and winning increases the probability of remaining in power, then the faction in power may choose to fight rather than agree to a settlement. In effect, the faction-in-power prefers the larger share of the smaller pie that fighting brings to the smaller share of the larger pie that it expects to get through negotiation.

To sketch a simple formal model highlighting this kind of commitment problem, suppose that the status quo is q and that the probability that state *1* prevails is p. As before, fighting destroys a fraction d of the resources, so *1*'s payoff to fighting is $p(1-d)+(1-p)0 = p(1-d)$ and *2*'s is $(1-p)(1-d)$. Hence both states prefer the territorial division x to war as long as $x \geq p(1-d)$ and $1-x \geq (1-p)(1-d)$ or, equivalently, as long as x is in the interval $p(1-d) \leq x \leq p(1-d)+d$. If q is in this interval, both states, when taken to be unitary actors, prefer the status quo to fighting.

Suppose, however, that state *1* is not a unitary actor. Rather *1* is composed of two factions, α and β. Faction α is currently in power and decides whether to fight and how to divide the state's resources between the two factions. To simplify matters, assume that the faction in power must give the out-of-power faction a share of at least $\lambda < \frac{1}{2}$ of the state's resources. We can think of this as the minimum necessary to buy off the out-of-power faction and dissuade it from launching a civil war or coup. (See Acemoglu and Robinson (2000, 2001, 2004) and Fearon (2004) for formulations along these lines.) Finally, let the probability that α retains power be r if there is no war and r' if there is a war and state *1* prevails. (If *1* is eliminated, both factions receive zero.)[34]

Faction α's payoff to accepting x is $(1-\lambda)x$ if α remains in power and λx if it loses

[33] This, of course, turns the anarcy-versus-hierarchy distinction between international and domestic politics on its head. For a discussion of this distinction, see Waltz (1979).

[34] On the effects of war on the fates of leaders, see Chiozza and Goemans (2004) and Goemans (2000, 53-71).

power. Agreeing to x therefore brings α an expected payoff of $r(1-\lambda)x + (1-r)\lambda x$. If by contrast α fights, its payoff if state *1* prevails and α remains in power is $(1-\lambda)(1-d)$ and $\lambda(1-d)$ if it loses power. Neither faction gets anything if state *2* prevails. This gives α an expected payoff to fighting of $p[r'(1-\lambda)(1-d) + (1-r')\lambda(1-d)]$.

Thus, both α and state *2* prefer x to war only if $p(1-d)[r'(1-\lambda) + (1-r')\lambda]/[r(1-\lambda) + (1-r)\lambda] \leq x \leq p(1-d) + d$. No such allocations exist if this bargaining range is empty, i.e., if $d[r(1-\lambda) + (1-r)\lambda] < p(1-d)(r'-r)(1-2\lambda)$. The expression on the left of the inequality is always positive, so this condition can only hold if fighting rather than settling increases α's chances of holding on to power (i.e., if $r' - r > 0$). When it does, this condition is more likely to hold the more likely state *1* is to prevail (the higher p), the lower the cost of fighting (smaller d), and the more valuable having control of the state is (the smaller λ).

Figure 4 models this situation as a game. State *2* begins by attacking or making an offer to state *1* which the faction in power, α, can accept or reject by fighting. If α accepts, it retains power with probability r. Thereafter the faction in power can try to buy off the out-of-power faction who can lock in a share λ of the domestic pie. If α fights, state *1* is eliminated with probability $1-p$ and both factions receive zero. If *1* prevails, α retains power with probability r' and the faction in power once again has the chance to buy off the out-of-power faction.

Strictly speaking, inefficiency condition (1) does not apply to this game because there is only one period, and there are more than two players. But the condition can be applied roughly by noting that there are only two players in the subgame starting with α's decision to fight or to accept an offer x and by taking the discount factor to be one. If $x > p(1-d)$, accepting increases the size of the domestic pie to be divided between α and β but also leads to a decline in α's power and an increase in β's. Condition 1 says that the bargaining will breakdown in inefficient fighting if the increase in β's power measured in terms of its minmax payoffs is larger than the bargaining surplus.

Faction β's minmax payoff if α accepts x is the probability that it comes to power times the payoff to being in power, $(1-r)(1-\lambda)x$, plus the payoff if α remains in power

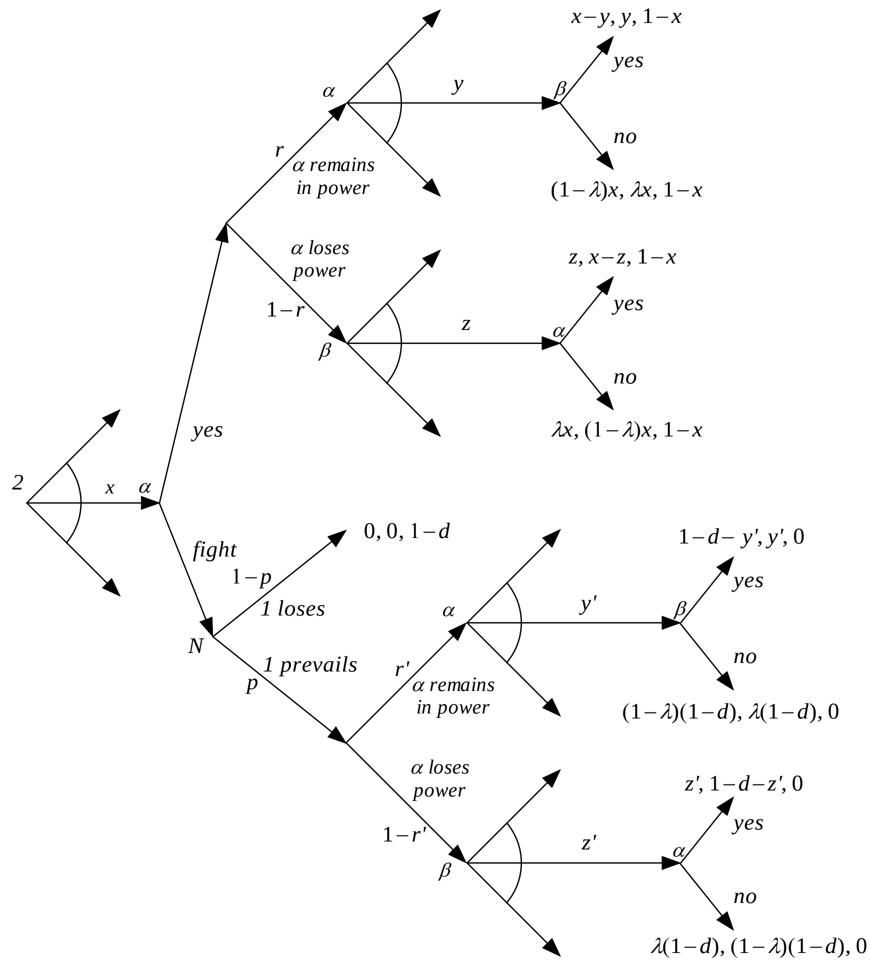

Figure 4: Shifting Power between Domestic Factions.

weighted by the probability that α retains control, $r\lambda x$. Similarly, β's minmax payoff if α fights is $p[(1-r')(1-\lambda)(1-d)+r'\lambda(1-d)]$. And, the bargaining surplus is $x-p(1-d)$. Substituting these payoffs into (1) and assuming $\delta = 1$ give

$$[(1-r')(1-\lambda)x + r'\lambda x] - p[(1-r')(1-\lambda)(1-d) + r'\lambda(1-d)] > x - p(1-d).$$

Simplifying yields $p(1-d)[r'(1-\lambda)+(1-r')\lambda]/[r(1-\lambda)+(1-r)\lambda] > x$. But, as shown above, α will only accept x if it is at least as large as the expression on the left of the previous inequality. Hence, α prefers fighting to accepting x whenever the inefficiency condition holds.

Once again, shifts in the distribution of power – this time at the domestic level – lead to bargaining breakdowns. Although the inefficiency condition can only be applied roughly, the fundamental idea underlying it helps explain the inefficient fighting. If $x > p(1-d)$, the domestic pie to be divided if α accepts is greater than if α fights. However, α's accepting leads to an adverse shift in the distribution of domestic power in that α's chances of remaining in power drop from r' to r. Because the pie to be divided is greater if α accepts, both factions would be better off if β could credibly promise to let α have as much as it could expect to get by fighting. But absent the ability to commit to divisions of the domestic pie, β cannot make this promise credible and α takes the country to war.[35]

This "domestic" commitment problem is closely related to Besley and Coate's (1998) analysis of political inefficiency. They identify three types of commitment problem which may prevent elected leaders from undertaking efficient investments in a representative democracy in which leaders cannot commit to following through on their election platforms.[36] First, a leader may not make efficient investments, i.e., investments which increase the present value of the flow of domestic benefits, if doing so adversely affects her probability of being re-elected or, more generally, of retaining power. Second, even if a leader's investment decision has no effect on the probability that one faction or the other will hold power, a leader may still act inefficiently if her investment decision affects the parties' future policy preferences. The party in power, for example, might run inefficiently high levels of debt in order to make its political opposition less willing to spend (on programs the party currently in power dislikes) should the opposition come to

[35] To see that both factions are better off, observe that difference between β's payoff to fighting, $[r'\lambda + (1-r')(1-\lambda)]$, and β's payoff to giving α its certainty equivalent of fighting, $x - [r'(1-\lambda) + (1-r')\lambda]p(1-d)$, is positive whenever $x > p(1-d)$.

[36] Persson and Tabellini (2000, 10-13) draw a useful distinction between models of pre- and post-election politics. In the former, parties or candidates are assumed to be committed to following through on their campaign positions. The median voter theorem is an example of this kind of model. In the latter, candidates cannot commit to their campaign pledges.

power.[37] Finally, a leader may face what is essentially the standard hold-up problem in economics.[38]

Although Besley and Coate focus on democratic states and economic investments, the commitment problems at the center of their analysis extend to other types of inefficient actions, like war in the example above, and to non-democratic states.[39] Indeed, the fundamental source of inefficiency in the model above is the same as Besley and Coate's first source. Acting efficiently – whether by investing or by refraining from fighting – adversely affects the chances that the faction in power remains there.

The Cost of Preserving the Status Quo: Finally, we turn to a very different type of commitment problem. A striking feature of the all of the examples above, and most formal work on war, is that fighting is costly but arming and securing the means to deter an attack are not.[40] Suppose more reasonably that states have to decide how to allocate their limited resources between guns and butter. Arming now entails an opportunity cost of foregone consumption.[41]

In these circumstances a state might face the following dilemma. State *1* can deter *2* from attacking by devoting a significant share of its resources to the military in every period. Alternatively, *1* can attempt to eliminate *2* by attacking and, if successful, be able to consume the "peace dividend," i.e., the resources it would otherwise be spending on deterring *2*. If deterring *2* is very expensive relative to the cost of fighting, *1* may

---

[37] See Alesina and Tabellini (1990), Persson and Svensson (1989), and Persson and Tabellini (2000, 345-61) for examples of this type of commitment problem.

[38] In the standard hold-up problem, the cost of investing is less than the investor's expected return because there is some chance that someone else will decide how to allocate the gains from the investment. The act of investing, however, has no effect on these chances or on the preferences of those making the allocation decisions. See Salanie (1997) for an introduction to the hold-up problem.

[39] For example, the authoritarian elites in Robinson (2003) fail to undertake efficient investments because they make it easier for the opposition to depose them.

[40] Slantchev (2004b) is a recent exception.

[41] Interestingly, models of conflict developed by economists generally do include a resource trade off but not an explicit decision to fight or attack (e.g., Hershleifer 2001, Garfinkle and Skaperdas 2000) whereas those developed by political scientists typically do not include a resource trade off by do include an explict decision to attack (see Powell 2002 for a review).

prefer attacking.

President Eisenhower appears to have weighed this option in the context of launching a preventive war against the Soviet Union before it acquired a large nuclear force. Writing to Secretary of State Dulles in 1953, Eisenhower worried that the United States

> would have to be ready on an instantaneous basis, to inflict greater loss on the enemy than he could reasonably hope to inflict on us. This would be a deterrent – but if the cost to maintain this relative position should have to continue indefinitely, the cost would either drive us to war – or into some form of dictatorial government. In such circumstances, we would be forced to consider whether or not our duty to future generations did not require us to *initiate* war at the most propitious moment that we could designate.[42]

Note that Eisenhower apparently believed that the United States would be able to deter the Soviet Union. But the cost of doing so over a prolonged period would be so high that going to war might be preferable.

Powell's (1993, 1999) guns-versus-butter model can be used to illustrate this type of commitment problem. Suppose that in each period states *1* and *2* have to allocate resources r_1 and $r_2 = 1 - r_1$ between consumption and defense. If, for example, *1* spends m_1 on the military, then its payoff is $r_1 - m_1$ in that period. Taking $p(m_1, m_2)$ to be *1*'s probability of prevailing given allocations m_1 and m_2, *1*'s payoff to attacking is $A_1(m_1, m_2) = r_1 - m_1 + p(m_1, m_2)[\delta(1 - d)/(1 - \delta)]$. The difference $r_1 - m_1$ is *1*'s consumption during the current period during which the states are fighting. The last term is the expected payoff to fighting. With probability p, *1* eliminates *2*, takes control of *2*'s resources, and reallocates all of them to consumption. This gives *1* a per-period payoff of $1 - d$ where d is the fraction of resources destroyed by fighting. State *1* loses and receives a payoff of zero with probability $1 - p$.

Both states prefer living with the allocation (m_1, m_2) to optimally arming for war and attacking if $(r_1 - m_1)/(1 - \delta) \geq A_1(m_1^*, m_2)$ and $(r_2 - m_2)/(1 - \delta) \geq A_2(m_1, m_2^*)$ where m_j^* maximizes A_j.[43] Conversely, at least one state prefers fighting to living with the status quo (m_1, m_2) if $r_1 + r_2 - m_1 - m_2 < (1 - \delta)[A_1(m_1^*, m_2) - A_2(m_1, m_2^*)]$. Simplifying matters

[42] Quoted in Gaddis (1982, 149).

[43] Powell (1993) shows that these inequalities bind in a peaceful equilibrium and this pins down the equilibrium allocations.

by assuming the players are very patient (i.e., letting δ go to one), the previous inequality reduces to $d < m_1 + m_2 + (1-d)[p(m_1^*, m_2) - p(m_1, m_2^*)]$.

This relation formalizes the commitment problem. At least one state will be dissatisfied and prefer attacking if the cost of fighting, d, is less than the cost of preserving the status quo, $m_1 + m_2$, plus the cost of being on the defensive rather than offensive.[44] Even if these latter costs are negligible, at least one of the states will prefer war to peace whenever the cost of fighting is less that the burden of defending the status quo. Bargaining does not breakdown in war in this mechanism because of a large, rapid shift in the distribution of power but because deterring an attack on the status quo is too expensive.

Conclusion

There are two rationalist approaches to the inefficiency puzzle inherent in war. A purely informational problem exists when states fight solely because of asymmetric information. Were there complete information, there would be no fighting. By contrast, a pure commitment problem exists when states have complete information and still fight.

Most formal work has treated war as a purely informational problem, and this approach has yielded important theoretical and empirical results. But the implicit assumption tht states would not fight if there were complete informatin creates an analytic blind spot which leads to odd readings of some cases. Fighting often does not seem to result from some residual uncertainty about an adversary. Rather, war comes when a state becomes convinced it is facing an adversary it would rather fight than accommodate.

Uncertainty abounds in international politics and many situations are likely to combine significant informational and commitment problems. Indeed, some private information is likely to be present in cases which are fundamentally commitment problems. The prevalence of uncertainty presents emprical and theoretical challenges to studying commitment

[44] The difference $p(m_1^*, m_2) - p(m_1, m_2^*)$ measures the change in *1*'s probability of prevailing if it optimally rearms for war and attacks or its adversary does. This difference times the resources surviving a war, $1-d$, is the expected loss of giving an adversary the offensive advantage of optimally arming for war.

problems.[45]

One way of studying commitment problems theoretically is to isolate them from informational problems by investigating the inefficiency puzzle in the context of complete-information games. This complete-information approach abstracts away from informational issues and focuses directly on the strategic mechanism through which the inability to commit leads to costly fighting. The goal – hope – of this approach is that it will be possible to identify a handful of mechanisms which explain a significant number of cases.

The present analysis describes two mechanisms. In the first, large, rapid shifts in the distribution of power undermine peaceful settlements. In order to induce its adversary not to fight, a temporarily weak state must promise its adversary at least as much as it can get by fighting. But when the once-weak bargainer becomes stronger, it will exploit its better bargaining position and renege on its promise. In effect, the shifting distribution of power limits the amount that the weak bargainer can credibly promise to give its adversary. If this is less than what that state can get by fighting, there will be war. This mechanism can be sseen to be at work in each of Fearon's three commitment problems. A closely related mechanism operating at the domestic level may also cause war. Here a shifting distribution of power between domestic factions can lead to inefficient fighting if these factions cannot commit to divisions of the domestic pie.

Finally, a second mechanism emphasizes the cost of deterring an attack rather than a shifting distribution of power. Bargaining models of war often abstract away from resource-allocation issues. As a result, fighting is costly but procuring the means needed to fight is not. This makes it impossible to compare the cost of deterring an attack on the status quo with the cost of using force to try to eliminate the threat to the status quo. When these costs can be compared, a state may prefer fighting to living with the status quo if deterring an attack is very costly.

[45] Ausbel and Deneckere (1989) show that one can get folk-theorem-like results in bargaining games, i.e., almost any behavior is consistent with some equilibrium. This means that it will be very difficult to rule out informational accounts deductively when the actors have private information. Ultimately, one will have to judge which mechanisms seem to provide a more compelling account of a set of cases. These judgments will have to await a better theoretical understanding of commitment problems.

Appendix

The effects and interpretation of the role of a discontinuity in p in Fearon's (1996) analysis of bargaining over objects that influence future bargaining power depend on which state is dissatisfied and which has the bargaining power. To develop these points, observe that costs of fighting relative to the size of the benefits in Fearon's specification goes to zero as the discount factor goes to one. That is, $\lim_{\delta \to 1}(c_1 + c_2)/[1/(1-\delta)] = 0$. In effect, fighting becomes costless and the bargaining surplus vanishes as the discount factor goes to one. Because the surplus disappears, the states' relative bargaining power which affects who gets how much of the surplus is of no consequence. Suppose, however, that the costs of fighting relative to the total benefits do not go to zero. Assume more specifically that the costs of fighting are modeled in terms of the fraction of resources destroyed as in the other examples above.

Fearon's model allows for the possibility that either state is dissatisfied (which state is dissatisfied depends on the initial distribution of territory x_0), and he happens to consider the case in which the dissatisfied state also has all of the bargaining power (i.e., state *1* makes the take-it-or-leave-it offers and the $\{x_t\}$ are increasing. A discontinuous jump in p of any size can lead to fighting in these circumstances even if he cost of fighting does not go to zero.

Formally, *2* prefers fighting to agreeing to x_t if inequality (3) is rewritten as:

$$[1 - p(\widehat{x})]\left(\frac{1-d}{1-\delta}\right) > 1 - x_t + \delta[1 - p(x_t)]\left(\frac{1-d}{1-\delta}\right). \tag{4}$$

This reduces to $\delta p(x_t) - p(\widehat{x}) > (1-\delta)[d - x_t]$ which again holds as long as p increases discontinuously at $\widehat{x}$ and the discount factor is close enough to one.

Suppose, however, that *1* has all of the bargaining power but *2* is dissatisfied (i.e., $\{x_t\}$ is decreasing). Then the discontinuous change in p needed to trigger fighting is $d/(1-d)$. To establish this, simplify the analysis by assuming that p is continuous from the right instead of the left as above, i.e., $\lim_{x \downarrow \widehat{x}} p(x) = p(\widehat{x}) > \lim_{x \uparrow \widehat{x}} p(x) \equiv p^-(\widehat{x})$. To compare the equilibrium condition to the inefficiency condition in these circumstances, note that

1 prefers to fight rather than satisfy *2*'s incentive compatibility constraint (2) if:

$$p(\widehat{x})\frac{1-d}{1-\delta} > x_t + \delta\left[\frac{1}{1-\delta} - (1-p(x_t))\frac{1-d}{1-\delta}\right].$$

The left side of this relation is *1*'s payoff to fighting with a probability of prevailing $p(\widehat{x})$. The right side is *1*'s payoff if *2* accepts x_t and *1* then gets all of the surplus after giving *2* its certainty equivalent to fighting. Simplifying and taking the limit as δ goes to one shows that *1* prefers fighting to accommodating *2* at $\widehat{x}$ if $p(\widehat{x}) - p^{-}(\widehat{x}) > d/(1-d)$.

Thus, a discontinuous jump of at least $d/(1-d)$ is needed to ensure fighting when war destroys a fraction d of the flow of benefits. The intuition underlying the importance of the locus of bargaining power is that the more bargaining power a state has, the larger the share of the surplus it gets and the greater cushion it has against adverse shifts in the probability of prevailing or its minmax payoffs. When *1* has the bargaining power and faces adverse shifts (because it must make concessions to a dissatisfied *2*), it takes larger shifts in the distribution of power to trigger fighting.

Inefficiency condition (1) is a sufficient condition: it specifies conditions which when satisfied are sure to result in fighting. This condition when applied to the modified model shows that a jump of $d/(1-d)$ is needed to ensure fighting. To apply condition (1) to the case where *2* is dissatisfied, note that *2* grows stronger at $\widehat{x}$ because p drops. Condition (1) then says that bargaining breaks down if the increase in *2*'s minmax payoff is greater than the bargaining surplus:

$$\delta(1-p(x_t))\frac{1-d}{1-\delta} - (1-p(\widehat{x}))\frac{1-d}{1-\delta} > \frac{1}{1-\delta} - \left(p(\widehat{x})\frac{1-d}{1-\delta} + (1-p(\widehat{x}))\frac{1-d}{1-\delta}\right)$$

which becomes $p(\widehat{x}) - p^{-}(\widehat{x}) > d/(1-d)$ in the limit. Hence, the equilibrium condition and the inefficiency condition are the same.

References

Acemoglu, Daron and James A. Robinson. 2000. "Why Did the West Extend the Franchise," *Quarterly Journal of Economics* 115(November):1167-99.

Acemoglu, Daron and James Robinson. 2001. "A Theory of Political Transitions," *American Economic Review 91*(September): 938-63.

Acemoglu, Daron and James Robinson. 2004. *Economic Origins of Dictatorship and Democracy.* Manuscript, Department of Government, Harvard University.

Alesina, Alberto and Guido Tabellini. (1990). "A Positive Theory of Fiscal Deficits and Government Debt," *Review of Economic Studies*, 57(July): 403-414.

Aster, Sidney. 1973. *1939: The Making of the Second World War.* London: Andre Deutsch.

Ausubel, Lawrence and Raymond Deneckere. 1989. "Reputation in Bargaining and Durable Goods Monopoly." *Econometrica* 57(May):511-31.

Bennett, D. Scott and Alan Stam. 1996. "The Duration of Interstate Wars, 1816-1985," *American Political Science Review* 90(June):239-57.

Besley, Timothy and Stephen Coate. 1998. "Sources of Inefficiency in a Representative Democracy," *American Economic Review* 88(may):139-56.

Busch, Lutz-Alexander and Abhinay Muthoo. 2002. "Power and Inefficient Institutions," Manuscript, Department of Economics, University of Essex (www.essex.ac.uk/~muthoo/agenda.pdf).

de Figueiredo, Rui. (2002). "Electoral Competition, Political Uncertainty, and Policy Insulation," *American Political Science Review* 96(June):321-33.

Chiozza, Gaicomo and Hein Goemans. 2004. "International Conflict and the Tenure of Leaders," *American Journal of Political Science* 48(July):604-19.

DGFP. 1949-64. *Documents on German Foreign Policy, 1918-1945*, series D (1937-1945), vol. 1-14. Washington, D.C.: U.S. Government Printing Office.

Fearon, James D. 1995. "Rationalist Explanations for War," *International Organization* 39(Summer):379-414.

Fearon, James D. 1996. "Bargaining over Objects the Influence Future Bargaining Power." Unpublished manuscript, Department of Political Science, Stanford University (www.stanford.edu/~jfearon/papers/booifbp.pdf).

Fearon, James D. 1998. "Commitment Problems and the Spread of Ethnic Conflict." In David A. Lake and Donald Rothchild, eds., *The International Spread of Ethnic Conflict.* Princeton: Princeton University Press.

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

Filson, Darren and Suzanne Werner. 2002. "A Bargaining Model of War and Peace." *American Joural of Political Science* 46 (October): 819-38.

Gaddis, John. 1982. *Strategies of Containment.* New York: Oxford University Press.

Garfinkle, Michelle and Stergios Skaperdas. 2000. "Conflict Without Misperceptions or Incomplete Information," *Journal of Conflict Resolution* 44:(December): 793-808.

Goddard, Stacie. nd. "Uncommon Ground," Manuscript, Department of Politics, Prince-

ton University (www.Princeton.edu/~segoddard/apsaedit1.pdf).
Goemans, Hein. 2000. *War and Punishment.* Princeton: Princeton University Press.
Gul, Faruk and Hugo Sonnenschein. 1988. "On Delay in Bargaining with One-Sided Uncertainty," *Econometrica* 56(May):601-611.
Halder, Franz. 1988. *The Halder War Diary 1939-1942.* Charles Burdick and Hans-Adolf Jacobsen, eds. Novato, CA: Presidio Press.
Hassner, Ron. 2003. "To Have and To Hold," *Security Studies* 12(Summer): 1-33..
Hassner, Ron. 2004. "The Path to Indivisibility," Manuscript, Department of Political Science, UC Berkeley.
Hirshleifer, Jack. 2001. *The Dark Side of Force.* New York: Cambridge University Press.
Jervis, Robert. 1978. "Cooperation Under the Security Dilemma," *World Politics*, 30(January):167-214.
Kydd, Andrew. 2003. "Which Side Are You On?" *American Journal of Political Science* 47(October): 597-611.
Lukacs, John. 1976. *The Last European War.* London: Routledge, Kegan, Paul.
Martel, Gordon. 1999. *The Origins of the Second World War Reconsidered*, 2nd ed. New York: Routledge.
Parker, R.A.C. 1993. *Chamberlain and Appeasement.* New York: St. Martins.
Persson, Thorsten and Lars Svensson. 1989. "Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences," *Quarterly Journal of Economics* 104(May):325-45.
Persson, Torsten and Guido Tabellini. 2000. *Political Economics.* Cambridge: MIT Press.
Powell, Robert. 1993. "Guns, Butter, and Anarchy," *American Political Science Review* 87(March):115-32.
Powell, Robert. 1996a. "Bargaining in the Shadow of Power," *Games and Economic Behavior* 15(August):255-89.
Powell, Robert. 1996b. "Stability and the Distribution of Power," *World Politics* 48 (January): 239-67.
Powell, Robert. 1999. *In the Shadow of Power.* Princeton: Princeton University Press.
Powell, Robert. 2002. "Bargaining Theory and International Conflict," *Annual Review of Political Science* 5(June):1-30.
Powell, Robert. 2004a. "Bargaining and Learning While Fighting." *American Journal of Political Science* 48:(April):344-61.
Powell, Robert. 2004b. "The Inefficient Use of Power." *American Political Science Review* 98(May):231-41.
Rich, Norman. 1973. *Hitler's War Aims: Ideology, the Nazi State, and the Course of Expansion.* London: Andre Deutsch.
Rich, Norman. 1974. *Hitler's War Aims: The Establishment of the New Order.* London: Andre Deutsch.
Robinson, James. 2003. "When Is a State Predatory?" Manuscript, Department of Political Science, UC Berkeley.
Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model," *Econometrica* 50(January):97-110.

Salanie, Bernard. 1997. *The Economics of Contracts.* Cambridge: MIT Press.

Schelling, Thomas C. 1960. *The Strategy of Conflict.* Cambridge: Harvard University Press.

Schwarz, Michael and Konstantin Sonin. 2003. "Conflicts and Commitments." Manuscript, New Economic School, Moscow (www.nes.ru/~ksonin/wars62.pdf).

Slantchev, Branislav. 2003a. "The Power to Hurt," *American Political Science Review* 47(March): 123-35.

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

Slantchev, Branislav. 2004b. "How Initiators End heir Wars," *American Journal of Political Science* 48:(October): 813-29.

Slantchev, Branislav. 2004. "Military Coercion in Interstate Crises and the Price of Peace," Manuscript, Department of Political Science, UCSD. (Available at: ???).

Smith, Alastair and Alan Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* (forthcoming).

Taylor, A.J.P. 1962. *The Origins of the Second World War.* New York: Atheneum.

Taylor, Telford. 1979. *Munich.* New York: Doubleday.

Thorne, Christopher. 1968. *The Approach of War, 1938-1939.* London: Macmillan.

Toft, Monica. 2002/3. "Indivisible Territory, Geographic Concentration, and Ethnic War," *Security Studies* 12(Winter): 82-119.

Sontag, Raymond. 1971. *A Broken World, 1919-1939.* New York: Harper & Row.

Trevor-Roper, H.R. 1964. *Hitler's War Directives, 1939-45.* London: Sidgewick and Jackson.

Waddington, G.T. 1996. "*Hassgegner*: German Views of Great Britain in the Later 1930s," *History* 81(January):22-37.

Wagner, R. Harrison. 2000. "Bargaining and War," *American Journal of Political Science* 44(July):469-84.

Wagner, R. Harrison. 2004. *War and the State.* Manuscript, Department of Government, University of Texas, Austin (www.la.texas.edu/~hw/drafts/introirbook1col.pdf).

Waltz, Kenneth. 1979. *Theory of International Politics.* Reading, MA.: Addison-Wesley.

Watt, Donald Cameron. 1989. *How War Came.* New York: Pantheon.

Weinberg, Gerhard. 1970. *The Foreign Policy of Hitler's Germany, 1933-36.* Chicago: University of Chicago Press.

Weinberg, Gerhard. 1980. *The Foreign Policy of Hitler's Germany, 1937-39.* Chicago: University of Chicago Press.

Werner, Suzanne. 2000. "Deterring Intervention," *American Journal of Political Science* 44(October):720-32.

Wheatley, Ronald. 1958. *Operation Sea Lion.* Oxford: Oxford University Press.

Wittman, Donald. 1979. "How a War Ends," *Journal of Conflict Resolution* 23 (December): 743-63.