Military Threats

The Costs of Coercion and the Price of Peace

Branislav L. Slantchev

На дядовците ми, Георги и Петър.

Contents

Figures

Tables

Acknowledgements

This book has a long history. I started wrestling with the basic issues over five years ago and have worked on the book for three years now. I have lost track of the number of times I rewrote entire chapters from scratch.

First, and foremost, I would like to thank Bob Powell. As a participant in the book conference I organized in the summer of 2007 in La Jolla, he read the entire draft and provided copious comments on how to improve it. Beyond that, he has been instrumental in shaping my thinking about these issues and is responsible for the clear framing and articulation of many points in this book. Although I pride myself on the clarity of my thinking, I am still astonished how muddled my logic can be and how often I can miss the crucial point while focusing on secondary ones. Bob's comments have provided the necessary corrective and, more often than not, the cold shower for enthusiasm that outstrips the findings.

I cannot thank enough the participants in the book conference. David Lake, Ken Schultz, R. Harrison Wagner, Peter Gourevitch, Miles Kahler, and Kelly Kadera not only read the entire draft in all its messy ugliness but came to La Jolla to give me their comments, ideas on how to improve it, and advice about what to ditch. I made an extra effort to follow their careful suggestions and as a result the book is so much better than the one they read, it almost qualifies as a new one. Jim Fearon and Jim Morrow also read the 2007 draft and I am sure they will recognize the result of their suggestions in the book now. I must single out Jim Fearon for a special debt of gratitude because his work is responsible, I am sure inadvertently so he must be held blameless in this, for my career thus far.

My research assistants Jessica Weiss, Jihyeon Jeong, Hugh Dauffenbach, and Molly Bauer have been invaluable. I am surprised how much they managed to do with my vague demands and sometimes lackadaisical direction. Molly also read the entire draft and gave me numerous written comments. Graduate students who find the book easier to read should thank her.

I have also lost track of the candid discussions, heated arguments, and vigorous collaboration that have contributed to the final version of this project. Robert Walker is the designated victim of nearly all my flashes of inspiration which so often turn out to be instances of that muddled thinking I mentioned before. He has been patient to the point of getting qualified as a victim of mental abuse. I thank him for letting me wax keen about ideas before his probing questions made me realize how embarrassing it would have been to voice them in public. I have seriously benefitted from the generous input of Ahmer Tarar and Kris Ramsay with whom I have managed to butt heads on more than one occasion about these issues. My head hurts now but it was worth it! Colleagues that have also donated their time and thinking hats to puzzles that tormented me are well deserving of a heartfelt thanks. Among

these are Hein Goemans, Kristian Gleditsch, Erik Gartzke, Al Stam, Robert Rauchhaus, Andy Kydd, Shuhei Kurizaki, Matt Baum, Ron Hassner, Jahoon Kim, Joel Watson, Songying Fang, David Mares, and Damon Colletta. I would also like to thank Paul Huth who readily shared his data on deterrence and also sent me the full case summaries which made it so much easier to work with the historical cases.

As my advisors in graduate school at Rochester, Randy Stone and Curt Signorino must share responsibility for my academic career. I have no idea what kind of researcher I would have become without them and I really do not care to speculate. Suffice it to say that if I had to choice to redo my graduate studies now that I know more about the discipline and the profession, I would march right back to Rochester and pick the same committee. This experience would not be complete without Randy Calvert and John Duggan, two mentors whose mark on my thinking will remain indelible.

My colleagues at UCSD have given me the supporting environment that junior faculty can only dream of. The University gave me a generous course release and a full sabbatical which allowed me to spend an entire year lurking in coffee shops to work on this book. Portions of this research were presented in seminars at Stanford, Berkeley, Harvard, NYU, Rochester, Columbia, Dartmouth, and the University of California–Irvine. I am grateful to the participants for their input. The colloquium sponsored by the Institute on Global Conflict and Cooperation at the University of California enabled me to share preliminary results with colleagues across disciplines in our system. I also gratefully acknowledge financial support from the Committee on Research Awards at the University of California and the National Science Foundation (grant SES-0518222).

My friends Jeff Lax and Katja Favretto helped me get through the inevitable rough patches and have made my life as an overworked academic battling occasional bouts of frustration a lot more enjoyable than it would have been. How much so? Let's just say that for a while this book circulated under the title *The Complete Idiot's Guide to Military Threats*, which unfortunately the publisher would not permit me to use. Copyright issues no doubt.

Christina Schneider also read the entire manuscript, some chapters more than once in their various incarnations. Her insistence on streamlining the presentation, clarifying the arguments, and dropping the cute digressions—something for which I exhibit unhealthy proclivities—has resulted in one tight book. On a personal level... well, it's too personal for print. She knows.

At this point, it occurs to me that more people may have read the manuscript than are likely to read the finished product! None of them is responsible for any remaining errors. I suppose these could be my fault.

I would also like to thank my cats but I don't have any.

Part I

Coercion and Credibility

1

Introduction

Preparation for war does not make war inevitable. On the contrary, prudent preparation for war, accompanied by a wise policy, provides a guarantee that war will not break out except for the gravest of reasons.

Count Sergei I. Witte

Military power is what gets one's voice heard in world affairs. Creating and maintaining armed forces is among the costliest undertakings for a nation short of their employment in hostilities. Even a casual glance at history reveals that whatever their defensive role is, armed forces are often used to menace others. More often than not, they are used indirectly, as an implicit or explicit presence in the background of negotiations, rather than directly in fighting.[1] States frequently find themselves on the opposite sides of disputes and in their attempts to wrangle concessions out of each other, they sometimes resort to military threats. The threat to use force can be verbal without any overt preparation to do so, or physical with all the measures—putting forces on alert, recalling reservists, mobilizing, dispatching the navy, deploying troops—required for its actual use. These physical measures, which I collectively refer to as *military moves*, do not have to be accompanied by an explicit warning. They are so menacing that the threat of hostile intent is implicit in their use. Sometimes these moves are nothing but necessary steps on the road to war. But more often, they are intended as a warning that war may come unless the adversary accedes to one's demands. War, with its enormous costs, pain, and risks, is not something to be contemplated lightly. But there are things worse than war and common sense dictates what history reveals: even state leaders who are averse to war can deliberately risk it to convince others to bend to their wishes.

It is the function of military moves as instruments to induce desired behavior in others rather than their proper application in the deadly arts of destruction that interests me. This is a book on military coercion. It is a study of how military threats can be employed in the pursuit of political goals. For a military threat to succeed as a coercive device, it has to accomplish two objectives: (a) it has to persuade the opponent that one is sufficiently likely to resort to violence if one's demands are not met, and (b) it has to render fighting sufficiently unpleasant for the opponent relative to the concessions demanded. What makes military threats effective? Why might they fail even if they are believable? Why would an actor forego the possibilities of militarized diplomacy and opt for war instead? How are military threats different from other instruments of coercion? These are all questions I address in this

[1] Goldhamer (1979, 9); Karsten et al. (1984, 3–5); Naroll et al. (1974, 1–2); Schelling (1966); Blechman and Kaplan (1978); Young (1968).

book. Although my interest is primarily theoretical, I will draw upon numerous historical cases to motivate the research and illustrate the logic of its findings.

The fundamental result is that military threats can be very effective tools of coercion. They can establish intent to wage war and can communicate that fact to the opponent in a way that he will believe it. Military threats can even reduce the likelihood that the confrontation will end in war relative to other coercive instruments. Unfortunately, these threats also tend to be expensive, especially if their intent is to coerce the opponent rather than wage war. Whereas this may discourage their use and thereby reduce the chances of a militarized dispute, it may also convince leaders that it is easier to settle the matter by force instead of trying to coerce the opponent with threats. This makes war more likely and underscores the need to distinguish between military moves that are a prelude to war and those that are designed to influence the opponent's behavior. These, as Count Witte observed, are not quite the same even though may take similar outward appearances.[2]

The findings have implications for international relations theory and policy. On the theoretical side, the results contradict a long tradition of arguing that nations with more powerful militaries tend to get their way more often than others but at the cost of having to risk war more often too. This may be so for non-military instruments but not for military threats. Through the judicious use of military threats, powerful states can secure better peaceful outcomes and lower the risk of war. Their task can be made more difficult if they misperceive the magnitude of the stakes for their opponent. Their overconfidence may prove to be their undoing if they fail to muster the resources necessary to coerce a determined adversary. However, even if they are pessimistic, their actions may make war more likely because they mistakenly believe that it would take too much effort to coerce the opponent and opt for war instead. In fact, the finding that the overall danger of war is mediated through the distribution of interests can help explain why attempts to link it directly to the distribution of power have generally failed. The likelihood of war depends on the extent to which one is prepared to use military threats to deter challenges to peace and compel concessions without fighting. The price of peace may be military establishments that are both costly and unused. These armed forces are not useless for their employment is indirect but nevertheless crucial.

I am more reluctant to draw conclusions with policy implications because no one is more acutely aware of the shortcomings of my theories than I am. However, even I cannot resist a couple of observations. Despite the attractiveness of the military instrument as a tool for coercion, one cannot have militarized coercion on the cheap. Gunboat diplomacy is unlikely to work unless it represents firepower that can make a difference in an actual engagement. In other words, military threats cannot be token in character if they are to succeed. They are not a cheap way for the powerful to throw their weight around. In fact, wealthier and more powerful nations may have to engage in relatively more aggressive behavior in order to make their threats stick. They may have to mobilize overkill capability compared to the issues at stake. Shooting flies with an elephant gun may well be the prudent thing for them to do.

The argument in this book depends on a series of theoretical models which all share the same basic assumptions. In this, they all stand or fall together, so it may be worthwhile to

[2] Ironically, Witte made this remark about the preparations right before the outbreak of the Russo-Japanese War of 1904–05 (Harcave, 1990, 308–9).

provide some justification for the choices I have made. I assume that a conflict of interest exists between two unitary rational actors who confront each other once to resolve it. A number of important assumptions are already buried in this simple statement.

I assume that the two actors are unitary and rational; that is, they behave as individuals with well-defined preferences. By "well-defined" preferences I mean that the actors can rank-order all the various possible outcomes of their interaction in a logically coherent way. More importantly, they can rank-order risky alternatives. For instance, suppose an actor is confronted with an ultimatum from his opponent and, for simplicity, suppose he has three options at his disposal: agree to the terms, launch a preemptive attack, or let the ultimatum expire to see if his opponent will attack. To decide on the best course of action, the actor must evaluate the likely consequences of the various options at his disposal. Capitulation to the opponent's demands avoids war but (presumably) imposes political and economic costs by forcing the actor to agree to unpalatable conditions. Launching a preemptive strike means going to war, with all the attendant risks and costs. There is no guarantee of victory but there is a chance to avoid the bad outcome. The third option is to let the ultimatum deadline lapse in the hope that the opponent will not attack. Unlike the outright capitulation, there may be a chance to avoid the bad outcome but at the risk of a war. Unlike launching a preemptive strike, there is a chance to avoid war but at the risk of foregoing whatever advantages such an attack would confer.

Each of these options has its own costs and benefits and each involves some trade-offs. We say that preferences are rational when they are logically consistent. For instance, it cannot be the case that the actor expresses a preference for adopting a wait-and-see stance to preemptive attack and preemptive attack to outright capitulation and then also be the case that he expresses a preference for outright capitulation over adopting a wait-and-see stance.[3] Throughout this book, we shall remain agnostic as to where these fundamental preferences come from. We shall take them as given and fixed.

This last assumption is actually less demanding than one might suppose. For instance, it does not imply that actors will not change their minds about what they want to do in a given situation when they obtain new information. To see that, suppose that we begin with the above rank-ordering which implies that waiting is the most preferred course of action. Suppose then that the actor receives information that if he lets the ultimatum deadline lapse, his opponent is almost certain to attack. As a result, he launches a preemptive strike. One might think that this indicates that the actor's preferences have changed, which would imply that taking them as fixed would be a serious problem. However, this is not so: all it means is that our original specification of preferences is not quite right for it misses an important bit that determines the trade-off between waiting and preempting. In this instance, the actor prefers to wait if there is some reasonable chance that his opponent will not attack (because this would avoid war) but prefers to attack himself if war seems unavoidable. His estimate of the probability that his opponent will attack if concessions are not forthcoming is part of the expected consequences of the actions and as such must be included in the preferences.

[3] In technical terms, preferences must be complete (i.e., actors must be able to consider all possible outcomes) and transitive (i.e., they do not admit logical contradictions like the one in the text). There are some more subtle requirements when it comes to rank-ordering risky choices. See von Neumann and Morgenstern (1947) for the classic treatment.

The correct way to specify the preferences, then, would be to give a full account of the contingencies.

One possibility is that the actor prefers to wait if there is at least a 50% chance that the opponent will not attack and prefers to preempt otherwise, with both of these being preferable to capitulation. Now, the reception of new information that causes that actor to revise downward his estimate of the probability of war if he waits may well cause him to choose to preempt even though he would have chosen to wait in the absence of this information. Loosely speaking, his preference for preemption over waiting has changed. Strictly speaking, this is not the case: the *choice of action* changed because his estimate of its likely *outcome* changed because of the new information. But notice that the preference ordering is:

- If the probability that the opponent will attack when the ultimatum deadline expires without response is more than 50%, preemption is preferable to waiting and waiting is preferable to capitulation;
- If the probability that the opponent will attack when the ultimatum deadline expires without response is less than 50%, waiting is preferable to preemption and preemption is preferable to capitulation.

The actor's initial estimate was that there was less than 50% chance of an attack if he waited, which meant that he would choose to wait. However, in light of the new information received, he has revised his estimate of that probability upward, and now chooses to preempt. Observe that his fundamental preferences have remained fixed even though his choice of action has changed. In other words, what the actor learns during the crisis can affect his behavior even though his fundamental preferences stay the same. In fact, this entire book is about how actors can alter the behavior of their opponent by manipulating information and the strategic environment.

In addition to having well-defined preferences, rational actors must pursue their goals to the best of their ability given the information they have and the constraints they must operate under. It is often supposed that rationality requires full information and the evaluation of all possible alternatives. That is not the case. As we shall see, the dynamics of military threats are highly contingent on uncertainty, both about the opponent's intentions and the outcomes of risky choices. It is true that to evaluate the best course of action, the actors will have to compare all the alternatives available to them but as analysts we have already simplified the world by limiting the actors' choices. It is very likely that in reality the actors are similarly constrained to just a handful of options and they do not consider all possible options. In that sense, the model's limitations are perhaps more realistic than one might suppose. It is a fascinating puzzle to see how actors frame the problem and decide what actions are simply not to be considered. In this book, we abstract away from that and assume that they have arrived at a particular frame of reference. Whether this simplification is distorting or not depends on how many relevant choices it leaves out, something that we would have to investigate in future work.[4]

[4] Karsten et al. (1984, 8–10) discuss the shift of cost-benefit calculations during crisis and conclude that "situations that are characterized by [. . .] lags in identifying the national interest indeed pose several problems for any assumptions made prior to the crisis concerning the nature of a rational response." Their argument is that "the attribution of rationality to the decisionmaking process presumes that the parties on each

There are very good reasons to assume that actors pursue their goals as best as they can. If this were not so, then behavior becomes unintelligible. We all tend to assume that on the average actors pursue their goals the best they can given the resources and information they possess and the constraints they must labor under. We then form some assumptions about their preferences over these goals and the assumption that they pursue these goals enables us to form expectations about their behavior. Some interactions are so routine and involve preferences so stable across the population that we do not even have to think about it: our behavior is guided by rules of thumb rather than conscious decisions. We tend to avoid bumping into other pedestrians on the sidewalk because we know this sort of thing will be unpleasant for both but we do not really do it consciously. More importantly, we assume that this preference is shared, which is why we do not expect to be bumped into as well.

More care has to be taken in situations that are riskier. For instance, we also tend to assume that a driver would rather not hit us when we cross the street. However, for him to act on this preference, he must be able to see us in time to react and be able to avoid us when he reacts. When we cross, we take these factors into account by asking how far the car is, whether the driver is likely to see us, and whether he will be able to avoid hitting us. Most of us are quite risk-averse when crossing the street but even in situations when we run the auto gauntlet—as some of us who grew up in large cities with near constant traffic know—we assume that the drivers do not want to hit us. We then form an expectation about their behavior (that they will do what they can to avoid hitting us) and then we decide whether to cross and when to do so. Of course, we know that drivers know that we do not want to be hit either. The danger is that they might assume that we will jump aside to avoid the accident precisely when we are assuming that they will swerve for the same purpose. Which one of us has not deliberately turned his head away from the driver to demonstrate that he "cannot see him"? The implication is that if we cannot see the car, we cannot act on our preference to avoid being hit. This leaves the driver with the remaining option to swerve and we "win" the unequal confrontation. The point is that all this behavior is predicated on the assumption that choices are, to a large extent, predictable because they connect rationally to preferences. The fact that people are sometimes hit is not remarkable. The fact that they are so seldom hit given traffic density is.

It is possible that actors make mistakes because of faulty interpretation of information, or wrong decision-making under stress, or incorrect implementation of correct decisions. However, it is difficult for me to believe that mistakes are systematic. It seems much more fruitful to treat them as deviations from the optimal course of action that *may* occur but in a more or less random fashion. If mistakes were systematic, one would have to wonder why actors do not correct them. Actors do make mistakes, true, but they also learn from these mistakes. Whether this results in them making fresh mistakes—as the famous quip that the generals are always preparing to re-fight the last war suggests—is a depressing (but unlikely) possibility. If actors do not act in their own interest, then we cannot hope to understand their behavior, much less form expectations about it. Anything is possible in

side of the threat possess full information" and that "during periods of high tension, decisionmakers tend to adopt simplified cognitive structures; goals are reduced, and the range of perceived alternatives shrinks." In other words, rationality supposedly requires full information about all possibilities. As we have seen, rationality does not require full information and we will not be considering all possibilities. The very simplicity of the formal model—something that critics often fault it for—is in fact its strength here.

such a world and therefore nothing is comprehensible. Every action can be "explained" by assuming that actors are deluded or inept or both. If this were true and actions were divorced from preferences, then it is a mystery why decision-makers spend so much time trying to divine the intent of their opponents and search their actions for meaning.

Although the concept of rationality used here is rather thin, the assumption that the players are unitary actors is more problematic. States are not individuals, they are collectives that comprise groups that themselves may be composed of other groups, all the way down to the individual. One may think of domestic politics as a way of aggregating these individual preferences in some sort of collective preference. Different political systems enfranchise different segments of the population in various ways. In the end, however, all that matters from our perspective is how these individual (rational) preferences translate into state preferences. It is well-known that there is no way to guarantee that the preferences of a collective will be rational even if the collective itself comprises rational individuals. No way, that is, except taking one of these individuals' preference as the one for the collective (Arrow, 1970).

A complete theory of crisis behavior would have to take domestic politics into account. It would have to show how the (possibly competing) interests of various groups within the state coalesce to determine state behavior. I will not do so in this book for two reasons. First, the underlying logic that will arise between two unitary actors will also be present in the more complicated interaction albeit at the lower level of aggregation. Whether it translates into analogous behavior at the state level remains to be seen but the fundamental problem will remain whether or not the solution is the same. Second, when it comes to the types of disputes that may end in war—the crises where military threats are employed—decision-making is usually restricted to a small group of people at the highest level. Collective irrationality is less likely to arise in smaller groups, especially when their members are not too dissimilar in their preferences, which tends to be the case at the highest level of political power. Whether this assumption is too distorting remains to be seen but, as before, the individuals would have to confront the basic issues that arise from the unitary actor interaction regardless.

Throughout the book I will consider two-actor interactions only. This allows me to abstract away from many important considerations that would doubtless affect behavior in the real world. For instance, in the real world, decision-makers are likely to take into account the expected behavior of their allies, of potential other belligerents, or of non-aligned states that may be carefully monitoring the interaction. Limiting the model to only two actors serves to illuminate the features of military threats that have bearing on the puzzle of credible communication. This may not be the only concern policy-makers have in their confrontation but it will be among the most important ones. Hence, a thorough investigation of this isolated role of threats is a necessary first step toward a theory of their use.

The restriction of attention to a single encounter is made to remove any considerations for consequences of one's actions beyond the current crisis. Reputational concerns and long-term repercussions can enter this model only as part of the payoff specification. In other words, while it is possible to incorporate them, I will only do so by assuming that they can be reduced into the payoffs. A richer theory would model future interactions to see how the consequences one expects to follow tomorrow affect behavior today.

A more intriguing problem with the single-encounter assumption is its implications for equilibrium behavior. In this book, the analysis boils down to finding the optimal course of action in a crisis for each of two opponents. What we are looking for are strategies for

the actors that are mutual best responses: neither actor has incentives to choose a different strategy given what his opponent is doing. The strategies then form an equilibrium because neither actor would want to deviate from his strategy. This approach depends on actors making accurate conjectures about the strategies of the other players. It is known that common knowledge of rationality is not sufficient to guarantee that conjectures about behavior will be correct.[5] The upshot is that we do not know that rational players would necessarily choose actions that are prescribed by the equilibrium strategies. One common justification for expecting them to is that actors learn to play the game through repeating the interaction and successively refining their conjectures (Binmore, 2007). In our context, actors who have more experience with each other because they encounter the same game repeatedly will be more likely to behave how equilibrium logic predicts they should (provided our model is capturing the essence of the interaction). By assuming a single encounter, we effectively destroy the possibility for learning. If actors confront an entirely unfamiliar environment, then their behavior may deviate significantly from the equilibrium prescription.

I have several responses to this problem. First, as I explained above, the fact that I do not model repeated interactions does not mean that one cannot think of the model as representing one encounter among several similar ones. The model will make incorrect equilibrium prescriptions if it does not specify the actors' incentives properly but that has nothing to do with their ability to play the game. Second, in high-stakes encounters where military threats are possible, decision-makers have very strong incentives to analyze their options much more carefully than we normally would in everyday life. It is more likely that they arrive at the optimal course of action and expect their opponents to do so, which means they should be able to make conjectures that are more likely to be correct on the average. Third, even in single-shot encounters of this type, decision-makers are likely to bring their prior experience and their knowledge of the opponent's past behavior into their analysis. Moreover, at this level decision-makers are often socialized through years of experience within relevant bureaucracies or decision-making groups which are likely to have imparted code of behavior which is derived from the experience of the organization: corporate learning, if you will, that extends beyond the individual. In other words, decision-makers may be able to do quite well even in situations they have not faced before provided these situations are not totally unique and the decision-makers' background (or their advisors' backgrounds) include socialization within organizations that have longer memories and experience.[6] Finally, even if one does not buy into any of the defenses above, I am prepared to concede that this problem may limit the predictive power of game-theoretic models. However, in no way does it undermine their usefulness as tools to clarify the logic of the interaction. Furthermore, we can still treat predictions from equilibrium analysis as hypotheses that can be subjected to empirical testing. Whether actors do form the correct conjectures then becomes an empirical question (again, provided that the model captures "enough" of their interaction). Without an alternative logically consistent theory to explain their behavior, this may be the best we can do.[7]

The models in this book also presuppose there exists a conflict of interest between the

[5] Pearce (1984); Bernheim (1984); Brandenburger (1992).

[6] See Farkas (1998) for an argument that, if stretched a bit, may be used to support this line of reasoning.

[7] See Powell (1999, 23–39) for an excellent recent discussion of the use of formal models.

actors. I begin by investigating what happens after one of them has challenged the status quo by initiating a crisis. In then extend the model to analyze the conditions under which one would choose to do so, and the consequences this has for crisis behavior. The one thing I will not analyze is where these conflicts of interest come from. Indeed, such a thing is well beyond the scope of this study although one may well wonder why disputes do not get resolved before they reach the acute stage where the use of force is contemplated.

This book makes heavy use of game-theoretic models. However, I have made a determined effort to minimize the algebra and maximize the exposition of the logic it represents. Although some prior knowledge of game theory will certainly be helpful, it should not be a prerequisite for understanding my arguments.[8] I have moved the technical material to various appendixes, and even then I have omitted most of the gory and tedious detail; the little math that remains in the text should be accessible to any reader with basic algebra. The book can be read without reference to these appendixes and I have provided plenty of examples and figures to illustrate the insights and intuitions that emerge from the formal analysis.

[8] Morrow (1994b) provides an accessible introduction to game theory with an emphasis on political science in general and international relations in particular. McCarty and Meirowitz (2007) is a more rigorous treatment at approximately the same level.

2

Commitment and Signaling in Coercive Bargaining

What you cannot enforce, do not command.
Socrates

In this book, I analyze military threats: how states can use them to establish commitments and credibly communicate intent. My purpose is not to create a theory of interstate crisis behavior but rather to illuminate the logic of military threats insofar as they are intended to deal with the above two concerns. To assess the utility of using military threats, we must begin with a look at the fundamental strategic problems that confront decision makers in international crises. This chapter lays the groundwork for the theory to follow. Most of the material is well-known and I have no wish to rehash widely available results. However, one does not develop a theory in an intellectual vacuum, and to appreciate the argument it will be useful to outline in a general manner the crucial issues that animate most of our current thinking about crisis bargaining.

What follows is not a compendium of results, it is an attempt to provide a unifying framework for thinking about the various mechanisms for credible communication in crisis. I begin by constructing a stylized representation of a crisis that will serve to highlight the role of uncertainty, and to provide the baseline for theorizing about the use of the military instrument. I then explore the various mechanisms for credible signaling that scholars have proposed in the context of this basic model. When there is uncertainty about the commitment of a resolved actor, that actor has very strong incentives to attempt to communicate his resolve to the opponent to avoid bad outcomes in the crisis. The main argument in this chapter is that whereas scholars have suggested numerous ways in which this can be done, none of these mechanisms reflects what military threats can do. In other words, we simply do not know the logic of military threats. The rest of the book is about expounding this logic.

2.1 A Stylization of A Crisis

Assume a conflict of interest exists between two rational unitary actors and that they can resort to arms to resolve it. The crisis game illustrated in Figure 2.1 is among the simplest models of an interaction in which one of them confronts the other over the possession of a disputed good.[1]

[1] The simple model is similar to the baseline crisis model in Schultz (2001, Ch. 2), but not equivalent. One very important distinction I make is that the valuation of the disputed good—what Jervis (1979) calls "intrinsic interest"— is the fundamental source of uncertainty. Although the general thrust of the results holds, the specifics are different. As will become clear in the following chapters, this assumption about preferences is

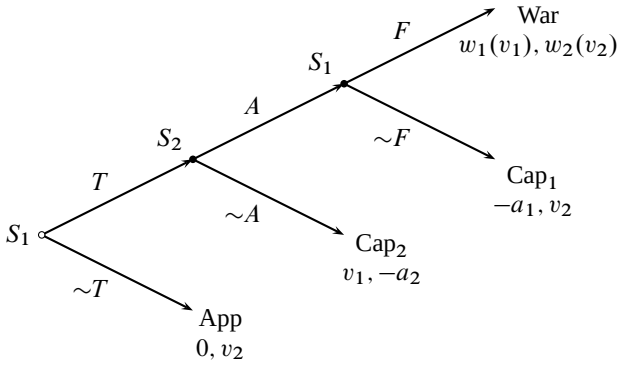


Figure 2.1 The Basic Crisis Game

In this game, a defender, S_1 is in possession of some valuable good—strategic territory, land with desirable resources, exclusive market access—that a potential challenger, S_2 wants to obtain.[2] Each player values the possession of this good at $v_i > 0$. For simplicity, I assume that this good is indivisible: states cannot share, if one has it, the other is necessarily denied any reward. The game begins when a crisis is already underway— S_2 has issued a demand for the good—and S_1 has to decide how to cope with the challenge. If S_1 does not escalate by threatening ($\sim T$), the game ends in appeasement and S_2 obtains the disputed good peacefully. S_1 is left with nothing, so his payoff is 0, and S_2 enjoys the rewards of possession, v_2. If S_1 threatens S_2, the challenger may either act (A) or give up ($\sim A$). Yielding to S_1's threat ends the crisis without the transfer of the good. The payoffs are v_1 for S_1 and $-a_2$ for S_2, where $a_2 > 0$ denotes S_2's reputational loss for conceding to her opponent's threat. If, on the other hand, S_2 resists in response to the threat, the defender must make the final choice between peace and war. He can yield ($\neg F$), ending the crisis with capitulation. In this case, the payoffs are $-a_1$ for S_1 (the reputational effect is reversed), and v_2 for S_2. If he fights (F), however, the crisis ends with a war. Each players receives his respective expected payoff from war:

$$w_1(v_1) = pv_1 + (1-p)0 - c_1 = pv_1 - c_1$$
$$w_2(v_2) = p0 + (1-p)v_2 - c_2 = (1-p)v_2 - c_2,$$

where $p \in (0,1)$ is the probability that S_1 will win the war, in which case he retains the good, $(1-p)$ is the probability that S_2 will win, in which case she obtains the good, and

also the basis for extending the model to account for militarized coercion properly. For other models very close to the one I present here, see the three-stage crisis in Wagner (1982), the Unilateral Deterrence game in Zagare and Kilgour (2000, Ch. 5), and the economic sanctions model in Drezner (2003). Furthermore, Bueno de Mesquita and Lalman's (1992) international interaction game embeds a crisis subgame, and Fearon (1997) builds on a similar model. This representation of a crisis is also consistent with the implicit non-formal models used, *inter alia*, by Huth (1988), and Huth and Russett (1993). The model essentially incorporates all the features that O'Neill (1999, 133) enumerates for what he calls "crisis signaling models."

[2] Here, and throughout the book, I will refer to odd-numbered players as "he," and even-numbered players as "she" to facilitate exposition. I also use the male pronoun when referring to generic abstract players, and generally use the female pronoun when referring to countries.

$c_i > 0$ are the costs of war the players must pay when they fight. Because we want to avoid trivial cases where states go to war because they both prefer it to living without the good, I make the following assumption:

ASSUMPTION 2.1 (War is Costly). Fighting for the good is worse than not having it. That is, $w_i(v_i) < 0$ for all v_i.

Since this is a game of complete and perfect information, we can analyze it by backward induction from the terminal node.[3] S_1 will attack only if doing so is better than capitulating:

$$w_1(v_1) > -a_1. \tag{CR$_1$}$$

This is the defender's *credibility constraint*. If this condition is satisfied, then he will go to war if the challenger resists his threat. This means that in order to determine what the challenger will choose to do, we have to examine two cases. First, suppose (CR$_1$) is satisfied. If S_2 resists, she must expect S_1 to attack, so the outcome will be war. The challenger will act if, and only if, war is preferable to capitulating:

$$w_2(v_2) > -a_2. \tag{CR$_2$}$$

This is the challenger's credibility constraint. If she is sure that resistance will result in war, she would only resist when this condition is satisfied.

Now suppose that (CR$_1$) is not satisfied. If S_2 resists, she must expect S_1 to capitulate, so the outcome will be peace. The challenger will resist if, and only if, peace is preferable to capitulating: $v_2 > -a_2$, which is always true because $a_2 > 0$. Hence, if the defender does not have a credible threat to attack, the challenger will always resist his initial threat.

We conclude that whenever the defender's credibility constraint is satisfied, the challenger's response to a threat will depend on her own credibility constraint (she will give up if it is not satisfied and will act otherwise). If, on the other hand, the defender has no viable threat, then the challenger will always resist his threat regardless of her own credibility constraint.

Turning now to the defender's initial choice, there are three cases to examine. First, suppose that (CR$_1$) and (CR$_2$) are both satisfied. If S_1 threatens, S_2 will resist, and the outcome will be war because S_1 will attack. He will threaten, and only if, war is preferable to the loss of the good: $w_1(v_1) > 0$. However, by Assumption 2.1, this condition never holds. Therefore, S_1 will never threaten and the status quo will be peacefully revised in S_2's favor.

Second, suppose that (CR$_1$) is satisfied but (CR$_2$) is not. If S_1 threatens, S_2 will capitulate because she knows that resistance will lead to war, which would be worse. S_1 will threaten if, and only if, retaining the good peacefully is preferable to losing it, $v_1 > 0$, which is always true. In this case, the defender has a credible threat but the challenger does not, and so compellence succeeds. The outcome is a peaceful maintenance of the status quo.

Finally, suppose that (CR$_1$) is not satisfied. If S_1 threatens, S_2 will resist, and he will be

[3] In technical terms, we will find the *subgame-perfect equilibrium* of the game (Selten, 1975). This equilibrium requires that actions be optimal at all points in the game, not just along the equilibrium path, as Nash (1951) equilibrium does. This requirement ensures that equilibrium threats are credible in the sense that actors will carry them out if they have to. Nash equilibrium behavior may depend on contingencies that never arise if the equilibrium strategies are followed. The requirements for Nash equilibrium cannot assess the optimality of actions in such contingencies and therefore cannot guarantee that these threats are credible.

forced to capitulate. S_1 will threaten if, and only if, capitulating is preferable to giving up the good peacefully, $-a_1 > 0$, which is never the case. That is, when the challenger is the only one with a credible threat, S_1 would avoid incurring reputational losses and would not even attempt to compel his opponent. The following proposition puts these results together.

PROPOSITION 2.1. *If actors are completely informed about the credibility of their commitments, then war will not occur. If both commitments are credible, then the defender will never threaten because doing so would lead to war. Neither will he threaten if he cannot commit to fight. He will only threaten if he has a credible commitment but his opponent does not, in which case she will capitulate with certainty.*

To summarize our findings, compellence can fail in two general cases. First, if the defender has no credible threat to attack if he is resisted, the challenger will act regardless of her preference for war, and, foreseeing this, the defender gives up. Second, if the defender does have a credible threat to attack, compellence success will turn on the challenger's credibility: if her threat to fight is credible, the defender will avoid war and capitulate at the outset. The only time compellence will succeed in this model is when the defender has a credible threat but the challenger does not. Note that under no circumstances do the players actually go to war in this game. This makes sense given the game's simple structure and the players' ability to foresee perfectly the consequences of their moves because of complete information.

There is some disagreement in the literature about the meaning of credibility. Many analysts take it to be synonymous with believability: a threat is credible if the opponent believes it.[4] Morgan (2003, 15) states it flatly: "Credibility is the quality of being believed.... it was not a state's capacity to do harm that enabled it to practice deterrence, it was others' *belief* that it had such a capacity. What deterred was not the threat but that it was believed."

I think that this conflates two logically distinct issues: an actor may be committed to a course of action and yet his adversary may not believe him; alternatively, an actor may be bluffing and yet his adversary may believe him. Can we then say that the threat was not credible in the first case but was in the second? To avoid this confusion, I will call a commitment credible (or genuine) if the actor would carry out the course of action it prescribes. This is an attribute that is known to the actor but may or may not be known to his opponent. Whether this opponent believes a commitment is a separate issue entirely, and one that I will investigate at some length. In our basic model, an actor whose credibility constraint is satisfied has a credible commitment. Of course, because this is a model of complete information, believability is not an issue.

Before we turn to cases where it is, however, I should also note that in the deterrence literature in particular, credibility is taken to be a multiplicative function of capability and intent (Singer, 1984, 56-57). That is, a threat is credible if the actor intends to carry it out *and* if it is capable of inflicting sufficient damage on the opponent. These are completely separable issues, however, and it conflating them has caused much confusion. In this book, we shall maintain a rigorous separation: a threat is credible is the actor is willing to carry it out. It may or may not be capable enough to alter the decisions of the opponent. The

[4] Schelling (1966); Mueller (1995); Freedman (2003). Zagare and Kilgour (2000, Ch. 3) provide an intelligent and very useful discussion of the concept of credibility and how it relates to rationality.

distinction, as it turns out, is quite important: we shall see threats that are fully credible but have no pretense to affect the expectations of the opponent and coerce him into settling without fighting. Conversely, we shall see coercive threats that may or may not be credible.

2.2 The Purely Informational Approach

The analysis demonstrates quite clearly that behavior in this model is largely determined by the credibility of the actors' commitments. The classical approach is to assume that actors either have credible threats or they do not, and then investigate the consequences of being unsure about each other's commitments. In the context of our simple crisis game, the question reduces to the actor's choice between capitulating under duress and fighting. Kilgour and Zagare (1991, 326) provide a cogent summary of this approach:

> by identifying the credibility of each player's threat to retaliate with the probability that a player prefers retaliation to capitulation, we maintain consistency with both the traditional strategic literature, in which credibility is usually equated with believability, and with the literature of game theory, in which credibility is usually taken to be synonymous with rationality.

For example, consider the remilitarization of the Rhineland in 1936 with Germany as the challenger and France as the defender. If Hitler's credibility constraint was satisfied, he could commit to fighting for the Rhineland if the French resisted his demand. The French, in turn, were sufficiently convinced that it was quite probable that it was satisfied, and hence an attempt to block the entrance of German troops would most likely result in a war. They were not prepared to fight such a war, and hence their credibility constraint was not satisfied. Hitler gambled on that and won: his demand was not resisted. The ironic aspect of all this, as we now know, is that Hitler was bluffing. He would not have attacked had the French resisted but they did not know it.[5]

There are two points that follow from this example and our model. First, the traditional approach assumes that a commitment is either credible or it is not (i.e., the actor would either fight for something or he would not). Second, the opponent may not actually know about it. Sometimes this works to the actor's advantage (in 1936 Hitler got away with militarization even though he did not have a credible commitment), but sometimes it works to his disadvantage (in 1939 Britain and France ended up in a war even though they did). The problem is one of information: at least one of the actors did not know if his opponent's commitment was credible.

We can summarize the traditional approach to credibility as follows. *Actors would either fight or they would not for a good. This preference is assumed to be fixed and known to the actor himself. However, it may or may not be known to his opponent. The difficulty lies with communicating these commitments credibly.*[6]

Now, some commitments are inherently credible and so are not interesting from an analytical perspective. As Schelling (1966, 35) succinctly put it, "No one seems to doubt that federal troops are available to defend California. I have, however, heard Frenchmen doubt whether American troops can be counted on to defend France." The US government presumably would have no trouble persuading an opponent that it would rather fight than give up

[5] Robertson (1967), Weinberg (1970).

[6] As we shall see later, many researchers are quite aware that it is possible to create commitments by modifying incentives.

California. But what is an opponent to think about its commitment to the defense of France?[7] The US government could prefer to fight, in which case it has a credible threat, and the task would be to communicate that fact in some persuasive manner. On the other hand, it may not prefer to risk losing New York to defend Paris, in which case it has no credible threat, although it may try to bluff on occasion as if it had. From the Russian perspective, the difficulty lies in distinguishing between these two possibilities: which "type" of US government is it facing: the committed or the bluffing?

This is the essence of the purely informational approach: the existence of this commitment is assumed in the preferences of the actors and the task is to manipulate the information available to the opponent to one's advantage. To illustrate this, consider the model under two-sided incomplete information. Neither actor knows if his opponent's credibility constraint is satisfied. Instead, each has a *prior belief* about the valuation of the other, and these beliefs are common knowledge. At the beginning of the game, both players know their own valuations. To represent their beliefs about the valuation of their opponent, assume that v_1 is drawn by the cumulative distribution function $F_1(\cdot)$ from the interval $[0, \overline{v}_1]$, and that v_2 is drawn by $F_2(\cdot)$ from the interval $[0, \overline{v}_2]$.[8]

Because S_1 is uncertain about S_2's valuation, he is uncertain about whether her credibility constraint is satisfied. Recall from (CR_2) that S_2 will prefer to war to capitulation whenever $w_2(v_2) > -a_2$, which we can rewrite as $v_2 > (c_2 - a_2)/(1-p) \equiv v_2^*$. Since S_1 is unsure what v_2 is, he does not know whether it exceeds v_2^* or not. However, because he has beliefs about it, he can estimate the probability that it does. This probability is $\Pr[v_2 > v_2^*] = 1 - \Pr[v_2 \leq v_2^*] = 1 - F_2(v_2^*)$. Thus, $1 - F_2(v_2^*)$ is S_1's prior belief that S_2's credibility constraint is satisfied. Players are assumed to form their prior beliefs on the basis of their historical experience with each other, observations of past behavior in similar situations involving other players, intelligence estimates of current capability, and so on (Morrow, 1989a).

The credibility threshold v_2^* and the analogous one for S_1 play such an important role in this analysis that is worth defining them explicitly:

$$v_1^* = \frac{c_1 - a_1}{p} \quad \text{and} \quad v_2^* = \frac{c_2 - a_2}{1-p}. \tag{2.1}$$

Since we are interested in cases where actors may or may not have credible commitments, we assume that $\overline{v}_i > v_i^*$ and $c_i > a_i$. These assumptions guarantee that it is possible, but not necessary, that players have credible commitments. For example, $\overline{v}_1 < v_1^*$ implies that the highest possible valuation for S_1 is less than his credibility threshold, which means that S_1 will never attack regardless of his valuation. Similarly, $c_1 < a_1$ implies that the credibility threshold is negative, and because we assumed that valuations are non-negative, this means that S_1 will always attack regardless of his valuation. In the first instance, the credibility constraint is never satisfied, and in the second it always is. We have already seen what happens when it is known whether the constraint is satisfied or not. The additional assumptions guarantee that there is some uncertainty about that.

We shall call a player whose valuation satisfies the credibility constraint *resolved* (or tough), and one whose credibility constraint is not satisfied, *unresolved* (or weak). We shall

[7] This is a Cold War example, so it refers to an attack by the Soviet Union upon France.

[8] The distribution functions also have continuous and strictly positive densities, $f_1(\cdot)$ and $f_2(\cdot)$.

maintain our assumption that war is costly, and so $w_i(\overline{v}_i) < 0$. This means that we are not assuming that a resolved player will escalate regardless of his opponent's behavior, only that if forced to choose between capitulating after escalation and certain war, he will fight. War is still worse than the status quo even for a resolved player.

Observe now that before S_2 decides whether to resist S_1's threat, she must estimate the likelihood that S_1 will attack if she does. Because S_1 will only attack if his credibility constraint is satisfied, S_2 must estimate the likelihood that it is. Now, recall that S_2 entered the game with a prior belief about S_1's valuation. However, now that S_1 has chosen to threaten, she may have learned something new about his valuation. In other words, S_2 will have a *posterior belief* which reflects what she learns from the observable behavior of her opponent. This is her new estimate of S_1's valuation.

Where does this estimate come from? It is natural to assume that S_2 will take into account both her prior beliefs and S_1's observable behavior, and she will somehow amalgamate all that into a new estimate of her opponent's resolve. As in statistics, we shall use Bayes rule to update old information with outcomes from new observations. This rule answers the question, "Given that S_1 has threatened, what is the probability that he is resolved?"

Why is it important to track how an actor would change his beliefs during the crisis? Because his behavior will depend on them, which means that his expectations are the crucial target that the adversary will attempt to manipulate. As Ellsberg (1975, 8) puts it, "To be effective, [the] threatened punishment need not be certain, only 'sufficiently likely.' " He then goes on to develop the idea of *critical risk*, which is the maximum risk of punishment one is willing to tolerate, and then shows how one can manipulate this critical level.[9] Although Ellsberg defines critical risk in terms of subjective estimates of how likely the other side is to follow through on its threat, he does not really talk about where these estimates come from and, perhaps more importantly, how they should change in light of new information players obtain from their interaction. As it turns out, this question is precisely the crux of the problem for if beliefs, and through them consequent behavior, are expected to change, then the opponent will attempt to influence them. But knowing this, would the actor then change his beliefs on the basis of what the opponent does? And if the opponent knows that the actor will ignore him, why would he attempt to pattern his behavior in a futile attempt to manipulate expectations? We need a way out of this circular morass.

We must ensure that beliefs are consistent with the strategy that S_2 thinks S_1 is playing. In equilibrium, S_2 knows the strategy of her opponent, which implies that her beliefs must be consistent with his equilibrium strategy. Because S_1 knows his own valuation, his strategy will depend on it. For instance, it may be "threaten only if $v_1 > 1/2$." If S_2 expects him to play this strategy and then observes a threat, she must infer that v_1 is at least $1/2$, otherwise no threat would have materialized. From her perspective, S_1 would have only threatened if his valuation is between $1/2$ and $\overline{v}_1$, and since he did threaten, it must be in that range. The probability that S_1 will attack if she resists is still $\Pr[v_1 > v_1^*]$ with the important proviso that $v_1 > 1/2$ too. Bayes rule answers the question "what is the probability that S_1's credibility

[9] Most of the strategies he suggests should be already familiar from Schelling (1960), although the formalization of some of them is novel. Some prominent applications of critical risk for analysis of international crises are Snyder and Diesing (1977) and Wagner (1982).

constraint is satisfied given that he threatened?" This conditional probability, $\Pr[v_1 > v_1^*|T]$, can be computed from the prior belief and S_1's strategy.

In our example, this is very easy: since only $v_1 > 1/2$ threaten, the probability of a threat is $\Pr[v_1 > 1/2] = 1 - F_1(1/2)$. As before, the probability that S_1's credibility constraint is satisfied is $\Pr[v_1 > v_1^*] = 1 - F_1(v_1^*)$. Assuming that the resolved types are among those who threaten, the updated probability is just $\Pr[v_1 > v_1^*|T] = (1-F_1(v_1^*))/(1-F_1(1/2))$. To get some intuition what this means, suppose $v_1^* = 3/4$ and that v_1 is uniformly distributed between 0 and 1. The prior probability that S_1 is resolved is $\Pr[v_1 > 3/4] = 1/4$. The probability of a threat is $\Pr[v_1 > 1/2] = 1/2$. The posterior probability that S_1 is resolved when a threat is observed is $\Pr[v_1 > 3/4|T] = (1/4)/(1/2) = 1/2$. Intuitively, S_1's threat and S_2's conjecture about his strategy have caused S_2 to revise her belief: whereas initially she believed there was only 25% chance that S_1 is resolved, she now believes that it is 50%. The valuations at which S_1 is resolved constitute only a quarter of the possible valuations he might have but fully a half of the valuations that cause him to threaten. Hence, S_2's belief that the chance of him attacking is 50% is consistent with her prior and her conjecture about his strategy. We shall use Bayes rule to ensure that beliefs are consistent in this way.[10]

I do not claim that this is how decision-makers learn in practice. However, even if one objects to using Bayes rule and consistency on the grounds that actual decision-makers do not use them for inference, any alternative proposed would have to include some of their properties. Most importantly, it has to incorporate "the strategic element of learning: since statesmen know that their behavior will influence the expectations of others, they have an incentive to take this into account in making their choices; but then others must take this incentive into account in deciding what inferences to draw from the actions in question" (Wagner, 1992, 139). This is precisely what the consistency requirement is designed to accomplish: when making inferences about his opponent, each player takes into account his expectations about the opponent's strategy which is itself consistent with the opponent's expectations about how the player would behave. This circular reasoning is very hard to disentangle without the aid of game theory (and even sometimes with it).[11]

The posterior belief that S_2 will form upon being threatened allows her to estimate the likelihood that S_1 will attack if she resists. Since we have not yet determined S_1's strategy, we shall keep her posterior beliefs abstract for now. Let $G_1(v_1^*) = \Pr[v_1 \leq v_1^*|T]$ be S_2's belief that S_1 will capitulate if she resists his threat. Given this belief, S_2 will resist if, and only if, the expected payoff from doing so is better than capitulation: $G_1(v_1^*)v_2 + (1 - G_1(v_1^*))w_2(v_2) > -a_2$. We can rewrite this as $v_2 > \hat{v}_2$, where:

$$\hat{v}_2 = \frac{(1 - G_1(v_1^*))c_2 - a_2}{1 - p(1 - G_1(v_1^*))}. \tag{2.2}$$

[10] We shall consider only consistent beliefs, that is, ones derived from the supposed optimal strategy of the defender, as plausible candidates for S_2's updated estimate. In technical terms, we are using the concept of *perfect Bayesian equilibrium* developed by Fudenberg and Tirole (1991b). It is a dynamic refinement of Bayesian equilibrium (Harsanyi, 1968), and is designed to eliminate implausible solutions analogous to the way a *subgame-perfect equilibrium* (Selten, 1975) eliminates unreasonable Nash (1951) equilibria.

[11] It is also true, however, that these solution concepts place rather heavy demands on what players are assumed to know about the game (McGinnis, 1992; O'Neill, 1992, 469-74). Levy (1994) provides an overview of the research on how leaders learn and change their beliefs.

In other words, given her updated beliefs, S_2 will resist if her valuation exceeds the critical threshold, $v_2 > \hat{v}_2$, and will capitulate otherwise. Observe now that if S_2 believes that S_1 will fight for sure, her posterior will be $G_1(v_1^*) = 0$. In this case, the critical threshold for resistance is the same as her credibility threshold: $\hat{v}_2 = v_2^*$. Not surprisingly, if S_2 is certain that war will follow, she will resist only if she prefers to fight rather than capitulate. If, on the other hand, she believes that there is some chance that S_1 will back down when resisted, her posterior will be $G_1(v_1^*) > 0$. In this case, the critical threshold for resistance is *lower* than the credibility threshold: $\hat{v}_2 < v_2^*$. Substantively, this means that S_2 will be willing to risk war by resisting even if she is not actually resolved to fight. The reason is that although there is a chance of war, there is also a chance of getting S_1 to capitulate.

This behavior allows S_1 to estimate the consequences of making a threat. Since S_2 capitulates if her valuation is less than the critical threshold for resistance, from S_1's perspective, the probability of capitulation to his threat is $\Pr[v_2 \leq \hat{v}_2] = F_2(\hat{v}_2)$. This implies that that the more likely is S_1 to back down when S_2 resists, the more likely is S_2 to resist.

It is crucial to realize that S_2's behavior is critically dependent on her beliefs, which in turn depend on S_1's behavior. Given that S_2 will update her beliefs based on what she thinks S_1's optimal strategy is and that this will affect her behavior, it makes sense that S_1 will try to influence these beliefs with his actions. That is, he will take into account how his behavior will alter S_2's expectations, and will consider the indirect effect of his actions that are mediated by his opponent's beliefs. This is entirely in keeping with numerous analyses of crisis bargaining that conclude that actions are important not just because of their direct impact but also because of their effect on the expectations of the crisis participants.[12]

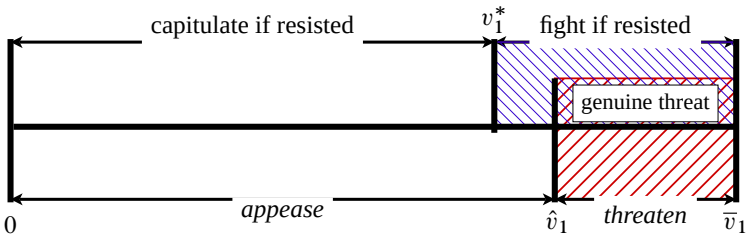
With all the machinery in place, we are now ready to analyze S_1's initial choice. S_1's expectation is that if he threatens, S_2 will concede with probability $F_2(\hat{v}_2)$ and will resist with complementary probability. If S_1 is resolved, he will fight when resisted, so the expected payoff from threatening is $F_2(\hat{v}_2)v_1 + (1 - F_2(\hat{v}_2))w_1(v_1)$. If, on the other hand, he is not resolved, he will capitulate when resisted, so the expected payoff from threatening is $F_2(\hat{v}_2)v_1 + (1 - F_2(\hat{v}_2))(-a_1)$. Hence, S_1 will threaten if, and only if,

$$F_2(\hat{v}_2)v_1 + (1 - F_2(\hat{v}_2)) \times \max\left\{w_1(v_1), -a_1\right\} > 0. \tag{2.3}$$

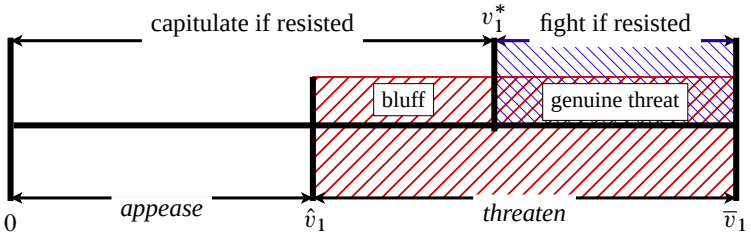
We need to examine (2.3) carefully because there may be no types for which it is satisfied. For example, if S_2 is extremely unlikely to capitulate—i.e., $F_2(\hat{v}_2)$ is close to zero—the expected payoff from threatening will be very close to the maximum of the war and capitulation payoffs. Since both are worse than accepting a revision of the status quo even for the highest-valuation type, S_1 will never threaten in equilibrium. If, however, (2.3) is satisfied for $\overline{v}_1$, then there exists some $\hat{v}_1$ such that it holds with equality. This type is indifferent between escalating and accepting a peaceful revision, and it plays a pivotal role in the analysis to follow. S_1 escalates if $v_1 > \hat{v}_1$ and appeases otherwise. To summarize:

- v_1^* is the credibility threshold for S_1: he attacks when his threat is resisted if, and only if, his valuation exceeds that threshold;
- $\hat{v}_1$ is the escalation threshold for S_1: he threatens if, and only if, his valuation exceeds that threshold;

[12] Iklé (1964); Schelling (1966); Wagner (1992).



(a) Genuine Threats



(b) Bluffing Possible

Figure 2.2 Defender Behavior in the Crisis Game.

- $\hat{v}_2$ is the resistance threshold for S_2 given how likely she thinks S_1 is to attack if she resists: she resists S_1's threat if, and only if, her valuation exceeds that threshold. If she is certain that S_1 will attack, then $\hat{v}_2 = v_2^*$; that is, she will only resist if her credibility constraint is satisfied.

There are two general cases to consider: either $\hat{v}_1 \geq v_1^*$, which means that S_1 only escalates if his credibility constraint is satisfied, or $\hat{v}_1 < v_1^*$, which means that there are some valuations at which S_1 escalates but backs down if resisted. In the first case, all threats are genuine, and in the second case, the defender might be bluffing. Figure 2.2 illustrates the two scenarios. The central result can be summarized as follows (formal proofs in Appendix A):

PROPOSITION 2.2. *If actors are uncertain about the credibility of each other's commitments, then war occurs with positive probability whenever the defender makes a threat. In equilibrium, there are only two possibilities:*

- *The defender's threat is genuine, in which case the challenger resists if, and only if, she has a credible commitment to fight.*
- *The defender might be bluffing, in which case the challenger may resist even if she does not have a credible commitment to fight.*

Consider first the case in which threats are genuine, as illustrated in Figure 2.2(a). When a threat signals credible commitment, war only happens if the challenger is also fully committed. It is worth noting that some defenders who do have a credible commitment do not threaten at all: the possibility of resistance discourages them. No unresolved type threatens either. Compare this to the complete information case with (CR_1) satisfied. If (CR_2) is also

satisfied, S_1 never even threatens, a resolved challenger does not face any risk of war. With incomplete information, S_1 does not know if S_2's commitment is credible, and he risks escalation if he is resolved. This may end in war if the challenger happens to be resolved as well. Hence, compellence can fail even if both sides possess credible commitments. The problem is the inability to communicate them in a believable way without running a risk of triggering war. However, because of the high risk involved, only genuine defenders would escalate, and so only resolved challengers would resist an escalation. There is no bluffing in this equilibrium by either side. In fact, the risk deters even genuine defenders who do not value the issue sufficiently.

If, on the other hand, (CR_2) is not satisfied, then with complete information S_2 relinquishes her claim to the good peacefully: compellence succeeds without war. Uncertainty does not change much in the genuine threat equilibrium if the challenger happens to be unresolved: she will concede here too. The main difference is that a resolved defender who would have retained the good under complete information might be forced to give it up because he is afraid to risk escalation on the off chance that the challenger happens to be resolved as well. As Figure 2.2(a) makes clear, if S_1's valuation is $v_1 \in [v_1^*, \hat{v}_1]$, then he will appease even though he has a credible commitment to fight. Uncertainty can certainly benefit an unresolved challenger because it allows her to obtain the good in circumstances where she would have been forced to abandon her claim to it otherwise.

For the resolved challenger, uncertainty is a hindrance because she suffers from the defender's belief that her commitment may not be credible. If she could somehow convince him that this was not so, then war would be avoided. The problem is that there is no way to communicate her resolve such that S_1 would actually believe her, and the culprit are the low-valuation types that would not willingly reveal that they are not resolved: they benefit from S_1 thinking that they are, and S_1 knows this. So he would discount any cheap communication that attempts to convince him otherwise.

Consider now the case in which threats are not necessarily genuine, as illustrated in Figure 2.2(b). Because S_2 believes that there might be a chance that S_1 is bluffing, she is more likely to resist compared to the situation in which threats are genuine. That is, the resistance threshold is lower than her credibility threshold: $\hat{v}_2 < v_2^*$. This implies that if her valuation is all $v_2 \in [\hat{v}_2, v_2^*)$, then S_2 is unresolved but resists anyway. Under complete information, a challenger with such a valuation would either give up the claim to the good when S_1 is resolved or would obtain it if he is not. With uncertainty, however, these types will risk war because S_1's initial choice does not fully separate the genuine defenders from bluffers. When S_1 threatens, S_2 is still unsure if he is serious, and therefore resists. She still obtains the good if she happens to be facing a bluffer but ends up in a war if she is not. The problem is that S_1 did not have an opportunity to take the "correct" action initially—correct being defined as the action he would want to have taken after S_2 resists. When resistance is a fact, a bluffer would strictly prefer never to have escalated in the first place but the irony is that he escalated because he was unsure of S_2's commitment. That is, he would have preferred to know if S_2 is committed, and a challenger with valuation $v_2 \in [\hat{v}_2, v_2^*)$ would have preferred that as well! From the perspective of such a challenger, she obtains the good if S_1 is unresolved (same as under complete information) but has to fight a war if he is resolved instead of capitulating in peace as she would have done under complete information. There appears to be mutual interest in revealing one's resolve.

The problem, as usual, is that the defender might have a low valuation. Suppose that S_2 revealed that she is not committed, and would therefore act only if she was sure that S_1 is not committed himself. S_1 now has no incentive to reveal truthfully his valuation if he is unresolved because doing so means appeasing S_2 but a successful bluff means preserving the status quo and keeping the good without fighting. Hence, he would prefer to misrepresent his valuation, and, knowing this S_2 obtains no benefit from telling the truth initially. After all, since S_1 is expected to bluff, and the probability of this happening depends on his optimism (he is less likely to bluff is he believes S_2 is committed), S_2 would not want to forego the advantage of concealing her true valuation and pretending that she is resolved too. Although both could profit from truthful communication, their incentives do not allow them to reveal their commitments credibly.

The situation is even worse when we look at a bluffing S_1: this type has no chances of keeping the good under complete information regardless of S_2's commitments. Now there is a strictly positive chance that he can do it, albeit by running a risk of having to capitulate if the challenger resists. Still, if S_2 happens to be unresolved and with valuation lower than $\hat{v}_2$, she capitulates, and S_1's bluff is never called. Clearly, none of the bluffers would ever want to reveal their valuation for it unambiguously hurts their prospects.

We conclude that in a crisis opponents would not choose a strategy that would fully reveal their commitments. It is not that they necessarily do not want to—resolved actors would dearly love to be able to do that—but they cannot do it in a way that would convince their opponent. Credible communication is impossible because an unresolved actor would try to conceal the damaging information by mimicking the behavior of a resolved one. Because of uncertainty, resolved players find themselves in a bind, and they have to look for a way to reveal their resolve such that it is persuasive for the opponent. Effectively, this means doing something that an unresolved player cannot, or would not, do. That is, they must credibly *signal* their commitment.

It is worth emphasizing that crises can happen between two unresolved actors. From S_2's perspective, a threat may be genuine or it may be a bluff. To deter frivolous escalation, S_2 chooses a strategy that induces strategic uncertainty in her opponent's expectations. She does not commit clearly either to resistance (because doing so would not be credible—after all, genuine defenders would go to war and unresolved challengers would want to avoid that), or to capitulation (because this would encourage even unresolved opponents to attempt compellence). S_2 calibrates her strategy such that her risk of having to fight (S_1 turns out to be resolved) is balanced against the gains of obtaining the good peacefully (S_1 turns out to be unresolved). Because bluffing is a distinct possibility, S_2 sometimes will resist even when not committed herself.

What we found echoes Kydd's (2005) succinct conclusion: "If uncertainty is at the heart of crises, then communication is the key of resolving them. The problem is that ordinary communication does not work" (186). Proposition 2.2 characterizes what we should reasonably expect players to do in the stylized crisis game with incomplete information. The findings are not surprising but since they will serve as baseline to compare and contrast the theory I will develop, it is worth summarizing them.

First, difficulties with communicating commitments credibly are at the heart of compellence failure in this model. We assumed that war is worse than the pre-crisis status quo for both players regardless of their resolve. In other words, if they had the choice between fight-

ing and living without the good, both players would prefer to relinquish the good in peace. Since neither player values war for its own sake, they are "peace-loving," and because they prefer to live with the existing distribution of benefits rather than fight, they may also be considered satisfied.

However, as even this simple model clearly shows, *being satisfied with the status quo and being peace-loving in no way guarantees that war will not occur.* The issue here is not whether one is a war-monger but whether one is prepared to risk war to prevent an opponent to take advantage of his reluctance to wage it. When force is an instrument of statecraft and backing down in the face of a challenge carries even small reputational costs, countries may find themselves at war when the attempt at coercion backfires. Unfortunately, coercion must carry this risk of failure because it is precisely this risk that can dissuade frivolous threats. It is the only mechanism here that players can use to communicate the extent of their commitments.

Second, we observed that the crux of the problem for S_1 is to convince S_2 that he is committed not to back down if resisted. The difficulty arose from the consequences of such persuasion: if S_2 is expected to believe that escalation signals resolve, then there is nothing to deter unresolved types from exploiting that. Hence, the only way the challenger could discourage that is not to believe it fully. It is as if S_2 says "if I see escalation, I will increase my estimate that my opponent is resolved, but I will still believe that it is possible he is bluffing." With these beliefs, the challenger will sometimes resist (which she would not have done if she were fully convinced of S_1's resolve), and this in turn discourages some threats from the unresolved types, which justifies S_2's strategy. This is the circular logic that hinges on beliefs being consistent with the strategies players are expected to implement and strategies being rational given these beliefs.

This result illustrates *costly signaling*, the notion that to be believable, a signal must carry with it some disincentive for the "wrong" type to produce it. In this model, a threat (the signal) is costly because it carries the risk of war. It is just costly enough to be discouraging for some unresolved types, and induce them to separate partially from committed ones through their behavior. It is not fully separating because inducing the costliness of the signal is costly for S_2 as well—she must accept some risk of war—and hence there is limit to how high a risk she will be willing to generate. In our case, she would run risks that do not induce full separation. Compellence is a balancing act by both actors, and in Powell's (1990, Ch. 2) apt characterization, its essence is the *search for credibility*.

2.3 The Search for Credibility

There are at least three general ways one can deal with a problem: show that it does not exist, make it irrelevant, or solve it.[13] In our context, the problem is the (lack of) credibility, and the three approaches are:

1 A resolved player successfully communicates his commitment to his opponent (costly signaling);

[13] Obviously, one can also ignore it, at least for the time being. Procrastination could almost be considered a norm in British foreign policy during its imperial period, and much can be said for a wait-and-see attitude (Orme, 2004).

2 A resolved or unresolved player maneuvers himself into an observable situation such that the question about his commitment becomes irrelevant (burning bridges);

3 An unresolved player restructures his incentives and commits himself (tying hands).

The first case assumes that a real problem exists: a player is fully committed and yet has trouble convincing his opponent that this is so. This is the unenviable situation that a defender with valuation $v_1 > v_1^*$ finds himself in: he wants to threaten only if the challenger is not resolved but ends up running a substantial risk of war. This is a shared problem too because the unresolved challenger with valuation $v_2 \in (\hat{v}_2, v_2^*]$ ends up fighting rather than capitulating, and the resolved one with $v_2 > v_2^*$ has to go to war rather than enjoy successful revision of the status quo in peace. Communicating existing commitments is imperative for these players and yet the model suggests it will be fraught will difficulties. As we shall see shortly, a lot of scholarly energy and imagination has gone into finding ways in which this can be done.

Unlike the first situation that deals with the problems of resolved actors, the second case applies to unresolved ones as well. The idea is that eliminating a tempting option may put an actor in a strong bargaining position regardless of the prior credibility of his commitments. If he is unable to exercise that option, then it is irrelevant. For example, suppose that S_2 somehow managed to make capitulation impossible. The game then degenerates into a contest between S_1 with himself: escalation would immediately lead to the final attack choice, and because Assumption 2.1 implies that whatever he chooses there is worse than the appeasement, he relinquishes the good immediately. Uncertainty becomes irrelevant in this context as long as S_1 can observe S_2 eliminating her capitulation option. If S_1 is unaware of that action, uncertainty becomes fatal for the defender: a resolved type will cause war with his threats, and a bluffer will shoot himself in the foot because there is no chance of obtaining concessions. Obviously, it is in S_2's interest to reveal her action, but because saying she has done it would be advantageous regardless of her valuation and irrespective of whether it is true or not, the incentives to lie raise their ugly head again.

Because eliminating options altogether can be such a risky tactic under uncertainty, the third case presents a less extreme alternative. Bargaining has at least two roles: the informational we have already discussed, and the functional, which consists of attempts to rearrange one's incentives such that certain courses of action become more or less attractive (O'Neill, 1991). In our model such a move would consist of manipulating one's payoffs from fighting, capitulation, or both. Suppose, for example, that S_2 somehow manages to double the reputational costs, so capitulation would now yield $-2a_1$. The predictable impact is to decrease v_2^*, thereby increasing the set of valuations for which she would be fully committed to resist. Threats now become riskier for S_1 and this decreases (and may eliminate altogether) the proportion of bluffers in the mix, improving S_2's expected crisis payoff. Without changing her fundamental preferences (valuation of the good), S_2 has succeeded in restructuring the strategic environment to her own advantage.

It is this functional role of crisis behavior that many studies have tended to neglect, but which is fundamental for any theory about the coercive uses of force. As it turns out, the military instrument has both signaling and incentive-rearranging features, and so it is a mix between the first and third tactics. Before we can consider the implications of this simple

observation, however, we must look briefly at the mechanisms underlying the three ideal-type solutions to the credibility problem.

2.3.1 Costly Signaling

The discussion so far suggests an obvious approach to solving the credibility problem "caused" by uncertainty: the committed actor should look for a way to reveal his resolve. This focuses the attention on credible communication; that is, making the opponent believe one's statements when he does not know the resolve of the actor making them. Indeed, this is the path that most research has taken, and with fascinating results.

In his influential book, Jervis (1970) studies *signals*, which do not change the distribution of power, and *indices*, which are either impossible for the actor to manipulate (and so are inherently credible) or are too costly for an actor to be willing to manipulate. In modern terms, he distinguishes between *cheap talk* and *costly signaling*, although his emphasis on psychological factors that influence credibility blurs the fundamental differences between the two.[14]

It is well-known that the possibilities for credible revelation of information when talk is cheap are rather limited and depend crucially on the degree of antagonism between the actors (Crawford and Sobel, 1982). In our setting, the opponents share a preference for avoiding war but they diverge significantly on the terms of an acceptable agreement. Schelling (1960) calls this a "mixed-motive situation," where truth-telling would be scarce, and suggests exploring tacit communication through actions instead of words.

Sunk Costs

One action is to *sink costs*, that is, incur expenses that one has to pay regardless of the crisis outcome (Spence, 1973). Often referred to as "burning money," this signaling mechanism relies on deadweight losses whose role is purely informational. As an illustration, suppose you wanted to convince your opponent that you valuation exceeds some threshold $\underline{v}$. If you burned resources whose value total $\underline{v}$, then you have effectively communicated that fact. Only an actor whose valuation exceeds that amount would have been willing to pay these costs because he would still remain in the black. This act would be irrational for anyone who values the issue less because it involves losing more than giving up the good. Therefore, resolved actors can burn money because their willingness to bear these costs separates them from the unresolved types and becomes a credible revelation of their commitment.

Consider an extension of the simple crisis game that allows the defender to choose the level of escalation, $m \geq 0$, with $m = 0$ being equivalent to appeasement. Assume that the costs of escalating are increasing in the level, and that they are incurred immediately; that is, they are sunk. The modified version of the game is depicted in Figure 2.3.

Setting $a_i = 0$ makes this game equivalent to the model Fearon (1997) analyzes. Since his main results under two-sided incomplete information hold for $a_i > 0$ as well, it is sufficient to summarize them here. Fearon finds that multiple equilibria are possible but they all take

[14] Another interpretation of the two categories would reduce the distinction to whether the characteristic is at all manipulable or not. This is what Spence (1973) does in his own seminal article, giving a worker's educational choice as an example of a signal, and the worker's sex as an example of an index. However, it seems to me that Jervis is quite explicit in that indices *can* be manipulated, but that actors may not be *willing* to do it.

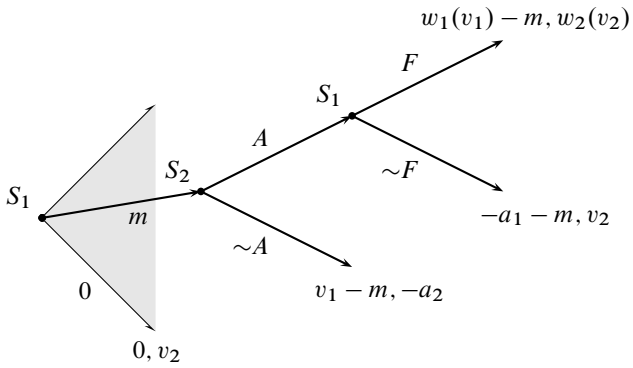


Figure 2.3 The Crisis Game with Sunk Costs

the same form. Despite the rich array of signaling options, only one signal is ever sent in equilibrium; that is, if escalation occurs, it is at one particular level, m^*. All $v_1 \geq \hat{v}_1$ threaten with m^*, and all others appease. As in the original model, only $v_1 > v_1^* \geq \hat{v}_1$ actually fight if resisted. S_2 resists if $m \neq m^*$ or if $v_2 > v_2^*$ and capitulates otherwise. Although multiple equilibria with this structure are possible, all but one of them are eliminated by an intuitive forward induction logic.

PROPOSITION 2.3 (Sunk Costs). *The escalation game with sunk costs has a unique intuitive equilibrium provided the defender's maximum valuation is sufficiently high. The defender threatens with a uniquely costly signal only if he has a credible commitment, and the challenger resists only if she has a credible commitment. The threat is sufficiently costly to make escalation unattractive to an unresolved defender.*

To see the logic behind this result, consider the basic crisis game we analyzed in Section 2.1 under uncertainty. Recall the scenario illustrated in Figure 2.2(b) where bluffing is possible. In that case, S_1 threatens if his valuation exceeds $\hat{v}_1$ even though $\overline{v}_1 < v_1^*$; i.e., he does not have a credible commitment to fight. By construction, $\hat{v}_1$ is such that a defender with that valuation is indifferent between appeasement and escalation, which means that his payoff from escalation is zero. Consider now some $v_1 \in (\hat{v}_1, v_1^*)$: a bluffer whose expected payoff from threatening is strictly positive. Suppose now escalation involved a sunk cost $m > 0$. If S_1 threatens with m, escalation will have two effects. First, it will uniformly reduce his expected payoff from escalation. In particular, this means that $\hat{v}_1$ will no longer be willing to threaten because his expected payoff from doing so is now strictly negative. But this reveals the second effect: S_1 is now less likely to be bluffing, which means that S_2 is less likely to resist. This increases the expected payoff from escalation and makes it more attractive than before. However, it turns out that the first effect is stronger than the second, which means that in general sinking costs will reduce the probability of bluffing in equilibrium.

The trouble with these strategies is that they can support a great many choices for the valuation at which S_1 is willing to escalate: one picks such a value and then derives the optimal signal and corresponding beliefs to make it work. As Fearon (1997, 76-77) argues, all but one of these values are unreasonable if one applies the Intuitive Criterion (Cho and Kreps, 1987). According to this logic, an equilibrium is *unintuitive* if (a) there exists a type

that could profit by deviating from the equilibrium strategy if that would cause the other player to change his belief and behavior, and (b) other types cannot benefit from changing their behavior even if it would lead to the same results.

This argument eliminates all equilibria except the one in which threats are genuine: $\hat{v}_1 = v_1^*$. To see why, note first that Bayes rule does not specify how S_2 should update her beliefs if she observes an unexpected escalation level $m > m^*$. This is a zero-probability event, and the theory allows the analyst to prescribe beliefs for any such event. To sustain these equilibria, we require that if S_2 observes such a high level, she infers that S_1 is unresolved. This is an odd inference indeed because if S_1's valuation is v_1, paying any $m > v_1$ is strictly dominated by living with appeasement: even if S_2 capitulated for sure (the best possible outcome), his payoff would be $v_1 - m < 0$. This now means that if S_1 sinks just enough costs to make the smallest-valuation type among the resolved ones, v_1^*, indifferent between appeasement and escalation, then S_2 should conclude that the signal could not have possibly been sent by any $v_1 < v_1^*$. That is, the probability that S_1 is bluffing if he escalates must be zero. This, in turn, implies that only resolved challengers would resist, and so $\hat{v}_2 = v_2^*$. Fearon proves that if his valuation is high enough, S_1 could profitably deviate to such a strategy as long as $\hat{v}_1 < v_1^*$, and hence all such equilibria are unintuitive.

In other words, as long as $\overline{v}_1$ is high enough, the Intuitive Criterion eliminates all bluffing equilibria analogous to the ones identified in Figure 2.2(b), and leaves only the analogue to the genuine threat equilibrium in Figure 2.2(a). However, if $\overline{v}_1$ is not high enough, which may be the case because of the restrictions implied by Assumption 2.1, then the forward induction logic will not be able to eliminate bluffing completely. This is because the maximum valuation of the defender may not be high enough to make the intuitive deviation worthwhile even for him. In that case, we would eliminate all equilibria except the one where $\hat{v}_1 < v_1^*$ is chosen such that $\overline{v}_1$ has no incentive to deviate. That is, the unique equilibrium will involve bluffing. It appears that in terms of strategies, sinking costs is not all that different from simple escalation.

The immediate question that arises is why this is the case: why doesn't S_1 make use of the signaling device that allows for an infinite variety of escalatory actions? To answer this, we must ask what it is that S_1 is supposed to be signaling. Although the uncertainty is about his valuation, this is not what needs to be revealed to S_2, at least not precisely. What S_2 needs to know is whether S_1 would actually fight when resisted, that is, she is only interested in whether his valuation is high enough to meet the commitment threshold. As we have seen, if the largest valuation is sufficiently high, then the resolved S_1 can choose a signal m^* that is just costly enough to make mimicking by unresolved ones unprofitable, and thereby convince S_2 that she faces a genuine threat. Obviously, any $m > m^*$ will just be a waste at this point—there is nothing to be gained by it because S_2 will conclude that S_1's commitment is credible with just m^*—so the costlier signal will not be attempted. Sending a less costly signal is also unsatisfactory because S_2 will retain doubts about S_1's commitment and will therefore resist with a higher probability. As Fearon (1997, 77) aptly puts it, "signaling dynamics drive the defender to signal 'all or nothing.' "

The close correspondence between sunk-cost signaling and simple crisis escalation may seem to imply that nothing can be gained from costly signals. Clearly, the ability to sink costs will not alter the situation depicted in Figure 2.2(a) in the slightest. Recall that in this scenario, S_1 escalates only if he has a credible threat and in fact for some valuations he

appeases despite having it. There is absolutely no reason for a costly signal here: the costless threat is sufficiently dangerous to deter all but the most resolved types from making it.

If, on the other hand the scenario is as in Figure 2.2(b), then the intuitive signal is lowest cost that makes the defender with a valuation at the credibility threshold indifferent between appeasement and escalation. Since sending this signal will make any unresolved defender strictly worse off than appeasement, making such a threat will be sufficient to convince S_2 of his resolve. This implies that if S_1 is able to send a signal that is costly enough, not only will the chance of appeasement increase (because only resolved types threaten), but the risk of war will decrease as well (because only resolved challengers stand firm). The ability to burn money unequivocally helps strong defenders and hurts the unresolved.

When $\overline{v}_1$ is not sufficiently large, the Intuitive Criterion cannot eliminate the bluffing equilibrium. Efficiency concerns would lead S_1 to choose the least costly signal in this case, and this would be m^* such that $\overline{v}_1$ is just indifferent between deviating in the manner prescribed above and sticking to his equilibrium strategy. Since $\hat{v}_1$ is increasing in m^*, the effect is to choose the smallest possible range of bluffers, and so the probability of bluffing with a sunk-cost signal will be strictly lower than the probability of bluffing without one. Since sunk-cost threats are more likely to be genuine, the minimum resolve that a challenger must have to resist will be higher, and so the probability of war will be lower. Even when bluffing is a distinct possibility, sinking costs can improve matters for a resolved defender.

Burning money does not affect the ranking of an actor's preferences over various outcomes. Since S_1 pays these costs regardless of whether he attacks when resisted or not, they are irrelevant for his ultimate decision, which still turns on the credibility constraint in (CR_1). Therefore, sunk costs cannot work as a commitment device and their function is purely informational. Except under conditions in which costless escalation is sufficiently risky, costly signals are worthwhile for resolved defenders and reduce the probability of war. Furthermore, when both cheap talk and costly messages are available, costly signals can improve the precision of communication (Austen-Smith and Banks, 2000). *Whereas sinking costs does not alter one's commitments, it may reveal their credibility.*

It may be difficult to find empirical examples that involve pure sunk costs. One possibility is the installation of nuclear missiles in Western Europe during the Cold War. Under the conditions of mutually assured destruction (MAD), these missiles would have no bearing on the likely outcome of an all-out nuclear war. Still, placing them in Europe and maintaining them there was quite costly to the United States and as such may have been done primarily for purposes of revealing the American commitment to European defense. Whether this was done more for the benefit of the Soviets or to reassure the European allies, is unclear.[15]

Threats That Generate Risk

Thus far, credibility turned on the difference between the expected payoffs from backing down and fighting. For S_1, this essentially reduced the calculation to his penultimate choice. He was said to have a credible commitment if war was preferable to capitulation. This very comparison, however, became extremely problematic when it came to global nuclear war between the two superpowers after the late 1960s, when the Soviet Union acquired second-strike capability. During the era of Mutually Assured Destruction (MAD), each side could

[15] I thank Barry O'Neill for suggesting this example. Personal communication.

absorb a surprise nuclear attack and then deliver a devastating counter-stroke. This retaliatory capability meant that whoever started the thermonuclear war, its end would always be the same: both sides utterly destroyed. In this context, a deliberate choice to attack became irrational.[16]

In terms of our model, this meant that (CR_1) is never satisfied, even with incomplete information. As we know, in this case compellence is certain to fail: S_2 always resists regardless of what she believes about her opponent because he will always capitulate. But when resistance is certain, then S_1 never issues threats because of Assumption 2.1. Uncertainty appears to be irrelevant.

As it turns out, however, this is not necessarily the case. The problem with escalation is that in case of resistance it leaves S_1 with a stark choice that he can never credibly threaten to make. But what if he could threaten with the risk that things may get out of hand, that war would come by accident, that mutual destruction could happen in spite of her efforts to prevent it? Schelling (1960) was the first to suggest the *threat that leaves something to chance*: the idea that whenever one cannot threaten to execute something that is too painful for him, he could still threaten with it happening anyway. To wit, if an actor could diminish the chance that he would be able to take the way out, then he could potentially threaten with an action that he would never rationally take.

One analogy is rocking the boat in the open sea when neither of the two people in it can swim. If the boat capsizes, then both drown, a disaster akin to mutually assured destruction. Of course, neither player can credibly threaten the other to capsize the boat on purpose. However, a player could start rocking the boat, increasing the chance that at some point it will destabilize enough and capsize anyway and neither would be able to stop it. The question that the other player must now answer is: How much risk is my opponent willing to tolerate, and am I prepared to run such a risk? As Schelling (1966, 103) characterizes this "competition in risk-taking,"

> In this way uncertainty imports tactics of intimidation into the game. One can incur a moderate probability of disaster, sharing it with his adversary, as a deterrent or compellent device, where one could not take, or persuasively threaten to take, a deliberate last clear step into certain disaster.

What does it mean for players not to be fully in control of events? In our analogy, an unexpectedly high wave could capsize the rocking boat, so rocking will be mostly safe although there is a chance that it could turn into disaster. Players have no control over the probability of such a wave arriving, and hence the risk of disaster is "autonomous" (Snyder and Diesing, 1977, 210).[17] Rocking the boat more vigorously increases the risk of capsizing because even moderately high waves could now cause it. By varying the vigor of rocking (the degree of escalation), players can manipulate even the autonomous risk.

For a less whimsical example of threats that leave something to chance, consider the Soviet fighter planes "buzzing" the transport corridor during the Berlin Airlift in 1948–49. Whereas shooting at the American and British planes would have been too provocative

[16] Brodie (1959); Kahn (1960, 1965), and Schelling (1966). Powell (1990) discusses the issues in the context of a series of formal models.

[17] See Schelling (1960, 188) for an explicit recognition that the risk must be beyond the collective control of the players, and Powell (1990, 16-20) for an extensive discussion of this issue and a critique of its empirical plausibility, a point to which I will return.

and escalation, putting fighters in the flight path of these transports increased the risk of collision but maintained plausible deniability. The Soviets could exercise pressure with the threat to down transports even when they would not willingly actually do it. The threat was not negligible: in April a British plane crashed after getting "buzzed", killing everyone on board (the Soviet plane also went down).[18]

One need not be fully aware of the degree of risk his actions are generating to put coercive pressure on his opponent. During the Cuban Missile Crisis in 1962, US President Kennedy thought he was carefully managing risk by opting for the less aggressive response (blockade) and then restraining advisors who preferred an early military response. In reality, however, the US administration was unaware of the actual risks it was running. Contrary to its intelligence that estimated that none of the missiles in Cuba were yet operational, some were. Unbeknownst to the administration, the Russians had also sent 42,000 combat troops with order to use the nuclear weapons in the event of an American invasion. There were also tactical nuclear weapons that were under the command of General Gribkov who could have used them had the US attempted a direct attack, not to mention the local commander General Pliyev who could have launched the nuclear-tipped FROG missiles that were under his personal command. All of this was known to Khrushchev, which meant that Kennedy unwittingly manipulated the risks of the crisis and managed to put much more coercive pressure on his opponent than he could have had he known the true state of affairs. Khrushchev was exceedingly fearful of the potential loss of control should unauthorized or accidental shooting begin (Frankel, 2005, 180). There was a distinct possibility of loss of communications once US direct action began, which could only have aggravated the problem. When Khrushchev found out about the American invasion plans (ironically, this intelligence report was actually incorrect), he "dropped a load in his pants" and reversed his brinkmanship tactics.[19] What to the White House appeared as reasonable and controlled escalation, was in reality a giant step toward nuclear holocaust in the eyes of the Russians. No wonder, then, that Khrushchev wrote the famous rambling first letter: from his perspective the Americans were racing toward a global war completely unaware of the risks they were taking, and he had no way to tell them about it (and if he did, they would not have believed him). In this

[18] Schelling (1966, 104, fn. 3) gives that example. In this particular instance, however, the tactic did not work: the British started sending escort fighters with orders to protect the transports by shooting down everything that appeared to threaten their safety. Continuing the "buzzing" tactic would inevitably result in the downing of a Soviet plane that would shift the escalatory onus back on the Soviets: they would have to retaliate. In the event, they stopped buzzing through the corridor.

[19] Cited by Taubman (2003, 568). The Russians also had to deal with Castro who urged a first-strike. Just when Khrushchev ordered that no US planes were to be shot at, Castro ordered the Cubans to open fire on them. This is in addition to the myriad accidental "escalations" that occurred during the crisis—from the unauthorized downing of the U-2 spy plane by a Soviet commander, and the straying of American aircraft into Soviet airspace, to the CIA sabotage operations in Cuba which the White House did not know about or forgot to cancel. There were more, potentially highly explosive, coincidences. For instance, Operation Mongoose (tacitly approved by Kennedy in March) was a covert action program for the possible invasion of Cuba and overthrow of Castro. Ironically, it was scheduled for October and had been infiltrated by Cuban intelligence who knew about the tentative plans. When the U-2 was shot down, Kennedy—who had received reassurances that Khrushchev was fully in control of the situation—interpreted it as the Soviets eliminating the reconnaissance so that they could complete the installations (Frankel, 2005, 148). Khrushchev himself heard a report designed to mislead him into believing that it was the Cubans who had shot down the plane (158). On the tactical nukes, see Gribkov and Smith (1993, 4, 27-28).

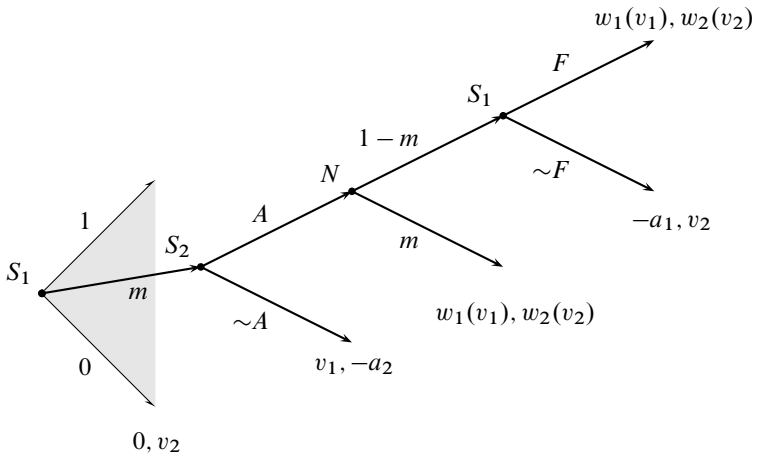


Figure 2.4 The Crisis Game with Randomized Threats

instance, the lack of information—and the fact that Khrushchev knew about it—worked to the advantage of the US administration: it was able to run up the risk of war far too high for the Russians to handle. Much higher, in fact, that it would have been willing to had it known about the details. The crisis would not have ended in such a lopsided victory then.

The formal analysis of randomized threats supports the basic insight that such threats can be used even in a MAD context, albeit with qualifications about their relationship to resolve. For instance, it is not true that the actor willing to run the highest risk under MAD would prevail (Powell, 1990).

Although the notion of threat that leaves something to chance arose in the context of mutually assured destruction where neither side would rationally choose to start a nuclear war, it is a much more general concept. Randomized threats can be used in non-nuclear crises where it is possible that the credibility constraint is satisfied. To see this, let us modify the basic escalation game to allow S_1's threat to involve an autonomous risk of war.

Figure 2.4 illustrates the new game. S_1 begins by choosing a level of escalation, which in turn determines the probability that war will come if S_2 resists and before S_1 could make the final rational move. To simplify matters, I assume that S_1 can calibrate his escalation to pick the probability of breakdown even though in reality such a precise manipulation of risk is implausible. However, as the boat analogy shows, one may vary the risk of disaster even when it is caused by an autonomous event over which players have no control. As before, both players are uncertain about each other's valuation.

Nothing changes for S_1's final calculation: he still attacks deliberately if, and only if, he has a credible commitment. From S_2's perspective, however, resistance is now riskier because it takes into account that it may lead to inadvertent war no matter what S_1's valuation is. The expected payoff is standing firm now is $mw_2(v_2) + (1-m)\big[G_1(v_1^*)v_2 + (1 - G_1(v_1^*))w_2(v_2)\big]$. Therefore, S_2 will resist if, and only if, this payoff exceeds $-a_2$. Hence, the largest valuation at which S_2 will capitulate is:

$$\hat{v}_2 = \frac{[1-(1-m)G_1(v_1^*)]c_2 - a_2}{1-p[1-(1-m)G_1(v_1^*)]}. \tag{2.4}$$

Setting $m = 0$ reduces this to (2.2), the threshold in the basic crisis game. As before, we need to consider the two basic scenarios shown in Figure 2.2. First, suppose that the situation is as in Figure 2.2(a) and that S_1's escalation signals a credible threat to fight. This means that S_2 will resist only if her credibility constraint is satisfied as well for doing so means certain war. The presence of a randomize threat changes nothing in that decision: either resistance automatically leads to war or it leads to S_1's decision which is guaranteed to be for war as well. Any type that escalates here can pick any risk, including no risk, and it will not affect equilibrium behavior. This is analogous to the dynamic in the sunk-cost game in the sense that under these conditions, the threat itself is sufficient to reveal S_1's commitment.

Consider now the scenario in Figure 2.2(b) where it might be possible for S_1 to escalate even if he is not resolved. The proposition that follows establishes the main result: when S_1 can make threats that generate risk, he will not make partial commitments. Instead, he will opt for an action that irrevocably commits him to war even if he would have liked to have the chance to back down if S_2 happens to resist his threat.

PROPOSITION 2.4 (Randomized Threat). *In the the escalation game with randomized threats, the defender threatens with an action that will automatically lead to war if the challenger resists even though he might not fight if he had the final choice. The risk of war is positive if the defender escalates and he might end up fighting even thought he would have preferred to capitulate when resisted.*

The logic behind this result is as follows. If the defender is resolved, then it does not matter to him whether S_2's resistance causes war by chance or through his own action. This type of defender will therefore maximize the probability that S_2 capitulates, which he can do by burning the bridge (i.e., choosing $m = 1$). This now implies that any choice of a less risky threat with $m < 1$ unambiguously reveals that the defender is unresolved. Unlike the basic and sunk-cost games, however, this does not mean that S_2 will resist for sure; after all, there is a positive probability that war might occur by chance. However, as Appendix A shows, the benefit from increasing the probability of her capitulation offsets the cost of a larger risk of war if she does not. This means that some, but not all, unresolved types would prefer to commit themselves fully. This, in turn, implies that S_2 will resist only if she has a credible commitment as well: even though she knows that her opponent might be bluffing, this is irrelevant for war is certain if she stands firm regardless of what S_1 might do in the endgame.

In contrast to the sunk-cost game, S_1 can threaten by burning a bridge even if he is unresolved. In contrast to the basic game, this threat will have the same effect on S_2 as if she believed him to be resolved. The ability to employ threats that leave something to chance restores the possibility for bluffing in equilibrium. From the perspective of an unresolved defender who values the issue sufficiently to make this threat, this ability unequivocally improves his position compared to sinking costs. Recall that intuitive sunk-costs equilibria do not permit bluffing, and therefore all these types will be deterred for sure, which is not the case here.

When it comes to crisis stability, however, randomized threats are definitely riskier than sunk-cost threats. The probability of compellence failure increases because S_1 can escalate even if he is not resolved. Even worse, so does the probability of war because if S_2 resists, which she does in both models with equal probability, war is more likely due to bluffers

not having a chance to capitulate. *The ability to run risks improves the expected payoff of some unresolved defenders, but does so at the cost of an increased probability of war.* As mentioned before, this is a result that extends to general crisis bargaining models.

Compared to nuclear deterrence under MAD, conventional deterrence where there are types who may have a credible commitment to fight changes the dynamics of randomized threats. As we have seen, instead of actually leaving anything to chance, defenders whose resolve is high enough fully commit to fighting in case of resistance. Just as in sunk-cost signaling, the escalatory move is an all or nothing proposition but unlike that case, it can involve bluffing. The defender effectively says, "I will fight whether or not I am committed because I will have no choice if you resist my demand." Because $m = 1$ makes this statement truthful, the challenger behaves as if her opponent is resolved even though she knows he might not be. Note, however, that not all types of defender will attempt this escalation: it involves creating a risk of war that $v_1 < \hat{v}_1$ are unwilling to bear. As such, randomized threats serve as costly signals.

It may appear that threats that leave something to chance run contrary to logic: after all, if the challenger resists, the unresolved defender would dearly love to be able to escape the escalatory commitment and back down. So *ex post*, he is in a worse position. However, as Schelling (1960) observed, making oneself weak may yield a bargaining leverage, as it does in this case.[20] By removing the option to capitulate, the defender makes his lack of resolve a non-issue. However, the extreme choice of $m = 1$ suggests that randomized threats are only analytically interesting in an environment where war is the worst outcome for players regardless of their valuation. As we shall see in the next section, constraining one's future options in this fashion is a rather strong commitment device.

Before we turn to that analysis, it is worth noting that there are strong substantive objections to randomized threats. Fearon (1994a, 579) and Schultz (2001, 43) both argue that treating the onset of war as a stochastic event over which actors have no control is somewhat unsatisfying. As many researches have repeatedly noted, this tactic is not only empirically rare, it is made more implausible by the fact that leaders often pursue strategies designed to minimize the shared risk of disaster, not increase it.[21]

2.3.2 Constraining Future Choices

The results from the randomized threat game already suggest that rendering oneself unable to exercise a tempting choice may be a useful commitment device. The classic example is "burning the bridge": if one cuts off his escape route, then one cannot retreat no matter how hard he wishes to, which means that he will have to fight even if he preferred flight

[20] He refers to this as "relinquishing initiative" (Schelling, 1966, 43-49). In our equilibrium, by escalating with $m = 1$, S_1 effectively relinquishes the choice to go to war to S_2, and is able to extract advantages from it.

[21] See, for example Young (1968) and Snyder and Diesing (1977) for general critiques, and Trachtenberg (1991) for a rather devastating skewering of the archetypal interpretation of July 1914 as a crisis in which leaders lost control of events and ended up in a war nobody wanted. McClelland (1961, 202) even goes as far as suggesting that leaders should exhibit "anxious attention to control problems, even to the extent of hidden collaboration on the part of the principal actors during the early phases of a crisis." Zagare and Kilgour (2000, 54-57) further argue that requiring chance to impose an outcome that an actor would not rationally choose does not solve the credibility problem. This is not an issue in this model, however, because resolved types do prefer fighting to backing down.

to a fight (Schelling, 1956; Ellsberg, 1975). Burning the bridge saddles the opponent with the unenviable task of choosing between certain war and backing down. It eliminates the temptation to flee, and makes the commitment to war credible even for unresolved types.

There is no shortage of examples of this tactic. When Hernán Cortés' plan to capture or kill Montezuma became generally known, there was defeatist talk of returning to Cuba—the enterprise seemed too risky and daunting. After foiling a conspiracy to kill him, Cortés ordered nine of the twelve ships anchored off Villa Rica to be sailed aground and stripped of all the rigging, sails, and guns. His excuse to the men was that the vessels had been rendered unseaworthy by the wood-beetle, but as many at the time recognized, without this drastic action he would have been unable to muster enough men for his expedition into the heart of Mexico. Cortés himself admitted as much later when he said that after the grounding the expedition "had nothing to rely on, apart from their own hands, and the assurance that they would conquer and win the land, or die in the attempt."[22] As Prescott (2001, 267) concludes, "The destruction of his fleet by Cortés is, perhaps, the most remarkable passage in the life of this remarkable man." (Incidentally, the famous "burning of the ships" never took place—they were beached.)

During the tumultuous twelfth century, the Minamoto clan was struggling to gain supremacy over the powerful Taira in Japan. In 1185, Minamoto no Yoshitsune engaged the Taira off the coast of Shikoku island. Prior to what became known as the Battle of Yashima, Yoshitsune's ally Kajiwara opined that it might be prudent to fit their boats with "reverse oars" so that they could maneuver more easily. Yoshitsune, who interpreted the boats' agility as facilitating desertion, vehemently objected:

> A soldier enters battle with the intention of never retreating. It is only after things have gone badly that he [even] thinks of turning back. What good can come from preparing one's retreat in advance? Your Lordships may fit these 'reverse oars' or 'turn-back oars' to your ships by the hundreds or thousands as you please. I myself am quite satisfied with the ordinary oars that have been used in the past.[23]

During the Second World War, Japanese *kamikaze* pilots took fuel that would only be enough to reach the American ships: even if one's nerve failed at the last moment, the alternative to ramming the enemy was drowning futilely in the open sea.

As obvious as this tactic may seem, sometimes armies get it so completely wrong that the effort is more farcical than heroic. During the Trojan War, the Trojans repeatedly tried to burn the ships of the invading Greeks. They failed but had they succeeded, they would have committed the horde to seeing the siege through. While the ships survived, escape remained an option: after Hector's fierce assault that breached the protective rampart, Agamemnon himself lost heart and ordered the ships closest to the shoreline hauled down and rowed into the sea and then, should the Trojans refrain from fighting during the night, using the cover of darkness to evacuate the remaining ships for there was "no shame in running, fleeing disaster, even in pitch darkness." Odysseus, who understood perfectly well that the retreat would turn into a rout, flew in anger and shouted at the king:

> Now where's your sense? With the forces poised to class

[22] Cited in Thomas (2005, 222–23). Even Emperor Charles himself would later explicitly acknowledge that the beaching of the ships was instrumental in forcing enough men to follow Cortés to Tenochtitlan (see fn. 82 in Thomas (2005, 223)).

[23] Morris (1975, 83).

> you tell us to haul our oar-swept vessels out to sea?
> [. . .]
> Achaean troops will never hold the line, I tell you,
> not while the long ships are being hauled to sea.
> They'll look left and right—where can they run?—
> and fling their lust for battle to the winds.[24]

The Trojans would have been better advised to help the Greeks build more comfortable "retreat" vessels rather than burn them!

One should be careful not to burn the bridge for one's opponent and thus leave him no choice but to fight to the death. As the 10,000 Greeks were making their way through the Persian empire during their abortive attempt to help the pretender Cyrus seize the throne, they had to fight numerous battles with hostile locals. During one operation in the mountains, the Greeks had to dislodge the Carduchi tribe. Xenophon, in command of the baggage train, was making his way to link up with the rest of the Greeks when he ran across the enemy occupying a ridge. The hoplites attacked the Carduchi taking care that "They did not attack from every direction but left the enemy a way of escape, if he wanted to run away" (Xenophon, 1950, 182-83). The advantage of committing oneself while ensuring that the enemy is not similarly committed is well-known to battle tacticians. It is for precisely this reason that Xenophon (1950, 286-7) counseled as follows:

> I should like the enemy to think it is easy going in every direction for him to retreat; but we ought to learn from the very position in which we are placed that there is no safety for us except in victory.

As any rule, this one has its notable exceptions. When most of the combat is hand-to-hand, a tightly-pressed encircled army may not be able to take advantage of its numbers with only soldiers along the perimeter being able to engage the enemy. In such cases, utter annihilation is a distinct possibility, as the Romans found out when Hannibal trapped them at Cannae. Cutting off the means of escape also means cutting off the means of supply, which usually renders a modern army helpless as well. When the Germans attacked Stalingrad, the Russians—with their backs to the Volga and retreat additionally discouraged by the NKVD shooting deserters—held until Zhukov could complete the encirclement and spring the trap on General Paulus. The Sixth Army held out for longer than one would have expected under the circumstances but in the end it had no choice but to surrender—the *Luftwaffe* was incapable of resupplying it in the bitter winter weather (Beevor, 1998).

To investigate the possibility of this type of commitment, we can modify our basic crisis game to allow S_1 to escalate in one of two ways: he can "burn a bridge" in which case resistance by S_2 means automatic war, or he can escalate by keeping his escape valve open. Since we have already analyzed these situations separately, putting them together does not require additional work. Burning the bridge ensures that the challenger will only resist if she has a credible commitment; that is, it guarantees the largest probability of her capitulation. A resolved defender would only choose to escalate without burning the bridge if doing so yields that high probability. This can only happen if S_2 believes that the threat is genuine. But if an unresolved defender can profit from this belief and escalate, then S_2 will have no reason to think that the threat is genuine. Therefore, a resolved defender will always burn

[24] Homer (1990, 372–73).

the bridge. It now follows that he may also choose to do so if he is unresolved. The logic is analogous to the one behind the randomized threat result in Proposition 2.4.

Everything that we found about randomized threats then applies to burning bridges. This includes their substantive implausibility, albeit for slightly different reasons. Eliminating alternatives that one may be tempted to take can result in enhanced bargaining strength but requires willingness to tolerate significant risk of failure (Schelling, 1960, 178–83). After all, should the enemy misperceive or fail to notice one's commitment, the result can easily be disastrous for both. The more absolute the constraint on behavior, the more credible the commitment, but the higher the risk one must run. But this is precisely what makes burning bridges unattractive to states during a crisis. As Young (1968, Ch. 9) observes, leaders always want to maintain freedom of choice, especially with respect to the final decision for war and are very unlikely to pursue any tactic that would lock them into a position from which they cannot retreat, even if doing so would confer a bargaining advantage. After noting that they could find no unambiguous instances of this tactic in their data set, Snyder and Diesing (1977, 213–14) go so far as to claim that the reason they even discuss it, "other than logical completeness… is that it is quite prominent in the existing theoretical literature," and so they expected to find many cases where it is used.[25] Finally, Petersen's (1986) statistical analysis also finds that the results are inconsistent with leaders firmly committing to a strategy that gives their opponents the "last clear chance" to avoid war. Rather, they are consistent with analytical behavior based on assessment of expected outcomes.

Given that leaders actively search for options during crises and "consciously use analytic decision-making procedures at least some of the time" (Herek et al., 1987, 218), this failure to find instances of them using the tactic is pretty alarming. If rendering the credibility problem moot is so extreme that it is substantively implausible, then perhaps a tactic that varies the degree of one's commitment may be more attractive? We have seen that randomized threats cannot help because they produce a result equivalent to burning bridges. Some subtler rearrangement of incentives is necessary.

2.3.3 Manipulating Incentives

I began the section on costly signaling by writing about "solving the credibility problem 'caused' by uncertainty." There is a reason I put the word *caused* in quotation marks. In my view, uncertainty is not nearly as big of a problem as not having a credible commitment in the first place. Focusing on communicating resolve side-steps the rather important issue of establishing the commitment. That is, whereas the informational approach takes the strategic environment as fixed and then analyzes the effects of changing information, it neglects one simple, and yet fundamental, observation: *actors can restructure their strategic environment*. While we shall continue to maintain the assumption of preferences fixed for the duration of the crisis, we shall permit actors to rearrange their incentives with their actions.

[25] Strictly speaking, they refer to a more general tactic that they label a "committing threat" where the move creates a commitment that did not exist before. Burning bridges is the clearest example of that, but their discussion makes clear that they have something more general in mind because this commitment is created "via the engagement of additional values (honor, resolve reputation, etc.) that would be lost if the threat is not fulfilled." As I note below, this is precisely the audience costs tactic, so their empirical criticism should be taken to refer to such strategies as well.

How can players rearrange their incentives and what are the consequences of doing so? Since the ultimate choice between fighting and backing down is at the heart of the credibility problem, let us consider S_1's final decision. Recall from (CR_1) that he has a credible commitment when his expected payoff from going to war is at least as good as the payoff from capitulating. This means that if he can manipulate one or both of these payoffs, he can change his incentives to go to war; that is, he can alter his commitment. In particular, any action that increases his expected payoff from war relative to the expected payoff from capitulation can serve as a commitment device.

Consider a defender with valuation v_1 and suppose that (CR_1) is not satisfied, and so he does not have a credible commitment. Using $a_1(v_1)$ to denote the expected payoff from backing down, this means that $w_1(v_1) < a_1(v_1)$. If S_1 can increase $w_1(v_1)$ or decrease $a_1(v_1)$ until the inequality is reversed, then he will have established a commitment not to back down. Recall that sunk costs do not change the relationship between these payoffs because they are subtracted from both sides of the inequality.

Threats that leave something to chance also do not change the incentives at this point. If anything, these threats appear counter-intuitive because increasing the risk of disaster actually *increases* the payoff from one's own capitulation because this is one way to relieve that risk (the other, of course, is the opponent's capitulation). In other words, the longer the crisis persists and the higher this risk gets, the more likely is S_1 to quit. Coercive pressure comes not from rearranging incentives here but from foregoing earlier opportunities to concede. It is precisely S_1's refusal to take the ever more tempting way out that reveals the extent of his resolve. Like sinking costs, threats that leave something to chance are about transmitting information.

Burning bridges does not as much rearrange one's incentives as it eliminates options altogether. Although this extreme tactic is probably not something that leaders often, or even sometimes, contemplate, it has the kernel of the commitment process. Rather than making the commitment absolute, actors may opt for a more graduated approach. What are some of the ways that S_1 can manipulate his relative ranking of fighting and capitulation while keeping his valuation fixed? As Schelling (1958) observes, an actor can establish a credible commitment by deliberately imposing on himself costs that he would have to pay if he fails to carry out the threat. If he does that, "he reduces his own payoff incentives to break the pledge—perhaps to the point where it would become irrational to break it" (Ellsberg, 1975, 26). Scholars have proposed what amounts to numerous variants of two general ways to accomplish this.

Domestic Political Audiences

Fearon (1994a) offers a dynamic commitment model where an actor can manipulate his payoff from capitulation after escalating. Leaders who choose to continue the crisis incur ever increasing *audience costs*; that is, the longer they escalate, the costlier it is for them to back down. If they prolong the crisis sufficiently, they will become locked into positions from which neither would recede, and the inevitable outcome will be war. The basic mechanism that enables them to *tie their hands* relies on progressively decreasing the benefit of peace until at some point war becomes the more attractive option. Fearon (1997) compares tying-hands signaling to sinking costs and finds that while bluffing also never occurs in

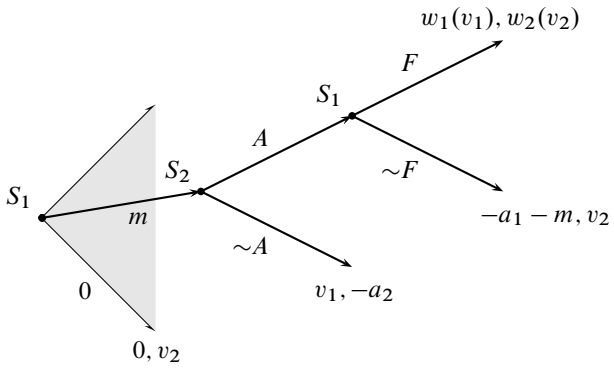


Figure 2.5 The Crisis Game with Audience Costs

equilibrium (just like the other scenario), leaders always do better by tying hands rather than sinking costs.

Before investigating the sources of these audience costs, it is worthwhile to outline the logic of Fearon's (1997) results. A slight modification of the payoffs in the costly signaling game in Figure 2.3 produces the audience-cost model in Figure 2.5. The only difference is that whereas sunk costs are paid regardless of the outcome, audience costs are paid only if S_1 fails to carry out his threat. Setting $a_i = 0$ makes this model equivalent to Fearon's, and since his analysis applies even with $a_i > 0$, I will only summarize his results.

PROPOSITION 2.5 (Tying Hands). *The escalation game with audience costs has an essentially unique equilibrium. The defender threatens with an act that incurs enough audience costs to commit him irrevocably to war, and the challenger resists only if she has a credible commitment.*

Observe first that whereas in the sunk-cost game higher levels of escalation had no bearing on S_1's commitment, they do affect it in this game. In particular, whether his credibility constraint is satisfied now depends *both* on his valuation and the chosen level of escalation. S_1 fights if, and only if, $w_1(v_1) > -a_1 - m$, or if her valuation exceeds the critical threshold that itself is a function of the escalation. This credibility threshold is defined as:

$$v_1^*(m) = \frac{c_1 - a_1 - m}{p}. \tag{2.5}$$

Comparing (2.5) to (2.1) shows that the commitment threshold with audience costs is lower, and that it is decreasing in these costs. That is, the higher the costs that S_1 can impose on himself, the lower the valuation necessary to make him willing to resist. By varying the level of escalation, S_1 can change these audience costs, and therefore *create a commitment where none existed before*. For example, take some $v_1 < v_1^*$ in the original model. This type does not have a credible commitment to attack: his valuation is $v_1^* - v_1$ short of the threshold. However, this unresolved type can close this gap by incurring audience costs of $m = p(v_1^* - v_1)$. Note that the required costs are smaller than the gap itself, reflecting the fact that increasing them lowers the threshold. If this type incurs at least m audience costs, then he has effectively tied his hands, and will go to war if resisted.

How would S_1 then escalate to influence S_2's beliefs? The most convincing escalation would only leave a resolved challenger resisting. Resolved types that are willing to escalate maximize the probability that S_2 will capitulate. Suppose they send a signal m^* that some unresolved types are willing to mimic. In equilibrium, S_2 will take into account that there is a chance S_1 might be bluffing, and will resist with a higher probability because war would not be certain. But a resolved type can always improve his payoff by incurring higher audience costs: he would not have to pay them because he never capitulates and S_2 will be more likely to capitulate herself because she knows that the higher the audience costs, the more likely is S_1 to be resolved.

Since there is no restriction on the amount of audience costs that the defender can generate, $m > 0$ can be arbitrarily large. Resolved types can increase their costs until it is no longer worthwhile for an unresolved type to mimic their behavior. Because S_2 always resists with positive probability, an unresolved type always faces the prospect of having his bluff called and ending with the capitulation payoff. If this payoff is sufficiently low, then escalation will be strictly worse than appeasement no matter how small the probability of resistance is. In other words, resolved types can generate high enough audience costs to provide this disincentive, and hence in any equilibrium the probability of bluffing will be zero.

Since $m^* > 0$, it follows that $v_1^*(m^*)$ is strictly smaller than v_1^* from (2.1). Not only are there no bluffers in this equilibrium, but the range of genuine defenders is strictly larger than the range in the sunk-costs and the basic models. Incurring audience costs has enabled some of the otherwise unresolved types to create a credible commitment and communicate it effectively to S_2. By *effective communication* I mean that only resolved challengers choose to resist: when S_2 decides what do, she is fully aware that her opponent has a credible threat to fight. In other words, this is the best the defender can do under the circumstances.

As in the intuitive sunk-cost equilibrium, S_1's signal completely resolves the uncertainty about his commitment. The logic here is straightforward: because audience costs are only paid if S_1 is forced to back down from his commitment, he can generate costs that are as high as necessary to lock himself into a position from which he would not recede. There is nothing in the model that restricts that ability, and correspondingly high-valuation types can make m^* arbitrarily high. On the other hand, there is always a chance that S_2 will resist anyway, and since concessions are not worth much to low-valuation defenders, they will not want to commit themselves.

Comparing this model with risk-generation reveals the escalation threshold remains the same whether S_1 escalates with a randomized threat or by tying hands. This is perhaps not surprising since either method commits S_1 to inevitable war. It is just that with risk-generation S_1 may not be resolved: rather than manipulating his incentives, he removes the tempting option altogether. Because this escalation is essentially costless (burning the bridge is free as are audience costs in equilibrium), the same types escalate under either scenario. For this reason, I shall treat the two models as equivalent in all subsequent discussion.

In contrast, the escalation threshold with audience costs is *lower* than in the sinking-costs model. In other words, S_1 is much less likely to escalate by sinking costs than by tying hands. The reason for this is intuitive: although in the end S_1 is just as committed to fighting, it is costly to reveal that fact and the costs must be incurred in equilibrium, which

in turn deters some lower-valuation types from escalating. The probability that S_1 will opt for appeasement is higher when his only signaling option is to sink costs.

This now allows us to evaluate the equilibrium risk of war. In all three models, S_2 stands firm whenever she has a credible commitment. Since S_1 never capitulates after threatening, the probability of war conditional on escalation is the same: it is the probability that S_2 is resolved. The overall risk of war therefore is the joint probability that S_1 escalates and S_2 resists. The second quantity is constant across the models, which means that the difference in the likelihood of war turns entirely on the different probabilities of escalation. The higher probability under tying hands translates into higher risk of war as well. In other words, sinking costs is the safer signaling mechanism, as Fearon (1997) finds as well.

The audience cost mechanism has become quite prominent in theoretical and empirical work.[26] However, despite its popularity, it is silent on two critical issues: how does a leader go about generating these costs, and how does he signal their magnitude to his opponent? We have had limited success accounting for the mechanism's microfoundations; that is, the domestic politics that would generate these costs (Smith, 1998b). For example, why would citizens punish a leader who manages to avoid war (Schultz, 1999)? One possible answer is that they worry about the country's international reputation (Guisinger and Smith, 2002), but that requires an unstated assumption that citizens care more about it that the leader does (Slantchev, 2006).

As Fearon (1997, 80) emphasizes, the mechanism requires the demanding assumption that leaders are able to generate sufficiently high levels of expected audience costs; so high, in fact, that peace becomes worse than war. When one talks about things as vague and as amorphous as "national honor" and compares them to the blood, the destruction of lives and property, and the psychological scars a war invariably inflicts on its participants, this assumption becomes heroic indeed.[27]

Still, it may not be that hard to find suggestive examples of audience costs from history. In 1494, Charles VIII, the King of France, decided to press the Angevin claim to the thrones of Naples and Jerusalem, and in September invaded Italy. When he sent envoys to Florence to obtain permission for the French army's march through Tuscany, Piero de' Medici decided that the city-state should remain neutral. Even worse, he galvanized opposition to the Angevin cause and dispatched mercenaries to the borders to prevent the king's entry into Tuscany. His zeal was not, however, matched by the influential Florentines, some of whom even tried to enter into secret negotiations with Charles to remove the Medici from power. Within a month, Piero discovered that he was virtually alone in his opposition, with help forthcoming neither from within the city nor from without, for neither the Papacy nor Venice would stand in the way of the immense French army. With the King of Naples already suffering a defeat in Romagna, Piero abandoned his stand and went secretly to Charles hoping

[26] Scholars have used audience cost arguments to explain debt repayment (Schultz and Weingast, 2003), peaceful conflict resolution (Lipson, 2003, Ch. 1), alliance reliability (Gaubatz, 1996), economic sanctions (Dorussen and Mo, 2001), trade agreement compliance (Mansfield et al., 2002), international cooperation (Leeds, 1999), monetary credibility (Lohmann, 2003), crisis escalation (Eyerman and Hart, Jr., 1996), and militarized dispute outcomes (Palmer and Partell, 1999). Gowa (2001) studies the supposed variation of audience costs across regime types but has doubts about the unstated assumptions about the electoral process that underlie audience costs models.

[27] O'Neill (1999) will probably disagree with this brusque dismissal of the value of national honor.

to get him to spare the city through a voluntary submission. Piero capitulated to all of the king's demands but when he returned to Florence, the *Priori* refused to admit him through the main gate of the palace and soon crowds gathered around, shouting insults of his betrayal and even throwing stones. The frightened Piero fled the city with his family, and the *Signoria* formally banished the Medici from Florence in perpetuity, offered a huge reward for Piero's head, and plundered the rich family holdings. Then they submitted to Charles anyway (Hibbert, 1974, 182-88). If this is a rather serious example of an incurred audience cost, it does demonstrate the danger of inflaming public opinion to follow a bellicose course of action and then folding abjectly when one's bluff is called.

On the other hand, one frequently hears the importance of Kennedy's public threats in the Cuban Missile Crisis. But as Frankel (2005) and Gaddis (1998) have concluded, Kennedy was probably quite prepared to back down if that proved to be the only way to avoid war. Gaddis (1998) is explicit: "far from placing the nation and the world at risk to protect his own reputation for toughness, [Kennedy] would probably have backed down, in public if necessary, whatever the domestic political damage might have been." Of course "whatever the damage might have been," Kennedy certainly did not expect it to result in loss of his own life or property.

Sometimes, one may not have the luxury of backing down because one's audience costs may provoke the opponent into attacking. If one manages to commit oneself to a position unacceptable to the opponent then the outcome may not be the concessions one hopes for, it may well be a war one really wished to avoid. Something like this happened during the border dispute between India in China from 1958 to 1962. We shall discuss this case in some detail in Chapter 5. For now I just wish to draw attention to the effect of Nehru's policy. After he was forced to disclose his correspondence with the Chinese to Parliament in 1959, his options rapidly decreased. In fact, by late fall of that year he had effectively committed his country not to negotiate with the Chinese over the territories. The public outcry to what he had framed as Chinese intrusion into clearly Indian territory—a matter of national honor—and his repeated reassurances that the Indian Army was ready to throw them out by force if necessary (both were false) ended up saddling him with charges of pursuing a policy of appeasement when he did not pursue a more vigorous policy. The Opposition, quite reasonably, wanted to know why the government is not ejecting the Chinese from the disputed lands, an obligation implicit in its definition of the issue as a matter of the territorial integrity of India. The criticism intensified after the public learned that Nehru had invited Zhou Enlai for talks in the spring of 1960 right after reiterating that he would not negotiate. This involved some inspired casuistry on Nehru's part in which he managed to square the circle by arguing that talks and negotiations are very different things. Observers like Walter Lippmann commented on the pernicious effect of publicizing the dispute and how it made a peaceful resolution impossible. Nehru had succeeded in committing himself to the most expansive demands possible. The problem was that he also knew India could not fight China to obtain them. That the Chinese believed his commitment was clear: they finally gave up hope of getting India to negotiate—on October 20, they attacked and imposed by force what they had proposed to negotiate peacefully.[28]

[28] See Maxwell (1970) on the public storm after Nehru's invitation to Zhou (150) and after the letter in the summer of 1962 (244).

Audience costs can also work in exactly the opposite way: they could punish escalation and prompt the leadership into ill-advised conciliatory behavior that will undermine its coercive strategy. For instance, when public opinion polls revealed that a majority of Americans opposed the doubling of the forces in Saudi Arabia in the fall 1990 (an escalation designed to deal with the failure of the sanctions regime to evict Saddam Hussein from Kuwait), President Bush made a dramatic public offer to for direct talks between the United States and Iraq. The represented a drastic departure from previous policy that had consistently refused to engage Hussein without a withdrawal of his forces, and was perceived as a major concession by the Iraqi dictator who believed it revealed the lack of resolve to resort to force that he was counting on (Freedman and Karsh, 1993, 235–37). In this instance, instead of committing the administration to the use of force, escalation's audience costs compelled it to take steps to defuze the crisis or, failing that, demonstrate unequivocally that war was indeed the last resort. Indeed,

> the reaction to Bush's 'extra mile' initiative was far warmer among Saddam's allies and the non-aligned than it was among members of the coalition [which] indicated the widespread suspicion that the exchange of visits would be a prelude to a sell-out.[29]

The audience costs, far from enhancing the credibility of the original threat, undermined it, making it far more likely that Hussein would resist in the hopes that the coalition will fall apart before decisive action could be taken.

Another problem is that the government may not be able to restrict audience costs to the capitulation outcome. As Chiozza and Goemans (2004b) have shown, the hazard of losing an office after a war is not systematically higher than after a crisis, but the fate of a leader who is ousted after losing a war is significantly worse than getting ousted without fighting one. In other words, while it may be possible to generate audience costs, they affect the expected payoff from war in addition to the capitulation payoff. Under some circumstances (e.g. non-democracies), they may actually make war worse than concessions, that is, an empirical direction exactly opposite to the one postulated by the theory. In any case, the range of audience costs that the government can generate depresses, and consequently does not permit the full separation we observed in the original audience costs model. If governments are limited in the magnitude of audience costs they can generate or cannot restrict them to the capitulation outcome, then their signaling abilities decline precipitously. Coupling this with the theoretical difficulties in accounting for the generation of these costs in the first place only makes matters worse.

Reputation and International Audiences

If leaders cannot generate sufficiently high domestic audience costs, then perhaps they could rely on costs that other audiences can impose? Interstate crises are usually public affairs and every international actor is a potential audience even if he does not participate in the crisis himself. Leaders may contemplate the inferences that these spectators will draw, and the consequences this will have in future interactions with them. In other words, leaders who may not be overly concerned about being punished by a domestic audience, may well be worried about suffering at the hands of an international one. For example, if backing down in the present crisis convinces another adversary that one is unresolved, this may undermine

[29] Freedman and Karsh (1993, 240-41).

one's deterrent posture with respect to that adversary, and precipitate a fresh challenge with potentially devastating consequences. It is worth noting that the costs from new aggression are likely to be far worse than anything a domestic audience is likely to impose. Perhaps such reputational effects that concessions may have can help establish credible commitments?

Certainly history seems to be full of instances where a leader appears aware that his actions may cause others to revise their expectations to the detriment of his reputation, and where this seems to have caused him to alter his behavior correspondingly.[30] Charles V worried that challenges to the periphery of the Habsburg Empire that go unpunished would encourage his enemies to threaten its core (Hopf, 1991). Dean Acheson worried that if the US did not help Turkey stand up to the Soviets in 1946 over the straits, the "whole Near and Middle East" would collapse, the so-called Domino Theory (Acheson, 1987; Mark, 1997). President Kennedy worried that if the US did nothing to compel the Soviet Union to remove its missiles from Cuba after he had made a clear deterrent threat (that failed), the Russians would be encouraged to test American resolve in Europe by moving on West Berlin. Sagan (2000, 98) in fact uses this case to illustrate that because of the reputation costs of failing to follow through, a "deterrence threat does not just *reflect* a commitment to retaliate; it *creates* a commitment" (emphasis in original). This, of course, is precisely the point I am making here although I would not call the result the "commitment trap," if only because one is not really without an option: not starting a war certainly always remains a possibility, and reputational consequences be damned.

Reputation can be engaged by verbal, rather than physical, acts: "The usual way that leaders commit themselves is simply by their words. Someone who backs away from a clear statement will lose future credibility and reputation" (O'Neill, 1999, 127). From a theoretical standpoint, words are cheap, and *cheap talk* is supposed to be ineffective precisely because it is costless. We know, of course, that the less adversarial the setting, the more information would costless communication convey (Crawford and Sobel, 1982; Farrell and Gibbons, 1989). However, a crisis is the epitome of an adversarial confrontation, so why should words matter? Under the reputational hypothesis, cheap words acquire binding meaning because actors latch onto them and condition their future behavior on whether one abides by the statements he makes (Sartori, 2002). Guisinger and Smith (2002) make the intriguing argument that because of these reputational concerns on the international level, citizens would want to punish their leader when he is caught bluffing. The concern for the reputational costs then directly gives rise to domestic audience costs.

The widespread approach to building reputation is by engaging in activities that are inherently costly (Kreps and Wilson, 1982; Alt et al., 1988). If there are benefits from convincing the other actor that one is resolved, then a resolved actor must do something than an unresolved one would not want to do. Although most models seem to treat building reputation in terms of a weak player mimicking the behavior of a strong one, reputation is essentially about the strong one separating from the weak.[31] That is, building reputation is a way for re-

[30] Hopf (1994), Mercer (1996), and Press (2005) all present extensive studies that basically agree on two conclusions: leaders do worry about reputational consequences of their actions very often but, ironically, the reputational effect seems non-existed. In other words, whereas we have plenty of evidence that leaders think about reputation, we have no systematic evidence that reputation actually matters. Of the three, Press (2005) is most emphatic in the argument that leaders do not lose credibility if they back down in a crisis.

[31] The traditional method is to introduce a small probability that one's opponent is an irrational player who

solved players to prevent bluffing by unresolved ones, much like costly signals in our model. It is in this context that engaging one's reputation becomes useful as a signaling device. As Ellsberg (1975, 27) explains, the player

> can stake his honor, his prestige, his reputation for honesty—if he has any of these—on carrying out [the threat]... thus [their] importance... They can be pawned. They can be wagered, risked, put up as security; they are something to lose, which can make more credible choices designed to preserve them.

The idea is that a player can engage his reputation to such an extent that he is almost certain to do what he threatens to, and this risk discourages any potential bluffers from imitating the strategy. This, of course, is the very logic of imposing costs on oneself, and the committing effect is the same. The crucial question then is: To what extent can an actor actually manipulate his reputation? If his ability to do so is limited, then engaging reputation becomes a rather ineffectual strategy to signal resolve.

Like any other audience-cost mechanism I have discussed, the reputational one suffers from the costs being imposed only indirectly. Rather than arrange for a penalty that she will surely suffer if she fails to carry out the threat (as we assumed would happen in our formalization), the challenger tries to bind herself with beliefs of others. The reputational logic makes it quite clear that she is in effect threatening others with their own inferences: "If I do not retaliate, *you* will conclude that I am not resolved, which will undermine my credibility in the future, causing me further difficulties with *you* (or others); to prevent *you* from making this inference, I must retaliate today." In his critique of Ellsberg's lumping together of reputation and imposition of a penalty, Wagner (1992, 128) gets it exactly right: "the latter represents a direct alteration in the payoffs... The former, however, can only influence her payoffs indirectly as the result of some influence on the beliefs (and therefore the behavior) of potential victims."[32]

This places the mechanism largely outside the control of the threatener, and it should come as no surprise that the substantive impact of behavior on reputation seems to be largely overwhelmed by situational factors. Despite their protestations about being influenced by what others would think if they back down, leaders most often seem to react to the immediate characteristics of the crisis (Press, 2005). Reputation is only valuable in the long-run, but the future is fraught with uncertainties whereas risks and costs to maintain it must be borne right now. This is what Bismarck had in mind when the said that "preventive war is like suicide from fear of death."[33] Preventive war is based on the notion that one must strike

always chooses to escalate and then to investigate the conditions under which an unresolved player would escalate as well (Kreps et al., 1982; Kreps and Wilson, 1982; Milgrom and Roberts, 1982; Fudenberg and Levine, 1992). There is something troubling about a rationalist explanation that has an irrational automaton at its core. Mailath and Samuelson (1998) argue that reputation involves the strong separating from the weak rather than the weak imitating the strong. See Mailath and Samuelson (2006) for an exhaustive treatment of reputation in economic contexts, and the conclusion to Chapter 3 about the implications of the theory developed in this book for the traditional arguments.

[32] Schelling (1982) also argues that it is very hard to control behavior of others by manipulating their expectations when they have to make inferences based on one's reputation. Jervis (1970) emphasizes the difficulty in drawing inferences when reputational concerns are involved, and even gives examples of a negative feedback: where an action designed to support one's reputation actually leads to the opposite inference. Nalebuff (1991) notes that self-falsifying prophecies should not arise in a reasonable equilibrium.

[33] Cited in Levy (1987, 103). The shadow of the future hangs heavy on the preventive motivation for war, much

now before the relative balance of power shifts against him. The logic is akin to having to pay enough costs today in order to build a reputation that may prevent having to pay them tomorrow. When the circumstances under which the future contingency may arise depend on the opponent's expectations, the costly purchase of such an uncertain effect may be quite unattractive.

Finally, two other features of the audience-cost mechanism make it less interesting when it comes to studying military moves. First, an opponent can do very little, if anything, to hinder one's ability to generate audience costs. This tactic cannot be countered by a strategic move. Whether this is good or bad depends on whether one is the side generating these costs or the opponent. As we shall see, this is one important characteristic that is not shared by the military instrument. Second, audience costs are "free" in the sense that leaders only pay them if they back down without obtaining concessions from their opponents. This means that leaders can generate as much of them as they need if they think there is a good chance the other side will quit first. Audience costs do not really alter the situation the way a physical move might, at least not independently of how the crisis ends. I shall return to these issues in the next chapter.

The Expected Value of War

I have now argued that most studies have focused on how leaders can communicate their commitments credibly during a crisis, and in this some of the analyses have touched on the possibility for creating commitments in the process of communicating them. In the latter, scholars have emphasized that the ability to impose costs for failing to carry out the threat can commit one to a course of action. With all this discussion of penalties for not following through, it is somewhat surprising to see almost no analysis of rewards for executing the threatened action. That is, in all the talk about manipulating the costs of concessions, the benefits of attacking have been neglected. But as I explain at greater length in the next chapter, the military instrument's purpose is precisely to increase these benefits, and thereby affect the other side of the credibility constraint. By improving the expected payoff from the use of force, a government can establish a credible commitment.

The payoff from war includes, at the very least, two components in addition to the value of the disputed issue—the probability of winning, and the costs of fighting—but we can certainly think of more. Leaders may take into account the consequences of victory or defeat that go beyond the issue itself. For example, one pertinent question concerns their own fate if they lose. While democrats are easier to oust, their punishments are relatively mild compared to autocrats who are harder to remove but tend to suffer rather extreme punishments when they are deposed (Goemans, 2000; Bueno de Mesquita et al., 2003; Chiozza and Goemans, 2003). However leaders estimate their likely fate in these contingencies, they will take it into account when forming expectations about the war.[34]

like reputational considerations are supposed to affect today's decisions. How much of a problem this is seems an empirical question. After all, when President Truman decided to escalate the Turkish Crisis in August 1946, he remarked that "we might as well find out whether the Russians were bent on world conquest now as in five or ten years" (Mark, 1997, 383).

[34] This is an open question and there is a growing literature on this topic. That leaders have incentives that are different from the constituencies they serve is at the heart of rational choice explanations of diversionary war (Levy, 1989b). Leaders may "gamble for resurrection" by engaging in risky international adventures when

In general, anything that influences the victory and defeat payoffs differently will be relevant (costs are paid regardless of the outcome). Perhaps total victory is not very attractive because eliminating the opponent removes a valuable buffer between the state and an even worse adversary. Perhaps such a victory would provoke a concerted effort to deprive the state of it. Perhaps defeat would undermine the state's general deterrence posture and will make it an inviting target. Perhaps former allies will abandon the weakened state and gravitate toward its victorious opponent. All of these possibilities, and perhaps many others, may find their way in forming an estimate about the expected value of fighting. Almost all of them have one thing in common: they are very difficult for the actor to manipulate. This is why I concentrate on the two components I identified above as basic: both are somewhat manipulable, and since they do affect the expected war payoff, the actor can use them to manipulate his incentives.

Of the two, the expected war costs are probably harder to manipulate, but not impossibly so. War costs include human casualties (soldiers killed, missing, or wounded in action, as well as civilians), evacuation and medical treatment of wounded, transportation, attrition, and replacement of military stock, destruction of physical assets, degradation of overall economy, losses from trade, exhaustion of financial resources, and deterioration in credit worthiness. If one adds the long-term costs of stunted economic growth, reconstruction, psychological trauma of large segments of the population, political and civic turmoil, and takes into account opportunity costs, then a war's destructiveness is immense. However, there are actions that one can undertake to reduce it. For example, if casualties are of paramount concern, then switching to tactics designed to minimize them will be worthwhile. Such tactics may necessitate the restructuring of the armed forces, the development of new technologies, and improved training, but it is certainly within the grasp of certain nations. Civilian defenses are another obvious target: developing evacuation plans, building shelters, and providing proper training will also reduce casualties.

The probability of winning (or obtaining a favorable settlement) itself depends on a variety of factors, some of which can be manipulated. Obviously, an actor's chances of victory will depend on his military forces: their size, technological sophistication, state of readiness, and quality. This latter characteristic depends on the political organization of the state and its armed forces (which determines the pool from which soldiers are recruited), and on the economic resources (which determine how much the state can spend on its military). Total economic and human resources are also important because of their long-term mobilization potential, although this is mediated by the organizational efficiency of the state apparatus (Brewer, 1990). One should not neglect less tangible, but also critical, variables such as the quality of leadership, both civilian and military, and their faculties of initiative and imagination. Polity characteristics that are relevant will include, among others, the willingness to tolerate casualties, as well as the government's ability to maintain war-time political consen-

their chances of remaining in office are low but when the high variance of their outcomes might just save them at the ballot box (Richards et al., 1993; Downs and Rocke, 1994; Bueno de Mesquita et al., 1999). Smith (1996) studies the perverse incentives for leaders that may lead to such a war, Tarar (2005) investigates the consequences of strategic choice of targets, and Chiozza and Goemans (2004a) wonder about the dangers of targeting a country whose leader has diversionary incentives. On the other hand, Chiozza and Goemans (2003) show that incentives generally thought to lead to war proneness are actually empirically conducive to peace.

sus (which may require suppressing information, sanctioning dissidents, and the extensive use of propaganda). Finally, the behavior of third parties will also affect the probability of victory. One will have to take into account the dependability of one's allies and the utility of their expected contributions, as well as the potential for others to join the opposite side.

As we shall see in detail in the next chapter, a great many of these variables are within the partial reach of states. At the very least, mobilizing one's forces during a crisis improves readiness, minimizes the risk of surprise attack, and improves the chances of victory. The existing theories that depend on rearranging incentives, however, usually neglect the war payoff (taking it as fixed), and instead focus on manipulating the costs of concessions.[35]

2.4 Conclusion

Although credible commitments are at the heart of deterrence theory and crisis bargaining in general, the process of establishing them has largely been missing from much of the theoretical analysis.[36] Formal work has identified general consequences of inability to commit (Powell, 2004) and how to communicate existing commitments effectively. However, it has not, as a rule, dealt with where commitments come from. Non-formal work has studied consequences as well (Walter, 2002), in addition to communication problems (Jervis, 1970), threat (mis)perception (Lebow, 1981), psychological pathologies or cognitive limitations (Jervis et al., 1985), and tactics that emphasize accommodation rather than coercion (George and Simons, 1994).[37]

The traditional approach has led to some confusion about credibility. The rational deterrence debate, for example, seems to have become an agreement to disagree about intent—the bone of contention being the answer to the question if a side in a conflict was "truly" intending to fight or not.[38] Although much of the archival research designed to divine these hidden preferences is useful, it is beside the point because it tacitly assumes that these preferences

[35] Morrow (1994a) and Kydd (2000) are exceptions.

[36] On informal theories of deterrence, see Kahn (1960), Snyder (1961), George and Smoke (1974), Morgan (1977), Jervis (1979), and George and Simons (1994). The debate in *World Politics*, 41:2 (January, 1989) is especially illuminating. Among the best empirical works are Huth (1988) and Huth and Russett (1993). The crisis literature is voluminous although only a fraction of it deals with crisis bargaining. The canonical works are Kahn (1965), Hermann (1972), Snyder and Diesing (1977), and Lebow (1981), with Brecher (1993) and Leng (1993) providing modern critiques. Banks (1990) uses mechanism-design to demonstrate properties of any game-theoretic crisis bargaining model where the expected payoff from war is fixed (this includes almost all existing models).

[37] See Freedman (1998), and Morgan (2003) for critical overviews of deterrence theory that include discussions of credible commitments. Zagare and Kilgour (2000) provide a comprehensive formal treatment of deterrence, and Powell (1990) explores credibility problems with the threats to use nuclear weapons. It is worth emphasizing that none of these works analyzes the use of military moves to create commitments. Even works specifically dedicated to the use of the military instrument for political ends, such as Karsten et al. (1984) and Cimbala (1994), do not deal with the functional aspect. The classic Snyder and Diesing (1977, 199–200) mentions that possibility in passing. As usual, Schelling (1966) is the authoritative exception.

[38] See the extensive exchange between Paul Huth and Bruce Russett on one side and Richard Ned Lebow and Janice Gross Stein on the other: the opening salvos in Huth and Russett (1984) and Lebow and Stein (1987), the rejoinders in Huth and Russett (1988) and Lebow and Stein (1989), and the final (for now) round by Huth and Russett (1990) and Lebow and Stein (1990). Adding oil to the fire are Orme (1987), Levy (1988), and Harvey (1995). This debate has been very important in clarifying testing methodology, case coding, and

remain largely unaffected by the crisis itself. Whether a leader would actually fight or not is not as relevant as whether his opponent thinks he is ready to fight. The postulate of unalterable preference has quietly propagated itself to formal models of crisis bargaining via the assumption that the expected payoff from war remains fixed throughout the crisis.

The possibility of rearranging incentives has not been entirely neglected, as the studies of domestic audience costs and reputation clearly show. However, both approaches provide leaders with levers that they barely have access to in practice. As a consequence, the ability to commit derived from such devices is suspect. On the other hand, the fundamental process of creating and communicating commitments reveals that the expected payoff from war is just as important as the expected costs from backing down. The premier function of the military instrument is to improve one's chances of a favorable outcome if fighting occurs, and hence the use of force primarily affects the expected payoff from war. To understand the coercive effect of threats to use force, we must examine how effective the military instrument is in the crisis bargaining context outlined in this chapter.

Unfortunately, we cannot use any of the existing analyses because the military instrument does not fall neatly into any of signaling or commitment categories that we have studied. As I argue in the next chapter, the preparation to use force (that is, the threat embodied in a physical deployment of armed forces) is an act that combines features of sunk costs and incentive-rearrangement. If we are to understand how threats are used, communicated, and perceived, if we are to investigate what effect they have upon crisis participants and on crisis outcomes, we must analyze them carefully, just like we have done with the other mechanisms.

Crisis bargaining models that ignore the commitment-creation process do not capture the empirical substance of the situation they are designed to represent. Existing abstract ideas about how such commitments can be created are unattractive because some are never used in practice, and others make fairly heroic assumptions about a leader's ability to make the peaceful outcome worse than war by manipulating variables over which he has almost no control. Moreover, since the inability to create dynamic commitments has grave consequences for crisis stability, if we fail to incorporate the strategic interdependence of such attempts, we should expect that analyses based on such models are likely to lead to wrong substantive interpretations, and policy prescriptions that will be radically incorrect in their estimation of the relationship between military moves, the probability of war, and the prospects for peace.

competing explanations. My point is that both sides have tacitly agreed on certain shared assumptions that turn out to be untenable.

Part II

A Theory of Military Threats

3

A Model of Military Threats

Diplomacy without arms is like music without instruments.
Frederick the Great

The arguments in the previous chapter outlined the general approach to crisis bargaining scholars have employed for decades. By constructing a stark stylized world of the basic crisis game, we were able to explore the interrelated problems of credibility, commitment, and communication that must lie at the heart of any theory of coercion. Although I was at pains to argue about the theoretical and substantive shortcomings of many of the proposed mechanisms, the real motivation for a new investigation of military coercion comes from the simple insight that the military instrument does not fall into any of the idealized tools of statecraft we have discussed. This is somewhat surprising because when we speak of coercion in a serious interstate crisis, we usually mean threats to use force, military force to be exact.

Threats to use force can be delivered verbally, or they can be tacit, implicit in the physical deployment of military units, or they can be a combination of the two. Physical military moves are a very often an indelible feature of crises. Consequently, I will focus on such preparations to use force because they can be used for signaling and commitment; that is, they can be used for bargaining purposes. Military moves are a suitable candidate for coercive bargaining because they have both informational and functional aspects that do not suffer from the empirical implausibility of other commitment tactics.

3.1 Characteristics of a Military Threat

Before constructing a model of military coercion that builds on the basic crisis game but that takes physical military moves seriously, I provide the substantive justification for making the assumptions of that model with respect to the effect military moves have on the payoffs. George (1994, 10) sees coercive diplomacy as "an alternative to reliance on military action." This, however, is thoroughly misleading for without the military moves one can hardly think of exerting serious coercive pressure. Before examining how military moves differ from economic sanctions and diplomatic coercion, it is worth discussing their characteristics, especially in relations to audience costs, all of which combine to make them much more plausible as candidates for coercive bargaining.[1]

[1] Military moves may actually increase audience costs. For example, Schultz (2001, 211) reports that French Defense Minister Maurin argued against a bluffing show of force to compel German withdrawal from the Rhineland in 1936 because "the government risked public ridicule if it were to mobilize a million men, only

For our purposes, a military threat is any physical move that satisfies the following two criteria:

1) it is inherently costly, and
2) it changes the distribution of power during the crisis.

Such a threat is *inherently costly* because the actor must pay to make it regardless of how the crisis turns out. As such, it is a variety of a sunk cost. Unlike pure sunk costs, however, military threats delivered through physical moves also affect the actor's probability of victory should war break out. In this context, the *distribution of power* can be usefully summarized by the probability of prevailing in an armed conflict. Clearly, it depends crucially on the economic, military, social, and political organization and capabilities that pre-date the crisis itself. This, in fact, is how theories of crisis bargaining usually treat it, and this is how we have treated it in Chapter 2 as well. But, as I will argue here, intra-crisis behavior can affect this distribution of power, at least at the margin.

Before going on to substantiate these two assumptions empirically, let me give some examples of physical moves that I have in mind and some that do not fit these criteria. Any move that increases one's preparation for fighting would satisfy the two requirements:

- moving naval, air, or army units close to a potential theater of operations,
- putting nuclear forces on alert,
- canceling leaves of military personnel and calling up reservists,
- performing military maneuvers and demonstrations in strategically sensitive areas,
- occupying or fortifying of strategic positions,
- reaching agreements with third-parties to permit troop passage or give over-flight rights,
- establishing military bases,
- financing or supplying third-parties hostile to the opponent, especially if they are currently engaged in conflict,
- ordering full-scale general mobilization.

On the other hand, actions that have only medium to long term consequences for the distribution of power would not be considered a physical military move no matter how provocative they are. For instance, increasing the military budget or passing legislation to increase the draft or implementing economic sanctions that impair the opponent's ability to wage war would no doubt affect the probability of victory in the future but since they have no appreciable effect in the short-term, such actions would not change the distribution of power during the crisis itself. Furthermore, purely political activities, such as bellicose speeches or dramatic public confrontations of will, would not satisfy the two criteria. Likewise, some defensive measures such as evacuating civilians from threatened areas, would also not count because they do not affect the distribution of power even if they do affect one's costs of fighting and as a result do increase one's expected payoff from war.

This last example illustrates the important and demanding requirement of the second criterion: it is not enough that the action affects one's own expected value of war, it must also

to march them back home if Hitler did not cave in." This just underlines the point I made about the magnitude of such costs: are we to believe that "public ridicule" is somehow worse than fighting a bloody war? Of course not, and this is precisely why the French could not meet Hitler's bluff with one of their own. They could not generate enough audience costs to actually commit to pushing his forces out of the Rhineland.

affect one's opponent's value of war. Actions that only reduce one's costs of war do make fighting relatively more attractive for one but since they do not necessarily affect the expected payoff of one's opponent they do not directly reduce the relative attraction of war for him. The distribution of power is a "common value" in the sense that whenever one's actions change it, the change is reflected in the opponent's payoffs as well. Some defensive measures, such as hardening of missile silos or introduction of better armor for tanks, will affect this common value because they make it harder for the opponent to destroy one's forces and therefore reduce the probability that he will prevail in the military confrontation. To make things a bit more concrete, I now provide several examples to justify the two criteria.

3.1.1 Military Moves Are Inherently Costly

Physical military moves represent sunk costs, which means that leaders are not free to generate them without limits, and in fact may find it hard, if not impossible, to pay for sufficient mobilization when necessary. Crises, of course, are very different from wars when it comes to costs. It is easy to see what disagreement entails while fighting goes on: continued pain and destruction. But what does prolonged disagreement in a crisis cost? When it comes to physical military moves, the price tag may include direct and opportunity costs, costs in terms of increased vulnerabilities, alert fatigue, accidents, and behavior of economic agents. All of these are sunk, as McClelland (1961, 200) nicely puts it:

> In the mobilization of a crisis, many sectors of the national social organization not usually involved in the "normal work-flow" of international relations will become agitated and active, while in the demobilization of the crisis there will be a rapid falling-off of such activities and a return of affairs to routine channels. These processes should take place whether or not issues are settled, problems are solved, or relations are "improved," although we may find that crises recede behind a veritable smokescreen of conflict resolution promises and of problem-solving talk.

Direct and Incremental Costs

Direct costs must be paid for salaries (including hazardous duty pay), subsistence items (e.g. food and water), and transportation of personnel and equipment (including contracts with commercial air and shipping). Examples of such costs are too numerous to list, but here are some illustrative cases from various periods of history.

Throughout the Middle Ages, Venice routinely had to resort to forced loans, special taxes, and even the sale of hereditary aristocratic titles or lucrative government posts to finance its mobilizations. During one of its many wars with Genoa, Venice found itself facing a victorious enemy fleet under Paganino Doria advancing in the Adriatic in 1354. The general mobilization was costly because in addition to the usual expenses incurred by the citizens, "a great boom of tree-trunks and iron chains was run across the Lido port between S. Nicolò and Sant' Andrea." The measures proved successful and Paganino advanced no further (Norwich, 2003, 221-22).

Provisioning the army is an incredibly expensive undertaking. According to one estimate, a 17th century army of about 60,000 would require "45 tons of bread, over 40,000 gallons of beer and the meat from 2-3000 cattle *every day*. Its animals consumed 90 tons of fodder (the equivalent of 400 acres of grazing) and each of its horses needed 6 gallons of water per day to remain healthy" (Tallett, 1992, 55, emphasis added). Such an army would require

"40,000 horses between cavalry, artillery, and baggage. [. . .] Overall consumption would thus amount to no less than 980,000lbs. per day" (Van Creveld, 1977, 24). These staggering requirements produced a system of "regulated plunder" of territory even if it was friendly or neutral and affected military strategy (Lynn, 1999). The demands increased exponentially if the forces had to march to their point of deployment for now one had to reckon with animals to haul the luggage, food, and munitions. In 1606, the Spaniards had to provide between 2,000 and 2,500 wagons for a force of 15,000. The inability to meet such demands with the limited resources available caused desertion and loss of control, which impaired military effectiveness (Van Creveld, 1977, 6-8). Desertion could cripple an army more than losses in combat or to disease. "Commanders routinely expected to lose 10 or even 20 per cent of their men annually in this way, but wastage rates could be far higher" (Tallett, 1992, 116).

In their preparation for the Second Boer War, the British House of Commons authorized $50 million for mobilization of 47,000 soldiers for four months.[2] The "1974 ten-day Greek mobilization [after the Turkish invasion of Cyprus] is said to have cost about 650,000,000 German marks, not taking account of costs involved in lost working hours" (Goldhamer, 1979, 41).

During the intense diplomacy to regain the Falklands after Argentina seized them on April 2, 1982, the British began costly preparations for war:

> With a British invasion of the Falkland Islands appearing imminent, Britain has now mobilized about 22,000 men in a task force of more than 70 ships strung out over the Atlantic Ocean at an estimated cost of more than half a billion dollars. That means Britain is using 12 men and spending at least $275,000 for each of the 1,800 islanders in its undeclared war with Argentina to regain the islands seized more than six weeks ago.[3]

To meet all this rather extravagant expense in a "non-inflationary way, consistent with its economic strategy," Thatcher's government had to consider higher taxes and cuts in public spending.[4] In early May, the government requisitioned or chartered commercial vessels for the task force, paying for them at commercial rates and covering the companies' operating costs.[5] The Argentines, who had already occupied the islands, were faring even worse. The cost of occupation itself was estimated at over $500 million and panicking investors "sent the price of dollars on the black market. . . 20% more than the official quotation."[6] In the event, the preparations on both sides did not prevent the conflict from escalating into open warfare.

This was not the case in late 1978 when Argentina was embroiled in a dispute with Chile over the Beagle Channel islands. As *Newsweek* reported at the time, "the war with Chile. . . did not take place last week. But it could break out any time. Over the past eight months, Argentina has spent nearly $1.2 billion to beef up its armed forces, which now boast more than 100 modern navy warplanes, a pair of aircraft carriers and a newly called-up reserve of 500,000 troops. Not to be outdone, Chile has paid $800 million to buy a fancy new

[2] *The New York Times,* October 21, 1899.
[3] *Washington Post,* May 19, 1982.
[4] *Financial Times,* April 10, 1982.
[5] *The Economist,* May 8, 1982.
[6] *Financial Times,* April 8, 1982.

French surface-to-air missile system and thousands of Israeli weapons."[7] Even though the conflict did not end up in fighting, both sides had to bear the burden of military preparations.

The costs of deployment of US forces to Saudi Arabia for Desert Shield in 1990 was averaging $28.9 million per day in August. Critics of the second build-up that doubled the troops cited the costs as a fundamental shortcoming that would not allow the United States to "settle down for a long haul," which would make a military solution more likely because diplomacy and sanctions would not be given sufficient time to work (Freedman and Karsh, 1993, 216,227).

The Indian mobilization after the December 13, 2001 terrorist attack on Parliament involved moving more than half a million troops to the border with Pakistan. The costs of this massive deployment were about 1.1 billion pounds and 800 soldiers dead in accidents during the process. In addition to the military moves, India suspended air and land traffic to Pakistan, a costly interruption of travel and transportation, with adverse consequences for economic activities. It is unclear what the overall costs of this mobilization were over the entire duration of the crisis, which lasted for nearly a year. Both soldiers and civilians died from landmines, many peasants appear to have lost their homes and access to their lands when the military moved through or mined the area.[8]

Sometimes a country may be overwhelmed by the financial demands of military moves. As pressure from the Soviet Union mounted in 1939, Finland had to resort to loans to finance its mobilization. The United States authorized an initial loan of $10 million in 1939, which it then extended to $20 million once war was underway.[9]

Depending on political system and type of military organization, mobilization may essentially involve the upkeep of a standing force that requires often excessive remuneration. Once organized, such a force may prove exceedingly costly because any decrease in the expected level of compensation may produce violent morale-sapping discontent, mutiny, or even outright rebellion. For example, in China, in 799 the new military governor refused to provide the provincial Xuanwu Army the gift of silk and hemp cloth that had become customary. The enraged soldiers sliced him up and ate him, perhaps the ultimate cost a government official can pay. As Graff (2002, 239) notes, at least ninety-nine mutinies during the late Tang period were caused by soldiers' discontent over payments. Mutiny was also chronic in the European armies of the turbulent 15th through 18th centuries. There were twenty-one in the Army of Flanders over less than a decade between 1596 and 1607. When mutinies occurred on a large scale, rulers often had to capitulate to the demands for higher pay or court military disaster. When the government did not have the means to meet its obligations, it could be forced to borrow from unwanted creditors, as Sweden did in 1641 from France (Tallett, 1992, 116-17). When France mobilized against Prussia in 1870, the chaotic plans failed to supply the fortress of Metz with adequate provisions of sugar, coffee, rice, and salt. Bucholz (2001, 170) summarizes the consequences:

> Troops were on their own. After they had eaten the locals out of house and home, they turned to pillage. Discipline lapsed. Soldiers did what they wanted. They wandered off, and came back as they pleased. Such an army was in no condition to fight a war.[10]

[7] *Newsweek*, November 13, 1978.
[8] Roy (2006), and *The Guardian*, Friday 15, 2006.
[9] *The New York Times*, December 11, 1939 and March 1, 1940.
[10] See also Howard (1981, 71).

The rapid deterioration of morale of a mobilized force that is not put to any use is well-documented and is, in fact, often used by opponents in what Liddell Hart (1991) termed the "indirect approach." In his contest with Pompeii who wielded forces much larger than his own, Caesar consistently avoided battle with the mobilized enemy and instead chose a strategy of "maneuvering repeatedly to inflict a series of pinpricks whose wearing and depressing effect on the enemy's morale was shown in the swelling stream of desertions." A few centuries later, the Byzantine general Belisarius was forced to attack the Persians because his own soldiers were getting restless and discontent with his (highly successful) strategy that avoided fighting but forced the invaders to abandon their designs on Antioch. The superior Persian force inflicted one of the very few defeats in the General's illustrious career (Mahon, 2005). The morale-sapping inaction and wearying suspense that accompanies also goaded the Romans into one of the most disastrous defeats in their history when Consul Varro, against the better judgment of his colleague Paullus, offered battle to Hannibal on terms heavily favoring the Carthagenians in 216 B.C. The slaughter at Cannae did not break Rome but the Republic lost almost its entire army of eight legions, the largest it had ever fielded (Goldsworthy, 2002).

The Romans were not the only ones tempted into action by rapidly deteriorating morale: in the summer of 1913, the commander of the Bulgarian army, General Savov, warned his government that they either had to fight or demobilize the army. The soldiers, under arms for nearly a year and exhausted by the fighting in the First Balkan War, were mutinous in the inaction. The American military attachè in Sofia argued that this "discontent will, however, cease to exist when the men are called upon to fight against Servia and Greece."[11] Increasingly isolated diplomatically, Bulgaria resorted to a desperate action and precipitated the disastrous inter-allied war.

Incremental costs arise from operation at increased levels of activity (more flying hours, higher fuel expenditures), maintenance (more repairs), and force reconstitution (replacement of worn or damaged equipment, replenishment of munitions stocks). For example, the September 2002 official estimate of the Congressional Budget Office put the deployment costs of American forces for possible operations in Iraq between $9 and $13 billion, plus between $5 to $7 billion to return the forces to their home bases.[12] The US Department of Defense provides for separate funding of Mobilization/Surge Costs to pay for mobilization and wartime surge capacity (facilities and equipment held in standby or idle status) that would not be incurred in peacetime. Costs also increase due to the inevitable accidents accompanying placing troops in a state of high readiness. For example, Betts (1995) reports that prior to Desert Shield mounting American deaths in nighttime helicopter accidents forced restrictions in the interests of safety, and "the number of deaths rose to more than a hundred *before* the war began in mid-January 1991."

Unpredictable events may make continued mobilization costlier still, both directly (human lives) and indirectly (by undermining the capacity for further action). For example when Venice sent an armada of about 120 ships to deal with Byzantium in 1171, Emperor

[11] Cited by Hall (2000, 103) who also documents the morale problem.

[12] These estimates did not include actual fighting and subsequent occupation, and were based on the costs incurred in operations in the Balkans, Afghanistan, and the prior war in Iraq. The report, "Estimated Costs of a Potential Conflict in Iraq" is available at http://www.cbo.gov/ftpdocs/38xx/doc3822/09-30-Iraq.pdf. Accessed January 11, 2005.

Manuel feigned a peace feeler designed to give him some breathing space to prepare to face the threat. Unaware of the duplicity, Doge Vitale Michiel accepted and instead of sailing to the Bosphorus, dropped anchor at the island of Chios to await the resolution of the negotiations. While the Byzantines dallied in Constantinople, plague struck the overcrowded Venetian ships, and within months the fleet was rendered incapable of any offensive action. Thus, when the ambassadors returned from Constantinople with the news of the complete collapse of the talks, the hapless Doge turned his decimated, demoralized, and nearly mutinous fleet back to Venice without ever managing to engage the enemy.[13]

Disease, in fact, was a common problem, especially before the advent of modern medical care. "With their dense concentrations of personnel, and their insanitary camps, billets and hospitals, armies offered ideal conditions fr the transmission of crowd diseases, such as typhus, smallpox and the bubonic plague, together with those diseases associated with poor hygiene, such as typhoid and dysentery. [. . .] one-third of the Marquis of Hamilton's force were dead within a month of arriving in Pomerania in 1631 following an outbreak of a contagion" (Tallett, 1992, 107).

When Romania invaded Bulgaria in 1913, it suffered no combat casualties because the Bulgarian military, wholly committed in action against Serbia and Greece, could offer no resistance. However, the Romanians lost about 6,000 soldiers to cholera, with the disease spreading upon return of the army home. The Ottoman Turks, who also used the opportunity to invade Bulgaria, similarly saw no combat but lost about 4,000 men to the disease (Hall, 2000, 118-19).

Opportunity Costs

Opportunity costs arise from removing a portion of the economically active population from productive use, and requisitioning of vehicles, railroads, and highways for military use. These are especially costly to countries without professional militaries that have to rely on conscripts.

In 1921, just a few months prior to the outbreak of the frontier crisis with Yugoslavia, the Albanian representative to the Council of the League of Nations complained that the border dispute was placing a considerable strain on the Albanian economy because the country was forced to maintain the mobilization of 10,000 troops (out of population of 1 million) just to defend the border.[14]

During the run-up to the Six Days War, the "price of the [Israeli] mobilization was staggering." The General Staff estimated that maintaining operational readiness without fighting cost the country approximately $20 million per day. The mobilization was so costly and the wait so risky for Israel, that the US offered to "furnish a number of items—100 half-tracks, Patton tank and Hawk missile parts, food and economic aid totaling $47.3 million, plus a $20 million loan—to tide Israel over" as long as Israel did not challenge the Egyptian blockade or precipitate war (Oren, 2002, 79,87,98). "As long as general mobilization was in effect, the Israeli society and economy came to standstill" (Maoz, 1990, 129).

Another type of opportunity cost arises from the increased vulnerabilities to third parties when one re-allocates forces from one potential crisis zone to another thereby exposing a

[13] Norwich (2003, 105-06). The Doge paid with his life upon his return to Venice; not only had he miserably failed in his military and diplomatic mission, he had brought back the plague with him.

[14] *The New York Times*, June 27, 1921.

weakened front that may be tempting to an adversary. Blainey (1988, 58) argues that the a "nation's decision to go to war always includes an estimate of whether outside nations will jeopardize its prospects of victory," and this equally applies to a decision whether to use the military instrument in a crisis if that would entail dangerous exposure elsewhere.

During the 1921 Albanian crisis with Yugoslavia, the Greeks seized the opportunity afforded by Albania's military preoccupation with defending its northern frontier and invaded in the South in an attempt to gain territory with disputed Greek minority settlements.[15] During the India-Pakistan Kashmir crisis of 1965, China took advantage of India's vulnerability, announced its support for Pakistan and delivered an ultimatum demanding that India dismantle its fortifications on the Sikkim-Tibet border or face war with China.[16] Even though Bolivia had affirmed its strict neutrality in a possible war during the Beagle Channel dispute between Chile and Argentina, both Peru and Bolivia had long sought to regain territories Chile had conquered during the War of the Pacific a century earlier. The Bolivians were not prepared but the Chilean government was worried about the Peruvians who engaged in military maneuvers along their southern border with Chile and stated that they will continue the maneuvers "as long as the possibilities of an armed conflict persist."[17]

Economic Costs from Market Agents

Since military moves are hostile acts, they are generally perceived as increasing the risk of war, causing economic agents to revise their expectations about the profitability of their investments. For example, a country that mobilizes and is expected to go to war may find it difficult to meet its debt servicing obligations, and hence government bonds will trade at lower prices, with the drop being especially acute if the country is also expected to do badly in the conflict. Frey and Kucher (2000) show that important foreign policy events before (during) World War II strongly affected bond prices of potential (actual) belligerents.

As *The Economist* reported on May 22, 1982,

> Nowhere is the impending invasion of the Falklands less popular than in the City of London. As military action seemed inevitable this week, the Financial Times index of industrial ordinary shares fell by nearly 29 points in three days. Sterling held up well, but government securities were weak. Bankers fear that the freezing of Argentine government assets held by British banks will damage the City of London's standing as a centre of international banking.

During the war scare over the Beagle Channel, "Argentina was a country of black-out drills, troop movements and fiery speeches.. . . There was panic buying in the supermarkets, a run on dollars and other hard currencies, and the stock exchanged plunged."[18]

One's own preparation for war may adversely affect the opponent's financial situation as well. Weidenmier (2002) tracks the Confederate Grayback price of a gold dollar during the American Civil War and notes how the passage of a conscription bill in the North (which increased its ability to mobilize soldiers) caused an increase in the gold premium.

[15] *The New York Times*, November 1, 1921.
[16] *Keesing's Record of World Events*, Volume 11, December, 1965, Page 21103.
[17] *Latin America Political Report*, LAPR XII, 50, p. 393. December 22, 1978.
[18] *The Economist*, November 11, 1978.

3.1.2 Military Moves Change the Distribution of Power

Military preparedness is an essential ingredient of one's ability to fight well, and so moves that increase it also increase the probability of victory. There are several factors actors must take into account when estimating their expected payoff from war: costs of fighting, probable outcome (whether victory is likely or if negotiations are to be expected), benefits from a favorable outcome, and losses from an unfavorable one. The impact of military moves is most readily apparent in the changing probability of securing a favorable outcome, and the costs of doing so. That is, they affect the distribution of power, which has two consequences: it improves one's own payoff from war and worsens the opponent's. In this way, *military moves can alter the strategic context by shifting the expected benefits from war for both sides.*

Local Distribution of Power

In general, victory in a total war belongs to the wealthier side that can mobilize more resources for war (Kennedy, 1987; Overy, 1995). However, most modern wars are limited and of relatively short duration: Slantchev (2004a) reports that for 104 interstate wars fought between 1815 and 1991, the mean duration was less than 14 months, and the median less than 6 months.[19] This means that in most cases, countries fight with their existing capabilities and there is no time to convert the industry to military use or rely on mobilizing resources from scratch. In other words, the probability of obtaining a favorable outcome depends heavily on the mobilized capabilities with which each belligerent enters the fray: a country that could potentially outlast its opponent once its resources are put to military use may find itself decisively defeated in the field and unable to resist and utilize that potential when it fails to mobilize sufficient forces at the outset. Military preparations may crucially affect the expected payoff from war even if there exists a significant resource asymmetry that renders the outcome of a protracted war fairly predictable *ex ante*.

For example, on the eve of the Russo-Japanese War of 1904, the distribution of power heavily favored the Russians who enjoyed over 5 : 1 superiority in military personnel, 3 : 1 superiority in both total population and energy consumption.[20] However, the theater of operations in Korea and Manchuria was very far from the European part of Russia where the bulk of its military was stationed. The 5,500 mile supply line with a 100 mile break at Lake Baikal made it exceedingly difficult to mobilize enough forces locally. East of Lake Baikal the Russians had only about 83,000 troops (from an army of 4,500,000) that the Japanese could counter with their land army of 283,000, which they could reinforce quickly with 400,000 reserves. Even though the Russians did manage to send reinforcements, in the final land battle at Mukden, the two sides were approximately evenly matched, allowing the Japanese to win the war (Dupuy and Dupuy, 1985, 1009-13).

Sending the air force to the potential theater of operations can have a serious impact on the distribution of power locally. The analysis of Argentina's dismal military performance

[19] This is certainly not the case for many early-modern wars which were protracted affairs of attrition that could go on seemingly interminably (Parker, 1996). The Thirty Years War is perhaps the most famous example but the later wars of Louis XIV illustrate the indecisiveness of what John Lynn calls "war-as-process" that characterized this type of warfare (Lynn, 1999).

[20] Figures from the Correlates of War National Material Capabilities Data Set, version 3.02 (Singer et al., 1972).

during the Falklands War reveals a number of key strategic errors, one of which was its "failure to move its air power from the mainland to the Falklands. Because of this failure, the potential superiority of the Argentine air force was canceled out since it was forced to operate from the mainland, 400 miles from the Falklands. The war could have turned out quite differently if the air force had been moved to the islands" (Hopple, 1984, 352).

Many wars end when opponents settle on a negotiated peace short of military victory (Slantchev, 2003b). In fact, some wars may be fought without military victory as a goal at all. Pericles' original strategy that Athens followed during the so-called Archidamian War did not envision defeating the Spartans in the field, but surviving long enough for their enemy to become convinced that it could not prevail, thereby softening them enough to offer better peace terms to Athens and its empire (Thucydides, 1996; Kagan, 1990). Similarly, the US entered the War of 1812 without the slightest intention of defeating Britain militarily. Instead, the goal was to persuade the British that continuing their trade embargo and impressment practices was too costly (Hickey, 1989). Perhaps the most famous example of a nation starting a war it knew it could not win is provided by the Japanese attack on Pearl Harbor: the Japanese harbored no illusions about what would happen to them once the US mobilized for war; they hoped to settle the conflict far in advance of that (Butow, 1961; Ike, 1967; Costello, 1981). Since an opponent who expects to suffer more is more likely to offer better terms, mobilization of forces that make such suffering possible and likely should increase the prospects of obtaining a favorable settlement even without military victory. Mobilization can increase the expected payoff from war because it increases the opponent's prospects of having to suffer great pain, and can therefore soften it into making concessions.

First-Mover Advantage

There are many ways in which military moves can change the distribution of power during the crisis. Van Evera (1999, 37–53) distinguishes between first-strike advantage (where the first to attack is more likely to prevail) and first-mobilization advantage (where the first to mobilize can start fighting before the other is ready). From our perspective these are equivalent: both constitute a *first-move advantage* because if a country manages to "get there" first, its expected payoff from war would be higher relative to its payoff if it ran second. If there exists an advantage to striking first, then preparing for such an attack makes it possible to execute it (one can hardly attack without adequate military prep), and therefore increases the probability of victory.[21] Although Germany eventually lost the Second World War, its opening gambit against the Soviet Union was brilliant. The *Luftwaffe* wiped out almost the entire Soviet air force (most of it on the ground), while the *Wehrmacht* encircled and destroyed numerous Russian armies who were still in the process of concentrating along the new border in recently carved up Poland (Erickson, 1999). Conversely, failing to prepare opens up the risk of disaster, which is what happened to the French in 1870 when Napoleon declared war full two weeks before his armies were ready to thrust into Prussia, a delay during which "trains from all over Germany poured uninterruptedly to the Rhine" shifting the advantage to the Germans, this time conclusively (Howard, 1981, 77).

Military moves may create extreme pressure during a crisis by restricting the amount of

[21] Van Evera (1999, Ch. 7) documents ample support for the notion that decision-makers feared even small delays in their mobilizations against those of their opponents during the July 1914 crisis.

time available for interaction—a feature that some authors actually include in the very definition of a crisis (e.g., Hermann, 1972, 10). Because of the costs, one cannot maintain a heightened state of readiness for too long, but since demobilization may render one quite vulnerable to a surprise attack, longer crises may provide the incentive to strike rather than risk exposure to that possibility (Schelling, 1966). Waiting also allows the enemy to prepare better to meet one's assault should war begin or even attack first, thereby lowering the expected probability of victory, which may also sustain a momentum for fighting. For example, in the last frantic days during the run-up to the Six Day War in 1967, Israel's military was increasingly worried about IDF's inactivity in the face of the significant Egyptian buildup in the Sinai. The General Staff argued that "every delay is a gamble with Israel's survival" and pressed for a preemptive strike (Oren, 2002, 98).

A first-move advantage differs from a *surprise attack*, where the advantage lies in catching the opponent off-guard and therefore less able to resist successfully (Kam, 2004). However, such potential vulnerability to surprise attack suggests that military moves that reduce the risk of being surprised would increase the expected payoff from war. On the other hand, one should remember that maintaining high readiness for a prolonged period of time may also result in alert fatigue with attending increased risks of being surprised as happened at Pearl Harbor in 1941 (Wohlstetter, 1962). This illustrates nicely the costs (fatigue) and benefits (preparedness) of military moves.

3.1.3 Military Moves Can Reveal Capability

It is worth noting that *military moves can, under certain circumstances, reveal one's strength unambiguously*. I discuss the strategic manipulation of the military instrument in the next section, here I wish to point out that a show of force may achieve something that words (or audience costs) never can: they can reveal to the opponent that its own estimate of expected yield from fighting is wrong because its perception of our capabilities is wildly incorrect. While words are cheap and everyone can claim superiority, actual demonstrations may be much more convincing as a non-manipulable "index" (Jervis, 1970). It is worth emphasizing that the believability of such a signal has nothing to do with its costs because it is not its affordability that convinces but the impossibility to mimic the capability being shown.

For example, during the Middle Ages, the Saracens besieged a fortress in Syria defended by a garrison of the Knights of Saint John. After several unsuccessful attempts to mine their way into the castle, they eventually managed to dig a mine all the way under the keep. They then invited the Franks to see the shaft, and when their own engineers reported that it would indeed collapse the keep when lit, the defenders surrendered and were allowed to withdraw from the castle unmolested (Gies and Gies, 1969, 197).

The high-speed missile test by Iran in April 2006 revealed its capability to strike targets within a wide radius, stoking Western fears about its intent to develop nuclear weapons that would be able to reach Europe.[22]

[22] Associated Press, April 2, 2006. Retrieved from http://www.usatoday.com/news/world/2006-04-02-iran-missile_x.htm?csp=34. There is also an apocryphal story, which is so good I wish I were able to verify it. During the Yom Kippur War of 1973, the Israeli air force supposedly paint-bombed the Aswan Dam to demonstrate its ability to destroy it as a last resort and flood the entire Nile valley if Israel's military situation became desperate.

Ironically, sometimes one need not have the actual capability as long as one can create the appropriate perception in the opponent. In the spring of 327 B.C., Alexander the Great arrived at the Rock of Sogdiana and discovered that the defenders of the fortress had stored provisions for a long siege, and that the sheer cliffs together with the deep snow made a direct assault impractical. He offered the Sogdianians safe passage if they surrendered, but they laughed at him fully convinced in the impregnability of their position, and ventured that only soldiers with wings could capture the rock. Alexander offered a prize of twelve talents to climbers who scaled the summit above the rock. Although thirty soldiers perished in the hazardous ascent, over two hundred Macedonians reached the peak and planted their battle flags. Alexander then informed the Sogdianians that he had, in fact, men with wings in his army and they had already captured the mound above their fortress, pointing to the plainly visible Macedonians on the top. As Arrian puts it, "the barbarians were astounded at a sight they had never reckoned on, and suspecting that the soldiers occupying the heights were more numerous and fully armed, they surrendered" (Arrian, 1976, 399–403).

In instances like these, the actual distribution of power does not change. Instead, the impossibility to mimic capability convinces one of the opponents that his original estimate of that distribution of power was wrong, which then causes him to reassess the desirability of fighting. Even though these military moves are physical (indeed, they have to be or else capability cannot be shown), they would not satisfy the second criterion above and as such are excluded from the present analysis.

To preview things to come, one major reason for the second requirement is that by affecting the adversary's payoff directly, the military instrument can potentially undo his commitment. As I noted before, one actor cannot counter its opponent's audience costs except by generating its own and increasing the risk of violent confrontation. However, one may be able to counter an enemy's military move, and either cancel out its commitment effect, or at least partially neutralize it. This is because one's expected payoff from war depends on the forces that the opponent can bring to bear, in terms both of the probability of obtaining a favorable military outcome, and of the ability to inflict costs in return. Even if one's military move increases the expected benefit from fighting (thereby creating a credible commitment to a bellicose course of action), the opponent may react with a military move of its own that would reduce one's expected benefit from fighting, therefore weakening one's commitment to such a course. An attempt to establish a credible commitment may thus trigger a counter-attempt to undo it, making the process very complicated because of this strategic interdependence of actions and their consequences. It also makes the process of creating commitments in such an environment an intriguing object of study, for which game theory is an especially suited tool for analysis.

3.2 Modeling Military Threats

The military instrument seems to be a good candidate for explaining the dynamic commitment creation process, both during a crisis and throughout the war that may follow. It has an informational role (because of its costliness and the risk it carries), but it also has the functional role missing from pure verbal bargaining. Our theories have not taken its empirical features properly into account because they do not allow actors to choose the level of costly

effort they invest in attempting to alter the environment to their advantage.[23] The closest thing we have is the economic literature on contests but it cannot be utilized readily because these models lack the essential decision to engage in the contest (fight) itself and the possibility to avoid it through concessions.[24] Also, these models are not designed to deal with information transmission (signaling) issues, which are critical for any crisis bargaining, and perhaps even war-fighting, model.

We now modify our basic crisis model just enough to incorporate the features of the military instrument identified in the previous section.[25] To facilitate the comparison with existing costly signaling and commitment models, I investigate the strategic use of the military instrument by a defender. In Chapter 5, I present a fuller theory that incorporates his opponent's initial decision to challenge the status quo.

3.2.1 The Technology of Conflict: Distribution of Power and System Militarization

In the traditional models we examined in Chapter 2, the distribution of power is summarized by p, the probability that S_1 will prevail if force is used, and remains static during the crisis. Since we want to explore the strategic use of the military instrument through its effect on that distribution, we must now define how military moves affect it. One straightforward way to think about the distribution of power is in terms of the distribution of military capabilities and how these translate into one's prospects of winning the war (Powell, 1993; Slantchev, 2005).

No state begins a potential crisis interaction without at least some existing military capabilities. Let $M_i > 0$ denote S_i's status quo military forces, and let $M = M_1 + M_2$ denote the total forces available to both actors. The status quo *distribution of capabilities* within the system is then denoted by (M_1, M_2) and we shall assume that it is perfectly observable. The *distribution of power* is determined with a simple formula: the probability of victory depends on one's share of the total military forces. Thus,

$$p = \frac{M_1}{M_1 + M_2}$$

is the status quo *distribution of power*; that is, the probability that S_1 would prevail in a war if no additional resources are mobilized. It is easy to see that $1 - p$ is then the probability that S_2 would prevail. Since the distribution of capabilities and their mapping into prospects for victory in war are common knowledge, the status quo distribution of power is common knowledge as well. Admittedly, the technology of conflict modeled through this ratio contest success function is exceedingly simple: it assumes that the marginal return from more capabilities is always diminishing, that there are no offensive or defensive advantages, that actors are equally effective in utilizing their capabilities, and so on. However, it is a common way

[23] Morrow (1994a) offers a model where the decision to ally incurs immediate costs but also enhances the war prospects. With military moves, the actors are also free to choose the level of commitment, and each can undermine the other's efforts.

[24] See Hirshleifer (1988, 2000) for applications of contests to conflict.

[25] This model is based on a one-sided incomplete information game that I analyzed for defense-dominant environments (Slantchev, 2005), and offense-dominant ones (Slantchev, 2004b).

of conceptualizing the link between capabilities and power and at the very least will serve as a useful baseline with which one could compare alternative richer specifications.[26]

In the models analyzed in Chapter 2, p remains constant throughout the crisis. However, as I have argued here, actors can affect the distribution of power, at least at the margins, through their behavior during the crisis. In particular, mobilizing additional military resources should improve an actor's prospects in war all else equal. Since mobilizing new forces essentially constitutes an addition to the existing base, one simple way of modeling this effect is to recalculate one's new share of the total forces. Thus, if S_1 allocates m new forces, his probability of winning will be:

$$p(m) = \frac{m + M_1}{m + M_1 + M_2}.$$

In contrast to the non-military models, the *intra-crisis distribution of power*, $p(m)$, is endogenous and for any $m > 0$ it will be different from the status quo distribution. To ease notation, it would be convenient to write $p_1(m)$ to denote the probability that S_1 wins, and $p_2(m) = 1 - p_1(m) = M_2/(m + M_1 + M_2)$ to denote the probability that S_2 wins.

Since p remains fixed in the traditional models, it is a sufficient statistic for the distribution of power and the precise distribution of capabilities is irrelevant. This turns out not to be the case in the military threat model. The marginal effect of additional mobilization depends on the existing allocations (and because its cost is fixed, its usefulness as an instrument will also vary), which implies that p is no longer sufficient to describe the strategic context and we must take into account the exact status quo distribution of capabilities.

To see that, observe that taking two different distributions of capabilities $(M_1, M_2) \neq (M_1', M_2')$ with their corresponding distributions of power p and p', it is quite possible to obtain $p = p'$. That is, the distribution of power may be the same even though the distributions of capabilities are different. For example, $(50, 100)$ and $(5, 10)$ both produce $1/3$ as the distribution of power but describe very different militarization levels within the system.

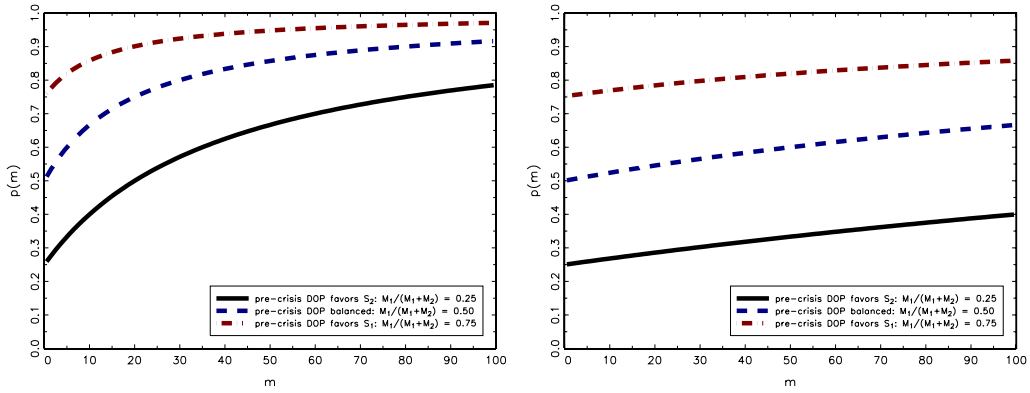
Our specification of the technology of conflict privileges additional mobilization when the status quo system militarization is lower. Given two distributions of capabilities that reduce to the same distribution of power, any given mobilization level will increase the probability of victory by a larger amount in the under-militarized system:

CLAIM 3.1 (More Bang for the Buck). *All else equal, the mobilization of additional forces is more effective when crisis participants are lightly armed.*[27]

It is perhaps illustrative to look graphically at how military moves can affect the distribution of power during a crisis using this technology of conflict and how their effectiveness changes with the level of system militarization. Figure 3.1 shows the marginal effect of additional mobilization of the distribution of power (the probability that S_1 will prevail in war). The left panel, Figure 3.1(a), exhibits a system where the adversaries are lightly armed, and the right panel, Figure 3.1(b), exhibits one where they maintain much higher armament levels. In either case, we look at three different status quo distributions of power: one which is

[26] See Hirshleifer (1991) for the contest success functions. Powell (1993) uses the ratio form in his model of the guns vs. butter trade-off.

[27] All formal statements and proofs are in Appendix B.



(a) Moderately Militarized System (b) Highly Militarized System

Figure 3.1 Marginal Effect of Mobilization on the Distribution of Power.

seriously advantageous to S_1 ($p = 3/4$), another in which the two actors are evenly matched ($p = 1/2$), and finally one in which the advantage belongs to S_2 ($p = 1/4$).[28]

As required, mobilizing more resources improves S_1's chances of victory but they do so at diminishing marginal rates. Two things must be noted here. First, the marginal effect of intra-crisis mobilization is much stronger in the moderately militarized system for any given status quo distribution of power. For example, if opponents are evenly matched at the outset, so $p = 1/2$, mobilizing $m = 30$ additional units produces a jump to $p(m)$ of approximately 0.75 if opponents are lightly armed but only to 0.55 if they are heavily armed. Second, in addition to the precise status quo distribution of capabilities, the benefit also depends on the status quo distribution of power itself. As Figure 3.1(a) illustrates, S_1 can more than triple his chances of victory from $p = 0.25$ with $m = 100$, but can only improve them by about 25% from $p = 0.75$ by mobilizing at that level.

I must emphasize that these figures illustrate the effect of mobilization on the distribution of power, not on S_1's expected payoff from war. Through its positive effect on the probability of victory, mobilization will tend to increase that payoff as well. However, its costliness will tend to decrease it. Because of the diminishing marginal returns on the distribution of power and constant marginal costs, it follows that in general it will not be optimal for S_1 to choose the highest possible mobilization level.

We now have specified the technology of conflict that allows us to endogenize the distribution of power through the use of the military instrument. In doing so, we noted that the relative direct effectiveness of this instrument depends on the precise distribution of capabilities in the system. Whereas different levels of system militarization will not affect any of the results in the traditional models with their exogenously-specified and static distribution of power, the military instrument is more attractive when the overall armament levels are low, and correspondingly the incentive to use it will vary with the status quo distribution of

[28] I have arbitrarily fixed $M_2 = 10$ for the moderately militarized system and $M_2 = 100$ for the highly militarized one. The values of M_1 are then derived to produce the three status quo distributions of power.

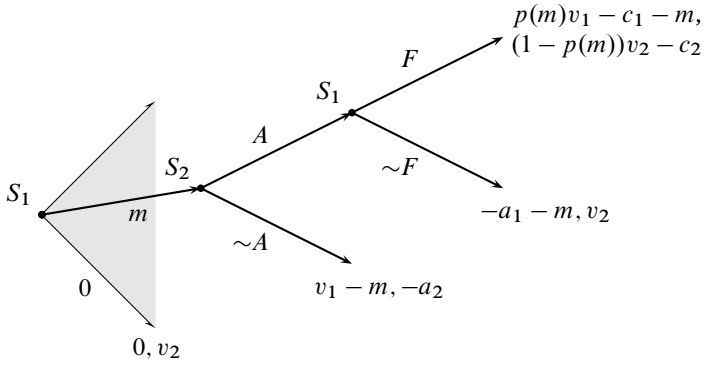


Figure 3.2 The Crisis Game with the Military Instrument.

capabilities as well. With these preliminaries in mind, we now turn to the specification of the military threat model.

3.2.2 The Military Crisis Game

As before, two states, S_1 and S_2, face a potential dispute over territory currently in possession of S_1, the "defender." Each player knows how much he values possession of the territory but neither knows his opponent's valuation. S_1 believes that the potential revisionist's valuation, v_2, is distributed uniformly on the interval $[t, u]$ with $0 \leq t < u$, and S_2 believes the defender's valuation is distributed uniformly on the interval $[0, \overline{v}_1]$. These distributions are common knowledge.[29] Assume that $u > c_i$ or else players would never fight.

The game begins with the defender's choice to issue a military threat by mobilizing more of his own forces, m, or appease his opponent peacefully. Mobilizing $m > 0$ is equivalent to escalation and its costs are immediately incurred; that is, they are sunk regardless of how the crisis ends. After observing S_1's mobilization level, S_2 must decide whether to act or capitulate. Capitulation ends the game with the peaceful preservation of the good in the defender's possession and the challenger suffers her audience costs. Resistance leads to S_1's final decision whether to fight or back down. If he backs down, the crisis ends and the good is peacefully transferred to S_2 while the defender incurs the audience costs. If he attacks, war begins, and the probability of victory is a function of the existing mobilization levels. Figure 3.2 shows the game tree.

The payoffs are as follows. If player S_1 retains the good peacefully, his payoff is $v_1 - m$, and if he relinquishes the good, his payoff is $-m - a_1$ if he escalates and capitulates, and 0 if he appeases S_2. That is, given a choice between appeasement and certain capitulation following escalation, he strictly prefers to appease even if he escalates without additional mobilization, $m = 0$. Player 2's payoff is v_2 if she receives the good peacefully (whether

[29] To compare and contrast results from this model to those in Chapter 2, I will use $t = 0$ and $u = \overline{v}_2 > 0$. However, I will derive the results in terms of an arbitrary interval because I will use them in Chapter 5 where this interval will change depending on S_2's initial decision.

by appeasement or by S_1's capitulation) and $-a_2$ if she capitulates following the defender's threat. The expected payoff from war is $p_1(m)v_1 - c_1 - m$ for S_1 and $p_2(m)v_2 - c_2$ for S_2, where $c_i > 0$ are the additional costs of war paid over the mobilization costs.

In keeping with the argument that audience costs cannot be arbitrarily high, especially in comparison to the costs of war, I will make the following assumption, whose effect is to guarantee that the defender does not have an inherently credible commitment which he could use by escalating without additional mobilization.

ASSUMPTION 3.1. The costs of war exceed audience costs: $c_i > a_i$.

As usual, this model is a reflection of the trade-off between realistic approximation of the empirical and a tractable stylization that we can use for transparent analysis. The specification of the payoffs formally captures the two fundamental characteristics of the military instrument. Mobilizing m is inherently costly for S_1 because he has to pay the price regardless of how the crisis turns out. However, it also changes the distribution of power and as such affects the war payoffs for both players directly. Some of the other simplifying assumptions are not that important, and I have made them mostly to be able to derive analytic solutions that I can do comparative statics on.[30] Even S_1's advantage that he can mobilize after knowing S_2's mobilization (and can therefore "tailor" his as much as possible) is somewhat balanced by the disadvantage of having to move last (in many situations the one saddled with the final choice for war is in a weaker bargaining position).

3.3 Threats with Complete Information

It will be instructive to look at this crisis game under complete information, so assume that both players' valuations are common knowledge. As before, we shall restrict attention to subgame-perfect equilibria (SPE, or "equilibrium") in which players can only make credible threats. As the reader probably already suspects, equilibrium behavior will be entirely dependent on the credibility of S_1's threat to fight when resisted and the credibility of S_2's threat to defy a military threat. By subgame perfection, S_1 will attack if resisted when the expected payoff from war given his mobilization m is at least as large as his payoff from capitulating. Formally, S_1's threat $m \geq 0$ is *credible* if, and only if, $p(m)v_1 - c_1 - m \geq -a_1 - m$, which simplifies to:

$$p(m) \geq \frac{c_1 - a_1}{v_1}. \qquad (\mathrm{CR}_1)$$

This expression makes three things immediately apparent. First, if S_1's war costs are too high, then no size of additional forces can ever make his threat credible. In particular, if $c_1 > v_1 + a_1$, then (CR_1) cannot be satisfied for any $m \geq 0$ because the right-hand side will be strictly greater than 1, and $p(m)$, as a valid probability, cannot exceed unity. Second, S_1

[30] The uniform distributions allow us to compute analytic solutions. The results could be derived with arbitrary distributions that specify certain regularity properties: non-decreasing hazards, and densities that do not increase "too quickly". The first property is common to many usual distributions. The second would ensure that optimal credible mobilization is increasing in type; that is, that higher types strictly prefer higher mobilization levels if S_2 is convinced of their resolve. Putting it another way, this ensures that an arbitrarily small increase in m does not drastically increase S_2's probability of capitulating. The uniform distribution has a constant density, so this is not an issue.

may be *inherently resolved* if (CR_1) is satisfied at $m = 0$. In this situation, S_1's threat to fight will be credible at the status quo distribution of power. As we shall see, even when this is true, S_1 may still find it necessary to mobilize additional resources for coercive purposes. Finally, when S_1 is not inherently resolved, he can *create* a credible commitment to fight when resisted provided his war costs are not too high. To see this, suppose that $c_1 < v_1 + a_1$, so it is possible to satisfy (CR_1) for some m. Since S_1's probability of victory, $p(m)$, is strictly increasing in his mobilization, it follows that for m high enough, (CR_1) will be satisfied. Even if S_1 starts out at a serious disadvantage at the status quo distribution of power—a situation in which he would have been compelled to appease S_2 in the original crisis game in Figure 2.1—it may be possible for him to redress the imbalance by mobilizing additional forces and establishing a credible commitment to fight.

The rest of the analysis now proceeds in a series of steps. An outline of the logic makes the argument easier to follow. I show that:

1) In any equilibrium in which S_1 makes a threat $m \geq 0$, he must be resolved at m. This implies that in any equilibrium in which S_2 is confronted with a threat, she will know that resistance will lead to certain war.
2) When S_2 knows that resistance means war, S_1 can undermine her commitment by mobilizing sufficient additional forces, and compel her to capitulate with certainty.
3) This implies that in any equilibrium in which S_1 makes a threat, the outcome will be either certain war or certain capitulation by S_2.
4) Because S_1's optimal mobilization is unique depending on whether he is preparing for war or compelling S_2 to capitulate, it follows that in the unique subgame-perfect equilibrium, S_1 will either appease S_2, prepare for and fight a war, or make a compellent threat that causes S_2 to yield.

Recall that if $c_1 \geq v_1 + a_1$, there exists no m that can satisfy (CR_1), and S_1 will capitulate when resisted regardless of the mobilization level he has chosen. This now means that S_2 will certainly resist when threatened, which implies that S_1 will never threaten in the first place because doing so would lead to certain capitulation in the endgame, which is strictly worse than appeasement. In this unique SPE, S_1 will appease S_2 immediately.

Assume, then, that $c_1 < v_1 + a_1$ for the rest of this proof. Observe that in any SPE in which S_1 does not appease, he must be resolved at the mobilization level m. If that were not the case, then he would never attack when resisted. Subgame-perfection then implies that S_2 would certainly resist the escalation, causing S_1 to capitulate and incur both audience and mobilization costs. This outcome is worse than immediate appeasement, which contradicts the supposition that S_1 does not appease in this equilibrium.

This now implies that in any SPE in which S_1 makes a threat, S_2 knows that resistance means certain war. In that case, she resists only when her expected payoff from war is strictly better than her payoff from capitulating. Formally, she resists if, and only if, $(1 - p(m))v_2 - c_2 > -a_2$, or whenever:

$$p(m) < 1 - \frac{c_2 - a_2}{v_2}. \qquad (CR_2)$$

This expression makes two things clear. First, if S_2's war costs are too high, she will capitulate for any $m \geq 0$ that satisfies (CR_1). In particular, if $c_2 > v_2 + a_2$, then (CR_2) cannot

be satisfied for any $m \geq 0$ because the right-hand side will be strictly negative, and $p(m)$, being a valid probability, cannot be less than zero. Because (CR_1) is satisfied, she knows that resistance would lead to war, so she will capitulate for sure. Second, and more importantly, even if S_2 is resolved at some $m' \geq 0$, S_1 can undermine her commitment by mobilizing sufficiently many forces. Intuitively, this happens because larger mobilizations by S_1 mean worse prospects for victory in the war for S_2; that is, $p(m)$ increases in m. Hence, for m large enough, (CR_2) will fail and S_2 will certainly capitulate even if she was resolved at $m' < m$. Even if S_2 starts out in a favorable position and is highly committed, S_1 can undermine her resolve by mobilizing additional forces (provided they also commit him to fight).

These results imply that in any equilibrium in which S_1 makes a threat, he either goes to war (because m is small enough for (CR_2) to be satisfied) or ensures S_2's capitulation. Let $\overline{m}$ denote the *assured compellence* mobilization for S_1; that is, $\overline{m}$ satisfies (CR_1) but not (CR_2), and therefore compels S_2 to yield. There can be only one such level in equilibrium because if there were more, S_1 would deviate to the smallest one and improve his payoff.

If S_1 is unwilling to mobilize resources sufficient to compel S_2 because $\overline{m}$ is too high to make it worth it, he may prefer to fight rather than appease her. In that case, he will mobilize m such that both (CR_1) and (CR_2) are satisfied. Making such a threat would lead to certain war, and the best S_1 can expect then is $p(m)v_1 - c_1 - m$. Let m^* denote the *optimal war* preparation for S_1—that is, it maximizes his expected payoff from war—and note that because the payoff is concave in m, m^* is unique.

We now conclude that in any equilibrium, S_1's initial choice reduces to a selection from three possibilities: appease S_2 immediately, prepare for certain war by mobilizing m^*, or ensure S_2's capitulation by mobilizing $\overline{m}$. Because the optimal mobilization levels are unique, the SPE is also unique. The following proposition, whose formal proof is in Appendix B, states these conclusions.

PROPOSITION 3.1. *The military threat game has a unique subgame-perfect equilibrium in pure strategies. In it, S_1 either appeases immediately, mobilizes uniquely at m^* for certain war, or mobilizes uniquely at $\overline{m}$ for certain compellence.*

The analysis of the military threat game under complete information leads to conclusions that presage several of the major themes of this book.

RESULT 3.1 *The military instrument can be useful to the actor making it because it sometimes allows him to compel the opponent's capitulation in circumstances when he would not have been able to do so at the status quo distribution of power without mobilizing additional resources. Moreover, the military instrument can be stabilizing because a military threat can sometimes reduce the probability of war relative to the crisis with non-military escalation.*

It will be instructive to examine numerical illustrations of these results. Fix $v_2 = 22$, $a_1 = a_2 = 0.5$ and $M_1 = M_2 = 2.5$ for all examples that follow. (All numbers rounded to the second decimal point for clarity.)

EXAMPLE 3.1 (Compellence Succeeds Only When Militarized). Let $v_1 = 22$, $c_1 = 12$, and $c_2 = 4$. Consider first a situation in which S_1 could not mobilize additional resources and had to choose between appeasing S_2 and escalating at the status quo distribution of power. His expected payoff from war at $m = 0$ is -1, which is worse than his expected payoff from capitulation (-0.5). S_1's threat is not credible, and S_2 will certainly resist when threatened. As a result, S_1 appeases immediately.

Consider now the scenario in which S_1 can make a military threat. Here, S_1 would attack at any $m \geq$

0.24, and S_2 would resist any $m < 10.71$ even when doing so would lead to war. Hence, by mobilizing $m = \overline{m} = 10.71$, S_1 can ensure S_2's capitulation. The payoff from this is $v_1 - \overline{m} = 11.29$, which is strictly better than appeasement. To see that S_1 would also prefer the assured compellence of S_2 to war with her, observe that the optimal war mobilization is $m^* = 2.42$ (which S_2 certainly resists), and S_1's expected payoff is 0.17. Even though this is also preferable to appeasement, it is worse than compellence. Hence, in the unique SPE, S_1 makes a militarized threat, and S_2 capitulates.

This example illustrates a situation in which S_1's threat would not have been credible at the status quo distribution of power, and which would have caused him to appease S_2. When he can militarize the crisis, however, S_1 can make himself resolved by mobilizing additional resources. Moreover, he can also undermine S_2's resolve sufficiently to compel her certain capitulation. From S_1's perspective, the military instrument is doubtless very useful.

As the example above demonstrates, the ability to militarize the crisis may be quite beneficial (under some circumstances) for the actor that can resort to such moves. However, the effect can also be positive from a social (collective) perspective as well. To see that, imagine a situation in which escalating at the status quo distribution of power would lead to war in equilibrium. (Obviously, this requires us to relax Assumption 2.1, which I will do for the sake of making the point.) Militarizing the crisis, however, may well lead to a peaceful outcome, which both actors prefer to war (even though it is unpleasant for S_2 because she is forced to capitulate).

EXAMPLE 3.2 (Militarization Can Be Stabilizing). Let all variables be as in Example 3.1 except $c_1 = 6$ (that is, all we have done is halve the costs of war to S_1). If S_1 were unable to mobilize additional resources to make a military threat, his payoff from war would be 5.0, which is strictly better than appeasement. Since S_2's payoff from war is 7.0, she would resist even if that leads to war because capitulating would yield only -0.5. Even though S_1's threat is credible, S_2's valuation is too high and the costs of fighting too low to compel her successfully. As a result, S_1 escalates, she resists, and he attacks. War is certain in that equilibrium.

Consider now the scenario in which S_1 can militarize the crisis. Since war even at $m = 0$ yields a strictly positive payoff, he would clearly attack at any $m \geq 0$ as well. As before (because the values of the parameters for her are the same), S_2 would capitulate at any $m \geq 10.71$. Compellence with $\overline{m} = 10.71$ yields S_1 a payoff of 11.29. Optimal war with $m^* = 2.42$ (which S_2 resists), on the other hand, only yields 6.17. Whereas this is obviously better than his payoff from war under the status quo distribution of power, it is much worse from the payoff from assured compellence. Therefore, in the unique SPE, S_1 makes a militarized threat at $\overline{m}$, and S_2 capitulates.

Recall that the only difference between Example 3.1 and Example 3.2 is that the costs of fighting for S_1 are much lower in Example 3.2. Compellence "costs" the same and is just as attractive as before, it is only war that has become much more attractive to S_1 (although it remains worse than compellence). The confrontation between two actors with relatively high valuations and low costs of fighting can lead to war unless S_1 can make a military threat. Mobilizing additional resources allows him to improve his war payoff but, more importantly, also undermines S_2's war payoff. If he mobilizes enough, she would capitulate. Assured compellence is preferred by both players to war, and in that sense peace is socially beneficial. Militarization can thus be stabilizing because it reduces the probability of war by making it sufficiently unattractive to the opponent.

These two examples make a crucial general point. Mobilization affects the war and capitulation payoffs for S_1 in different ways. Under certain conditions it may improve the war payoff sufficiently to make fighting preferable to capitulation. That is, S_1 can make himself resolved by mobilizing enough additional resources. More importantly, mobilization also

affects the war payoff of S_2 directly. Under certain conditions it may attenuate it sufficiently to make capitulation preferable to fighting. That is, S_1 can vitiate S_2's resolve by mobilizing enough additional resources. This result is central to the book, and is worth emphasizing it:

RESULT 3.2 *The military instrument can* create *a credible commitment for an actor where he had none and can* undermine *the commitment of his opponent.*

Before we get carried away with enthusiasm for the military instrument, I should note that the beneficial effects under some conditions must be tempered with rather serious disadvantages under other conditions. For instance, the ability to militarize the crisis could turn out to be destabilizing and lead to war in circumstances where inability to do so would guarantee peace at the status quo distribution of power.

EXAMPLE 3.3 (Militarization Can Be Destabilizing). Let all variables be as in Example 3.1 except $c_2 = 2$ (that is, all we have done is halve the costs of war to S_2). If S_1 were unable to mobilize additional military resources, his payoff from war remains at -1, and since this is worse than -0.5 (which is what he would get by capitulating), his threat is incredible, and in the unique SPE he must appease S_2 immediately. War will never happen in this equilibrium.

Consider now the scenario in which S_1 can militarize the crisis. To commit himself credibly to war, he would have to mobilize at least $m \geq 0.24$. However, because S_2's war costs now are much lower, her expected payoff from fighting at any given m is much higher, and consequently it takes a lot more to coerce her to capitulate. In this case, she would only do so if $m \geq 31.67$. Although S_1 can ensure her capitulation, doing so is too expensive: his payoff would be $v_1 - \overline{m} = -9.67$, which is strictly worse than appeasement. In other words, S_1 would rather give up the good than pay to compel S_2 to give up her claim to it. On the other hand, appeasement is worse than optimal war at $m^* = 2.42$, which would yield a payoff of 0.17. Since S_2 will resist this allocation, escalating at this level causes certain war. Hence, in this unique SPE, S_1 makes a militarized threat that commits him to war but fails to undermine S_2's commitment. As a result, the equilibrium outcome is certain war.

Compare this example to Example 3.1: the only difference is that the costs of war for S_2 in Example 3.3 are half her costs in Example 3.1. This means that her expected war payoff would be higher at any given mobilization by S_1, which in turn implies that it would take a lot more to compel her to capitulate. In this case, assuring compellence is so expensive that it makes it less attractive than appeasement. On the other hand, S_1's optimal war payoff is the same as before, and since this is better than appeasement, he escalates and the crisis ends in war. When S_2's resolve is very high, the additional military resources S_1 must bring to bear to undermine her commitment may be prohibitively expensive. However, because S_1 can still make himself resolved, he may make fighting too attractive an option. In that case, he will opt for war to avoid appeasing his opponent. Although the military instrument is still useful to S_1, its costliness prevents him from ensuring the peaceful submission of S_2.

Compare now this example to Example 3.2: the two differences are that S_1's costs of war in Example 3.3 are twice his costs in Example 3.2 and S_2's costs of war in Example 3.3 are half her costs in Example 3.2. Because of the higher costs of war in Example 3.3, S_1 does not have a credible commitment to fight at the status quo distribution of power, which means that peace is certain to prevail unlike the scenario in Example 3.2. On the other hand, S_2's lower costs of war in Example 3.3 also make it exceedingly costly to compel her capitulation. Coupling this with S_1's ability to make himself resolved with additional mobilization destabilizes the crisis leading to war. To summarize:

RESULT 3.3 *The ability to create a credible commitment may be destabilizing when the opponent's resolve*

is so high that it is prohibitively costly to undermine it. In that case, an actor can mobilize optimally for war and fight it to avoid appeasing the opponent.

The costliness of the military instrument is problematic when the opponent is so committed that the required compellence mobilization is prohibitively expensive. As we have seen, in this case it is entirely possible that the actor would instead prepare to fight a certain war. The costliness, however, is also a factor to be reckoned with even when it is possible to profit from compelling the opponent's capitulation. Although in that case peace will prevail in equilibrium, as it does in Example 3.1 and Example 3.2, the resulting mobilization can be quite large. For instance, in both examples S_1 spends nearly 50% of the value he attaches to the disputed good to ensure that S_2 will give it up peacefully. The reason in either case is that S_2 happens to have somewhat low costs of fighting, which make war relatively attractive, which in turn means that S_1 has to mobilize a lot of additional resources to undermine her commitment.

The price of peace, however, can also be quite high even when the opponent is relatively weak, as the following example demonstrates.

EXAMPLE 3.4 (Over-Mobilization for Commitment). Let $v_1 = 12$, and $c_1 = c_2 = 10$. Here, the costs of fighting are relatively high for both players, and S_1's valuation is much lower than before. Given the high costs of war, S_1 would need to mobilize at a high level to make fighting preferable to capitulation. In this case, he would only fight if $m \geq 7$. On the other hand, because S_2's costs of war are also high, it does not take that much to undermine her commitment: she would capitulate whenever $m \geq 0.79$. Note now that it takes a lot less to compel her than to establish a credible commitment for S_1. This should not be surprising: after all, S_1's valuation is also relatively low, which in turn means that the probability of victory in war must be quite a bit larger to make fighting sufficiently profitable. This situation creates an interesting conundrum for S_1: since S_2's resolve is shaky, it appears that he could mobilize $m = 0.79$ and compel her to capitulate. Unfortunately, this is not going to work: at such a low level of preparedness, his threat to fight when resisted is not credible, and S_2 would certainly resist that escalation. The minimal mobilization level that would compel her successfully is $m = 7$ because this is the smallest allocation at which S_1's threat will be credible. Mobilizing at this level would ensure S_2's capitulation and yield S_1 a payoff of 5, which is still better than appeasement. It is also better than war because at the optimal mobilization for war at $m^* = 0.48$ (which S_2 will certainly resist), S_1's expected payoff from fighting is -3.96. (Obviously, because $m^* < 7$, S_1 will not actually fight at this allocation. But since this is the best he can, in principle, obtain in war, any other allocation would yield even worse payoffs. For example, if he fought with $m = 7$, it will be -7.5.) Hence, in the unique SPE, S_1 makes a military threat that undermines S_2's commitment and compels her capitulation with certainty. The equilibrium outcome is peace.[31]

Notice an intriguing aspect of strategic mobilization: S_1 is forced to "over-prepare" for war in this example. S_2 seems relatively easy to compel—after all, it does not take much to get her to capitulate with certainty provided she also expects that resistance would lead to war. And herein lies the crux of the problem for S_1 for at such a low level of preparedness he will not actually fight if she resists. This lack of credibility, in turn, destroys the compellent function of the threat and S_2 would surely resist such an escalation. By mobilizing at the much higher level, S_1 is essentially paying for his own lack of commitment. Because his

[31] The equilibrium is also peaceful if S_1 cannot use the military instrument but it differs qualitatively in the resulting distribution of benefits. Fighting at the status quo distribution of power yields S_1 a payoff of -4, which is worse than capitulating (-0.5), so S_1 does not have a credible threat to fight. Consequently, S_2 would resist any escalation, which in turn implies that S_1 will immediately appease her. Even though the equilibrium outcome is peace, it also entails a revision of the status quo in S_2's favor whereas when militarization is possible S_1 can successfully retain the good (at a price).

valuation for the issue is relatively low, he cannot credibly threaten to fight for it unless he mobilizes enough additional resources to make victory very likely. Given the weakness of the opponent, this behavior may appear puzzling and especially wasteful. However, the logic of coercion is merciless: without this extravagant preparation for war, S_2 would disregard the threat and S_1 would be forced to appease her. In this scenario, the military instrument's main function is to create S_1's credible commitment rather than undermine S_2's resolve. Again, this makes peace quite expensive and without the logic of the model S_1's behavior may strike one as irrational and unnecessarily aggressive. Unfortunately, it is precisely this "overkill" capability that is required here to ensure peace through the capitulation of the opponent. These conclusions seem sufficiently intriguing to warrant their emphasis:

RESULT 3.4 *Even though the military instrument can help achieve a peaceful resolution of the crisis, the price of that peace can be very high because the player must mobilize sufficiently many additional resources to compel the opponent's capitulation. The stronger the opponent, the costlier the peace. When the opponent is not too resolved, an unresolved player will have to mobilize for "overkill" capability in order to create a credible commitment, without which the opponent would not capitulate.*

This result suggests that, ironically, S_1 will have to over-mobilize and behave quite aggressively in a situation in which he does not particularly care about the issue and where his opponent's costs of war are so high that he appears very easy to compel. The run-up to the Second Persian Gulf War in 2003 may be a striking instance of that dynamic. Although resolving its problems with Saddam Hussein was somewhat important to the US government, it certainly cannot be considered to have been a vital security issue.[32] The war would also be quite costly for the US because of the need to transport troops over large distances, maintain and supply them in a faraway theater of operations with minimal allied support over a long period of time, and do all of this when the economy was on a downward trend. Even without considering the costs of the occupation that followed, war was clearly going to be a costly enterprise. On the Iraqi side, Hussein's valuation of the issue must have been extremely high: yielding to the American demands would have destabilized his regime, undermined whatever prestige he had in the Arab world, and perhaps even exposed Iraq's actual weakness, a consequence of the First Persian Gulf war from which it had never recovered. Furthermore, war itself would be exceedingly costly given US capabilities and the fact that defeat would spell the end of his regime and probably his life as well.

This, then, appears to match the example's parameter configuration quite closely: an actor with a moderate valuation of the issue and high costs of war must make a military threat against an opponent whose valuation and costs of war are both very high. Given the military superiority of the US it should not have taken a very strenuous preparation to undo Hussein's commitment to resist, *provided that he believed the US would actually fight at that level of preparedness*. But at such low level of preparedness, the moderate valuation of the US would have left it uncommitted, rendering the threat incredible. As a result, the US had to engage in much more aggressive behavior and mobilize what appeared an excessively large amount of force to deal with such a weak opponent. The logic of coercion rationalizes such behavior.[33]

[32] This is declarations of Iraq being a "clear and present" danger notwithstanding.

[33] Since we have not considered the post-war occupation, the level of forces are only relevant for the war-fighting phase. As it turned out, a force sufficient to win a war may prove insufficient to create the post-war peace.

In the actual crisis the compellent threat did not work because even though the US had mobilized enough to commit itself to war, it had not mobilized enough to undermine Hussein's commitment. As we shall shortly see, this is exactly what can happen when S_1 is uncertain of his opponent's valuation and is unsure how much to mobilize to ensure S_2's capitulation. The costliness of the military instrument may cause him to balance the risk of war against the benefit of economizing on preparation. He ends up mobilizing less than what it would take to compel a highly committed opponent (but more than what it would take to compel a moderately committed one), so if it so happens that he actually faces such a resolved enemy, the equilibrium outcome is war.

One final observation to make here is that whereas S_1's optimal mobilization for war is a function of his valuation and costs of war, his compellent mobilization is much more involved. First, he must marshal enough resources to commit credibly to fighting when resisted. As we have seen, sometime this may involve substantial overkill capability given the type of opponent he is facing. Second, when he is not forced to over-mobilize to compensate for the inherent lack of credibility of his own commitment, the optimal mobilization level depends on *his opponent's* valuation, costs of war, and audience costs. The actual behavior of S_1 would reflect the type of opponent he faces rather than his own type.[34]

This dependence is critical for it implies that it will affect S_1's behavior when he is uncertain about the opponent's type. In that scenario, S_1 will not know just how much he must mobilize to ensure S_2's capitulation at the lowest possible cost. This opens up a whole new can of worms because now the desire to avoid unnecessary expense may cause S_1 to fail to mobilize at a level that is sufficiently high to compel S_2 successfully. In turn, S_2's uncertainty about the credibility of S_1's commitment may cause her to resist threats in the hope that S_1 is actually bluffing. In both cases, their actions will lead to war for reasons that are entirely separate from the optimal war that occurs in the model under complete information. In particular, the probability of war could be strictly positive in circumstances when the players would have avoided it for sure had they been completely informed from the outset. This means that uncertainty may lead to *war with regret*—war that players would not have fought if they were not asymmetrically informed about their valuations. It is to examine these possibilities that we now must turn to the analysis of the military threat game with incomplete information.

3.4 Threats Under Uncertainty

As it turns out, the analysis of the military threat game under incomplete information essentially amplifies the conclusions from the complete information case, with some intriguing nuance. The analysis itself is quite tedious because it involves handling every possible configuration of parameters. In the interest of clarity of exposition, I present the formal details in Appendix B and here provide the main logic of the argument along with results that are of most significance from theoretical and substantive perspectives. As in the previous chapter, the solution concept is perfect Bayesian equilibrium in pure strategies, or henceforth "equi-

[34] This is why the optimal mobilization for war is always smaller than the optimal mobilization for coercion, as we shall see later. It is worth exploring militarized coercion when one's expected payoff from war also depends on the opponent's valuation (e.g., because the latter influences how hard she will fight).

librium." Recall that this equilibrium requires that strategies are sequentially rational given beliefs, and beliefs are consistent with the strategies.

As in the complete information case, S_1 will attack if, and only if, (CR_1) is satisfied. The problem, of course, is that because S_2 does not his valuation, she may be unsure about the credibility of his threat. She will, therefore, attempt to infer S_1's level of commitment from his observable behavior—in our cases, this implies that S_2 will attempt to infer his valuation from his mobilization effort. Let $v_1^*(m)$ denote the *least committed* (or least resolved) type of S_1 at the allocation m; that is:

$$v_1^*(m) = \frac{c_1 - a_1}{p_1(m)}, \tag{3.1}$$

where I have rewritten (CR_1) and solved for v_1. As one would expect, $v_1^*(m)$ increases in c_1 and in M_2, which means that the higher S_1's costs of war and the more advantageous the status quo distribution of power is to S_2, the more S_1 would have to mobilize at any given valuation to commit himself. On the other hand, it is decreasing in m, which implies that the larger the mobilization, the lower the minimum valuation S_1 must have to become resolved to fight. Even a relatively low-value type can commit himself to war if he mobilizes enough additional resources.

The fact that S_1's credibility depends on his mobilization creates a serious problem for both actors because if S_2 remains uncertain about his valuation, she may not know whether the mobilization effort she has observed is, in fact, truly committing for her opponent. Because mobilization is costly, S_1 clearly has incentives to economize on it as much as possible. From S_2's perspective, this may translate into attempts to bluff her into submission by allocating fewer resources than he actually requires to commit to fighting. This may tempt her to resist and risk the possibility that her opponent is fully resolved, causing war in the process. None of this was an issue in the complete information case where S_2 knew with certainty whether S_1's threat was credible.

With uncertainty, when S_2 observes a mobilization m, she knows that if she resists, S_1 will attack if his valuation is $v_1 \geq v_1^*(m)$, and capitulate otherwise. Let $G_1(v_1^*(m)) = \Pr[v_1 \leq v_1^*(m)]$ denote her belief that S_1 will capitulate when resisted at m (that is, the probability that $v_1 \leq v_1^*(m)$). If she resists the mobilization m, then she will obtain S_1's submission with probability $G_1(v_1^*(m))$ and will end up causing war with probability $1 - G_1(v_1^*(m))$. Therefore, she will resist if, and only if, the expected payoff from resistance is greater than the payoff from capitulation, or when $G_1(v_1^*(m))v_2 + \left[1 - G_1(v_1^*(m))\right]\left[(1 - p(m))v_2 - c_2\right] > -a_2$. Solving this for S_2's valuation yields $v_2^*(m)$, the smallest valuation at which S_2 will resist given her beliefs and the observed mobilization by S_1:

$$v_2^*(m) = \frac{(1 - G_1(v_1^*(m)))c_2 - a_2}{1 - p(m) + p(m)G_1(v_1^*(m))}. \tag{CR_2'}$$

This, in essence, is the analogue to (CR_2) under uncertainty. S_2 resists m if her valuation is greater than $v_2^*(m)$, and capitulates otherwise. As it happens, this condition can be simplified quite a bit for all but one type of equilibrium in this game. As I will now show, in any equilibrium with positive risk of war, S_1's threat must be credible, which implies that whenever S_2 observes the equilibrium mobilization levels, she will know for sure that S_1 will attack when resisted because $v_1 \geq v_1^*(m)$. This implies that $G_1(v_1^*(m)) = 0$, which reduces (CR_2')

precisely to (CR_2) under complete information. In other words, *optimal strategic behavior* by S_1 will reveal sufficient information about his valuation to establish the credibility of his threat to fight in any equilibrium in which his threat leads to a positive probability of war. This echoes results from the sunk-cost and tying-hands models in Chapter 2: bluffing is not an issue in equilibrium in the sense that S_1 can, and does, credibly signal his resolve. The only time when he does bluff is when this carries no risk of war at all, a result new to the military threat model that I will investigate at some length later in this chapter.

RESULT 3.5 *The military instrument's potential to create a credible commitment for an actor and undermine the commitment of his opponent is fully utilized under uncertainty: risky military threats will always be credible.*

To establish the crucial result that risk-inducing threats must be credible, we begin by observing that if S_1 mobilizes m in some equilibrium, then he must attack when resisted with positive probability. If that were not so, then S_2 would resist him for sure. Because by supposition S_1 will not attack after m, it follows that if he chooses to threaten with m, he will have to capitulate with certainty in the endgame. But then S_1 strictly prefers to appease S_2, contradicting the equilibrium supposition that he threatens her with m. This implies that only two general types of behavior can be observed in any equilibrium in which S_1 mobilizes: either S_2 capitulates for sure (assured compellence) or she resists with positive probability and when she does so, war occurs with positive probability. The next step establishes the stronger claim that in any equilibrium that does not lead to assured compellence, her resistance will *certainly* to lead to war. Lemma 3.1 states this result, and because the proof is instructive (and short), I present it here in its entirety.

LEMMA 3.1. *Military threats that induce a positive probability of war are credible in any equilibrium with plausible beliefs.*

Proof A belief is *plausible* if an unexpectedly large out-of-equilibrium mobilization causes S_2 to conclude that S_1 is resolved.[35] Consider now some equilibrium threat m after which there is a positive risk of war. If S_2 were certain to resist such a threat, then S_1 would never bluff: any type that does so would have to capitulate later on and incur both audience and mobilization costs. Therefore, the credibility of the threat could only be a problem if S_2 resisted with positive probability but was not certain to do so. This now implies that there are valuations for which S_2 does not resist m in the support of S_1's beliefs. By sequential rationality, S_2 capitulates when $v_2 \leq v_2^*(m)$, and because such types exist, it follows that the least valuation type to resist is $v_2^*(m)$, which is indifferent between capitulation and resistance.

Suppose now that there is an equilibrium with a positive risk of war after m, and where some of the types mobilizing at m are bluffers. It cannot be the case that all of them are because in any equilibrium in which S_1 mobilizes m, there must be some types that are resolved at m. If this were not true, then S_2 would infer that he is not committed, and would resist for sure regardless of her valuation, forcing any bluffer to capitulate for sure. But this means that bluffers would strictly prefer to appease her immediately, contradicting the

[35] Plausible beliefs are formally defined and explained in Appendix B. Since I will only consider equilibria with plausible beliefs, I will henceforth refer to them as equilibria.

equilibrium supposition. Therefore, in any equilibrium in which S_2 resists m with positive probability less than one, bluffers must be pooling with resolved types on the common threat.

Consider now the highest-valuation type among the ones that pool and note that he must be resolved at m because if any type in that set is resolved at that allocation, then so must he. If he deviated to some $\hat{m} > m$, any plausible beliefs S_2 might have would lead her to conclude that he is still resolved. To see that, note first that if $\hat{m}$ happens to be the equilibrium mobilization of some other set of types, then it must be credible. To see that, suppose that another set of types (all of whom have valuations higher than the deviating type's) pool on $\hat{m}$ in equilibrium and some of them are unresolved at that level. This leads to a contradiction because we know that the deviating type must be resolved at m, which implies that any type with a higher valuation will also be resolved at m, which means that any such type will also be resolved at $\hat{m}$. Hence, $\hat{m}$ must be a credible mobilization level in equilibrium. It follows that $\hat{m}$ cannot be the equilibrium mobilization of a set that consists exclusively of bluffers. Therefore, it must be an out-of-equilibrium mobilization, in which case the plausibility requirement applies. The upshot is that S_2 will conclude that S_1's threat is credible and will resist with a strictly lower probability. To see the latter, recall that in the supposed equilibrium, the least-valuation type to resist is precisely indifferent between capitulation and resistance when she believes there is a positive chance that S_1 is bluffing. When S_2 revises her belief to conclude that resistance will certainly lead to war after $\hat{m}$, then some of the low-valuation types will no longer prefer resistance to capitulation. But the drop in S_2's probability of resistance, no matter how small, will increase the deviating type's payoff because the drop is discontinuous in $\hat{m}$; that is, even an allocation that is larger but arbitrarily close to m will produce it. But this means that this type would prefer to deviate, contradicting the equilibrium supposition. □

This is indeed a strong result that casts serious doubt on the what many analysts consider to be the crucial problem decision-makers face during a crisis—the search for credibility. What we have just found is that states can *always* find ways of making their threats credible. Military movies appear to be quite persuasive in that regard. In their analysis of seventy-seven crises involving military threats, Karsten et al. (1984, 54-7) found that these threats appear to be quite clear (98.9% of cases with respect to goals and 93.9% with respect to means). More importantly,

> The targets generally regarded the threats as credible—that is, in 69.5 percent of the cases, the targets appeared to believe that the threateners *did* intend to fight if their demands were not met. (In only 8 percent of the cases did the targets appear to believe that the threat lacked credibility. The remaining 22.5 percent of the cases did not belong to either of these clear-cut categories.)

The same study, however, also finds only weak correlation between credible communication and the success of threats (success being defined as the coercion of the target short of war). On one hand, it is not difficult to see how a threat that is perceived to lack credibility may not produce concessions. In the run-up to the First World War, the Germans heavily discounted the Russian threats. There were several reasons for this. Russia had made a similar threat during the Bosnian annexation crisis in 1908 but had backtracked when Austria-Hungary, propped by Germany, had stood firm. In that event, the Russians did not make any military moves and even pressured Serbia, which had mobilized, to back down too. The Germans had

calculated, correctly, that Russia was unprepared for war and had called its bluff.[36] Six years after this encounter, Russia was still recovering from the disastrous 1904 war with Japan, it was rebuilding its Baltic fleet and expanding its railways. The Germans were optimistic that Russia was bluffing yet again. They were wrong but the key question here is not why they did not believe the Russian threat but why they resolved to fight even *after* the Russian mobilization had made it plainly obvious that the threat was credible. In other words, once Russia moved militarily, its commitment to Serbia was revealed but as a result the Germans just mobilized for war themselves (Trachtenberg, 1991, 52–55).

Even a threat that is perceived as credible from the outset may not work. The military threat of the US-led coalition against Iraq over the invasion of Kuwait in 1990 was an instance of mobilizing overkill capability. Despite some mixed signals from the Bush administration (e.g., the ill-conceived 'last mile' diplomatic initiative), by late December Saddam Hussein was certain that the threat was credible. He may have hoped for some last-minute reprieve if he could just delay the outbreak of war long enough but even failing that he determined that the expected payoff of fighting to a military defeat would be better than withdrawal from Kuwait provided he survived the war. The problem was not that the threat was not credible—it was—but that Hussein's political calculations could not be influenced by military means whose objectives did not include the topping of his regime, something that was politically impossible for the coalition to aim at and maintain its cohesion (Freedman and Karsh, 1993, 275–78,434).

Why, then, do credible threats sometimes fail? Clearly, the problem is not that they are disbelieved, that the threatener has somehow failed to communicate the extent of his commitment. There must be something else. As we shall now see, this "something else" is the inability to undo the opponent's own commitment. Whereas the military threat can commit one to war and communicate this fact credibly, its costliness or the uncertainty about the extent of the opponent's commitment may prevent the threatener from mobilizing sufficient forces to induce that opponent to capitulate. Indeed, in 1914 the Germans took their estimate of Russia's lack of adequate preparation for war and translated it from thinking that the Russian threat was a bluff (because Russia would not fight a war unprepared) to thinking that Russia would certainly lose the war because it is fighting unprepared. The threat, while credible, was not capable enough to lower the Germans' expected war payoff to the point that would have induced them to negotiate. In 1990, Hussein had initially hoped that disagreements among coalition members and public opinion would constrain the Americans' willingness to use force. When this did not happen and the US proceeded with an enormous military build-up in the region, he convinced himself that these same factors would hamper the Coalition's ability to prosecute the coming war effectively: if he could only prolong it enough for casualties to accumulate, he would be able to obtain some concessions. The threat, while credible and capable did not reduce Hussein's expected payoff from war enough to induce him to capitulate.

Mobilizing for War, Coercion, and Compellence

Lemma 3.1 shows that in any equilibrium with plausible beliefs, any threat S_1 is willing to make must be credible when it runs a positive risk of war even if it involves pooling. That is

[36] Taylor (1971, 451-55), Lebow (1981, 122).

even if S_2 remains uncertain about his actual valuation, she will know that it is high enough to satisfy (CR_1). This now implies that any equilibrium mobilization m will lead to one of three possible behaviors by S_2, which we use to label the cases:

War Preparation: S_2 resists with certainty after m. Lemma 3.1 then implies that any such mobilization must lead to fighting. This is essentially the same as the complete information case: S_1's mobilization is sufficiently high to commit him to war but too low to undermine the resolve of the lowest-valuation type of S_2 he believes he faces. This causes S_2 to resist for sure even though she knows that the outcome will be war. As in the complete information scenario, when S_1 is certain that his mobilization will lead to war, his optimal preparation for it is unique for each valuation. There will be no pooling on a common threat here: S_1's war preparation will completely reveal his type. We shall use $w^*(v_1)$ to denote the optimal war preparation for S_1 with valuation v_1.

This is what Lebow (1981, 334) calls "justification of hostility crises" in which the decision for war precedes the apparently coercive military threats. Lai (2004, 216–18) also notes the difference between public coercive mobilizations and preparations for war (which are often done in secret): "Unlike public mobilization, private mobilization is designed to have little effect on the behavior of an opposing state in a crisis. It is not designed to coerce a state into backing down and agreeing to a settlement."

Coercion: S_2 resists with positive probability less than one after m. This is perhaps the most interesting case because it can only occur because of uncertainty about S_2's valuation. Even though her uncertainty about S_1's valuation will be resolved by his equilibrium behavior, *his* uncertainty about her valuation causes him to run a positive risk of war in return for a chance that she will capitulate. In principle, S_1 could always mobilize for assured compellence. However, doing so may be too costly because it requires that he mobilizes enough to undermine the commitment of the toughest opponent possible (highest-valuation type that he assigns positive probability to). Because there is a chance that S_2's valuation is lower, this mobilization may be too wasteful. Because it is also costly, S_1 will attempt to strike a balance: his mobilization will be smaller (than what it takes for assured compellence), which is cheaper but runs the risk that S_2 will resist and he will have to fight. Still, because he believes that the lower-valuation types of S_2 could be compelled profitably, he does not prepare for outright war: his mobilization will be larger (than his optimal war preparation that even the lowest-valuation type of S_2 would resist), which is more expensive but induces the possibility that S_2 will capitulate. In essence, S_1 faces a trade-off: spend more and obtain a higher probability that S_2 will give up or spend less and run a higher risk of war. The optimal mobilization balances these gains and losses.

This mobilization may involve pooling for reasons that are reminiscent of the "overkill" causes under complete information. If there are (low-valuation) types of S_1 that could profit from mobilizing at m (because S_2 capitulates with positive probability) despite the risk of having to capitulate in the endgame, then they would be tempted to mimic the behavior of the type that is resolved at m. But

Lemma 3.1 implies that such bluffing cannot happen in equilibrium, and indeed S_1 eliminates the possibility for bluffing by his optimal behavior: any type, the credibility of whose optimal coercive mobilization is threatened by potential bluffers, over-mobilizes to restore his commitment in a way that S_2 must believe. As we shall see, in equilibrium this means that all these resolved types pool on a common threat that is too costly for any potential bluffer to want to mimic. When potential lack of credibility is not an issue, the coercive mobilization level is unique for each valuation of S_1, and his behavior will completely reveal his type. We shall use $\widehat{m}(v_1)$ to denote the optimal coercive mobilization for S_1 with valuation v_1.

Assured Compellence: S_2 capitulates with certainty after m. As we shall see, this may involve bluffing. Intuitively, the logic of Lemma 3.1 does not apply because no type of S_1 will be willing to mobilize more than m to signal his credibility: there is no profit in doing so because S_2 is already certain to give up. From S_2's perspective, even though she might harbor doubts about S_1's commitment, the risk of war will be sufficiently high and her payoff there sufficiently bad to outweigh any gain she might have from resisting the threat. This is similar to the assured compellence equilibrium under complete information but is not the same because bluffing is possible only when S_2 is uncertain about S_1's type. As before, there can be at most one mobilization that ensures compellence in equilibrium because if that were not so, any type that chooses the higher compellent level can profitably deviate to the lower one. We shall use $\overline{m}$ to denote that unique compellent mobilization.

The rest of the analysis then reduces to the (rather tedious) exercise of determining which types of S_1 will choose a particular type of mobilization. In general, war preparation is cheapest, coercion is more expensive, and assured compellence is most expensive. This is easily demonstrated mathematically but the logic is transparent. By definition, any preparation for war happens when S_2 is certain to resist at that particular allocation even though she knows that S_1 is resolved. In particular, this means that the lowest-valuation type of S_2 would rather fight at any $m^*(\cdot)$ than capitulate. Certainly, she would do so at any mobilization $m < m^*(\cdot)$, therefore she could only be willing to give up after some $m > m^*(\cdot)$. Because the only way to obtain a positive probability that S_2 will capitulate is to induce some of the lower-valuation types to do so, it follows that any coercive mobilization $\widehat{m}(\cdot)$ must be greater than any war preparation.[37] Because S_2 resists $\widehat{m}(\cdot)$ with positive probability, it follows that her highest-valuation type is unwilling to capitulate at that allocation even though it is credible. The only way to induce her to do so is to undermine her commitment with an even larger mobilization. Therefore, $\overline{m}$ must be higher than $\widehat{m}(\cdot)$. Clearly, $\overline{m}$ is the largest mobilization that can be seen in equilibrium: because it ensures that S_2 will capitulate, there is no sense in spending more on unnecessary mobilization.

The equilibrium can involve mobilization of all three types. In general, the lowest-valuation types will find even the cheapest war preparation too onerous and will appease. Somewhat

[37] This result clearly depends on the assumption that the war payoff for one player is independent of the valuation of the opponent: if war is to occur, players do not care how much the enemy cares about the issue. An alternative assumption would be that higher-valuation types fight harder, so the payoff from war against them is lower. See Slantchev (2010) for a model along these lines.

higher-valuation types will be unwilling to spend on coercion but will prefer to avoid appeasement, and consequently mobilize for certain war. Even higher-valuation types can afford to spend more to get some chance of forcing S_2 to capitulate but find that ensuring that is prohibitively costly. They mobilize for coercion (which may involve some pooling). Finally, the highest-valuation types can afford the extravagant mobilization that causes S_2 to give up with certainty. These all pool on the unique assured compellence level, which makes it credible.

RESULT 3.6 *The costliness of the military instrument can sometimes make it unprofitable to employ it to minimize the risk of war. Its use must balance the costs of mobilizing the larger force necessary to achieve that against the gains from obtaining a higher probability of the opponent's capitulation.*

When Stalin challenged Finland in 1939 demanding the cession of strategic territory around Leningrad and the lease of the Hanko Peninsula for the creation of a base, the Finns could not have hoped to coerce the USSR by the threat of force. They had no way of making the threat capable enough to deter the Soviets from attempting to take by force what they could not gain through diplomacy. In a sense, a coercive threat was so prohibitively costly as to be entirely out of reach. Having judged their own interests in the issue to be vital, the Finns mobilized for war. However, despite their initial successes and the outpouring of diplomatic support for their cause, they could not prevail against a foe as tremendously more capable as the Russians (Jakobson, 1961; Trotter, 1991).

The October War of 1973 is another case in which the opponent was seen as being too strong and having too high a valuation to be coerced at an acceptable cost. In this instance, Egyptian military preparations that would underscore the threat, and make it credible and capable would be too extensive and too risky should the threat fail. It was more efficient to attack Israel despite the low probability of overall victory because even partial success in the war would shaken its image of invincibility and perhaps induce it to offer terms that it could not be coerced into offering with threats alone.[38]

When Xerxes invaded Greece in 480 B.C. and challenged the city-states to surrender, his army was too large for the disunited Greeks to be able to muster the military strength to deter him. With their pessimistic assessment of the likely outcome with such a severe asymmetry of power, most capitulated. Only Athens and Sparta decided to make a stand, and of these only the Athenians fully prepared for war. In this case, the mobilization had no intent to coerce the Persians—that would have been unthinkable—it was meant to fight a last-ditch attempt to prevent the subjugation of the peninsula. The Athenians did not have much confidence in their chances: watching their city burn from the temporary safety of Salamis, they had already made plans to evacuate to Italy should the Spartans fail to support the planned battle in the straits (Green, 1996; Strauss, 2004).

[38] It is also interesting to note that the costliness of mobilization may have prevented Israel from establishing a deterrent posture. Egyptian exercises in the Canal Zone were a constant source of anxiety for the Israelis. The IDF had been partially mobilized on several occasions because of them and after the last war scare in June ended without hostilities, there was criticism of the government for the economic costs of mobilization. Both Defense Minister Dayan and Chief of Staff Elazar were reluctant to order another mobilization in late September (Dupuy, 1978, 406–8). When Elazar became sufficiently worried about the somewhat vague intelligence signals and asked the government to authorize mobilization, it was for a preemptive strike or counter-attack, not deterrence (The Insight Team of the London *Sunday Times*, 1974, 114–23).

In both illustrative cases, the problem was in the threatener's perception of the opponent's valuation (very high) and military strength (very capable). This combination makes coercion very expensive and (although that factor is not present in the model) risky if it fails: an overt military preparation may make one a convenient target should the opponent refuse to capitulate. This suggests that militarily stronger states with clearly defined interests in the dispute may resort to coercion more frequently than militarily weaker states. This in itself is perhaps not surprising for it echoes common wisdom. However, what is less obvious is that the logic also suggests that even though strong states will appear more bellicose because they resort to military coercion more often, they will actually have to fight less often. That is, if a weaker state finds itself in a dispute with a strong opponent, it is likely to find it optimal to fight its adversary rather than attempt to coerce it. In asymmetric disputes, *the weaker side will be more likely to initiate hostilities.*[39]

The flip side of this argument is even more explosive: a strong state that does not engage in aggressive coercion essentially reveals lack of interest in the issue and invites attack from a weaker high-valuation opponent. This suggests that the stringent and unyielding demands the US made of Japan in 1941, seeing as they were not accompanied by military preparations to make American resolve credible, were practically bound to provoke what became the attack on Pearl Harbor. One is hard-pressed to see how it could have been otherwise, which is why it is not surprising that many have concluded that the US policy was deliberately designed to provoke such an attack: an overt display of force could have persuaded the Japanese to abandon their claims.[40]

To see more precisely under what conditions a threatener would prefer to attempt coercion instead of outright war, we must return to the model. When mobilizations of these different types occur in equilibrium, they always partition the valuations of S_1 in order of increasing valuation, as indicated above. It is possible that no type mobilizes for war, or that no type mobilizes for coercion, or that no type mobilizes for compellence, or any combination of these. For instance, it could be the case that in equilibrium some types appease, others prepare for war, and others mobilize for assured compellence (i.e., no type mobilizes for coercion). We know that the types who appease have the lowest valuations, the types who fight have moderate valuations, and the types who compel have the highest valuations. When neither of these mobilizations occurs in equilibrium, we have an instance of **assured appeasement**: S_1 appeases S_2 regardless of his valuation. Appendix B shows the conditions that determine which particular configuration of mobilization levels will exist in equilibrium.

Figure 3.3 illustrates graphically the equilibrium for a case in which S_1 appeases if his valuation is low, prepares for war if his valuation is moderate, mobilizes for coercion if

[39] Paul (1994) reaches similar conclusions but for slightly different, though not incompatible, reasons that emphasize the advantage of a surprise attack.

[40] Sagan (1994, 61-63) argues that the stations of the US Fleet at Pearl Harbor was too vague of a threat, especially in context of Roosevelt's domestic promises. The American ambassador to Japan reported that Japan will be deterred only insofar as it believes that the US is serious. Ironically, Roosevelt seems to have thought that an explicit verbal threat would be more provocative than a clear military move (80). As a result, the US opted for a slow buildup in the Pacific that only succeeded in putting severe pressure on Japan to jump the gun instead of surrendering in the future. The model here implies that had the US mobilized openly to persuade Japan that it would be embarking on a protracted war, Japan would have capitulated. The problem was that Roosevelt could not have ordered such a mobilization for domestic reasons. He also seems to have underestimated the probability that Japan would risk an attack.

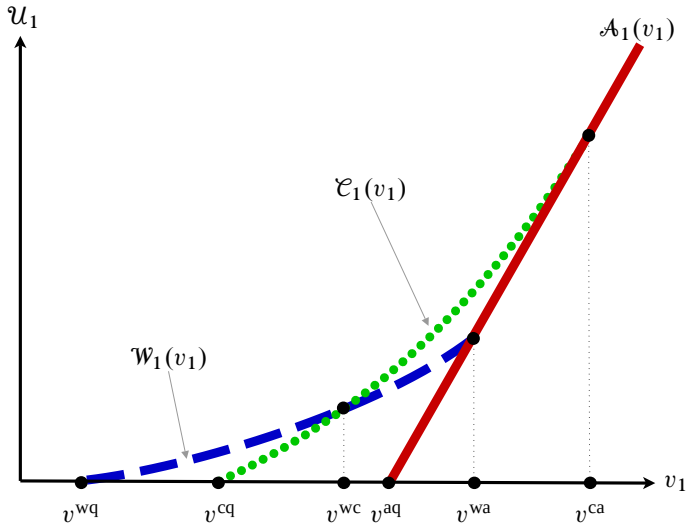


Figure 3.3 War Preparation, Coercion, and Assured Compellence.

his valuation is somewhat high, and mobilizes for assured compellence if his valuation is really high. The figure plots S_1's expected payoff from mobilizing optimally given what S_2 is expected to do after such a mobilization. Because S_1's behavior depends on his (privately known) valuation, the payoffs differ in v_1. For example, $\mathcal{W}_1(v_1)$ is S_1's payoff from certain war for which he has mobilized at the unique for his valuation best preparatory level $m^*(v_1)$; $\mathcal{C}_1(v_1)$ is his payoff from coercion with the unique $\widehat{m}(v_1)$ for his valuation; and $\mathcal{A}_1(v_1)$ is his payoff from assured compellence with $\overline{m}$, which of course is type-independent.

The shapes of these payoff functions are derived in Appendix B, which also establishes that as S_1's valuation increases, his payoff from assured compellence increases by more than his payoff from coercion, which in turn increases by more than the payoff from war. Intuitively, this follows from the fact we established earlier: the war preparation level is smaller than the coercive one, which is still smaller than the assured compellence level. Observe now that as v_1 increases so do $m^*(v_1)$ and $\widehat{m}(v_1)$. However, $\mathcal{C}_1(v_1)$ will increase by a larger amount than $\mathcal{W}_1(v_1)$ because while larger mobilization improves the payoff from war in both cases, in the coercive scenario it will also cause S_2 to capitulate with a higher probability. As a result, the overall improvement in S_1's payoff will be larger when he mobilizes for coercion than when he prepares for war. To see that $\mathcal{A}_1(v_1)$ increases by even more, it suffices to note that because $\overline{m}$ does not depend on S_1's type and there is no risk of war, the increase in S_1's valuation translates directly into a corresponding increase in his payoff from assured compellence. This is by far better than in the coercive case where the costs of the higher mobilization and the positive risk of war temper the improvement.

Appendix B also shows the derivation of the various special types of S_1 indicated on the horizontal axis. For our purposes here, it will be sufficient to note that the superscripts on these types consist of two-letter mnemonics designed to indicate which two actions the type is indifferent between. The codes are as follows: 'q' (quit for appeasement), 'w' (certain war), 'c' (coercion), and 'a' (assured compellence). For example, v^{wq} denotes the type that

is indifferent between certain war and appeasement, and the type v^{wc} is indifferent between war and coercion. Equilibrium behavior by S_1 takes the following form:

- if $v_1 \leq v^{wq}$, appease S_2 immediately;
- if $v_1 \in (v^{wq}, v^{wc}]$, prepare optimally for certain war by mobilizing $m^*(v_1)$; in that case S_2 resists for sure, and S_1 attacks;
- if $v_1 \in (v^{wc}, v^{ca})$, mobilize $\widehat{m}(v_1)$ as coercion; in that case S_2 capitulates with positive probability, and if she does resist, S_1 attacks;
- if $v_1 \geq v^{ca}$, mobilize $\overline{m}$ for assured compellence; in that case S_2 capitulates with certainty (and if she were to resist, S_1 would attack for sure).

In this scenario, when S_1's valuation is $v_1 \in (v^{cq}, v^{wc})$, he can actually profit from coercion relative to appeasement as well: $\mathcal{C}_1(v_1) > 0$ for them. However, because the coercing S_2 requires a relatively large mobilization, this option is unattractive. Consequently, these types mobilize $m^*(v_1)$, which is smaller, still credible, but insufficient to get S_2 to capitulate with positive probability, so certain war is the outcome. When S_1's valuation is moderate, he will be unwilling to pay to coerce his opponent, but he will be willing to prepare and fight a war instead. This behavior is roughly analogous to what we have already seen in the complete information scenario.

If S_1's valuation is moderately high, coercion becomes attractive relatively to certain war, so S_1 switches to a strategy in which he mobilizes a larger force but in turn reduces the risk of war and increases the likelihood of S_2's capitulation. Observe now that if $v_1 > v^{aq}$, then assured compellence is better than appeasement, if $v_1 > v^{wa}$, it is also better than certain war, and if $v_1 > v^{ca}$, it is also better than coercion. When S_1's valuation is $v_1 \in (v^{wa}, v^{ca})$, then mobilizing $\overline{m}$ for assured compellence is more attractive than either war or appeasement. However, he will still be unwilling to pay as much as necessary to ensure that S_2 capitulates. Instead, he economizes on the costs of mobilization and accepts a positive risk of war.

This analysis agrees with Kagan's (2003, 54) contention that Pericles' defensive strategy for Athens in the Peloponnesian War was seriously, if not fatally, flawed for the purposes of deterrence. Once the Spartans had made their final non-negotiable demand, Athens could not have persuaded them to give up short of war without a demonstration of capability that would give the conservative faction in Sparta enough strength to win the domestic argument against war with Athens. Instead, Athens essentially threatened to fight a war that posed no risk to the Spartans and imposed no significant costs on them. Inevitably, then, "without an obvious, credible, frightening offensive threat [Pericles'] diplomatic strategy of deterrence was crippled and doomed to failure."

It is not difficult to find examples of successful military threats. Take, for instance, the famous Fashoda Incident of 1898. When the French expedition under Marchand attempted to secure the area around Fashoda, it met the recently victorious Kitchener who was in the process of reconquering the Sudan for the British Empire's Egyptian client. Although their meeting was polite, it sparked a crisis in Europe. For two months, recriminations flew across the Channel, with war fever running high. Eventually, the British began earnest public preparations for war by mobilizing their fleet (Langer, 1935, 537-80). This forced the French to capitulate: despite their large army they were in no position to engage the British without naval superiority that they could not hope to achieve (Bates, 1984). Schultz (2001) argues

that British political unity was crucial in rendering the threat to use force credible, and it may have been so. In support, he cites Joseph Chamberline's claim that British victory was due "as much to the spectacle of a united nation... as it was to those military and naval armaments *about which the foreign press talks so much* and knows so little" (emphasis added). What this quote reveals, however, is the preoccupation of the opponent with the actual military preparations undertaken by the British. Lebow (1981, 326) is right to conclude that Britain's success came from "her greater willingness to use force and Salisbury's ability to impress this fact upon France."

It is important to emphasize that military threats can fail despite being credible. Many studies have unduly privileged credibility in crisis interactions, which has attracted serious criticism that usually aims at the wrong target. For instance, Lebow (1981, 274) attacks the focus on credibility by observing that

> efforts to impart credibility to commitments may have only a marginal impact on an adversary's behavior. Even the most elaborate efforts in this regard may prove insufficient to discourage a challenge when policy-makers are attracted to a policy of brinkmanship as a means of preserving vital strategic and domestic interests. [... In many cases,] the defending state not only did its best to buttress the credibility of its commitment, but the commitments in question represented interests of sufficient political or strategic magnitude to have given pause to any kind of rational adversary.

From this, he (and others) have concluded that credibility is not the biggest problem faced by decision-makers who are trying to make their threats stick. Although this finding somewhat disingenuously ignores the fact that deterrence theorists were concerned with the credibility of threats in the shadow of nuclear weapons, it is important to recognize the element of truth in it, and one need not jump to irrationality to do so. The problem is not that credibility is unimportant—it is, and as we shall see shortly one may have to pay dearly to maintain it—but that it is only one ingredient in the effectiveness of a military threat. As Figure 3.3 shows, the threat need not maximize the probability of capitulation: it must balance this against the costs of ensuring it. Fully effective threats that achieve assured compellence will be relatively rare precisely because they require the threatener to be deeply invested in the issue and able to demonstrate sufficient military capability. More often, military threats will be coercive in the sense that the threatener will have to accept some risk that they will fail despite their credibility (meaning a risk that he will have to go to war). He must balance this risk against paying so much for mobilization that even its success in ensuring peace would not offset its enormous expense. Peace through strength may be worse than life in a more dangerous world.

Mobilizing Overkill Capability

The costliness of the military instrument can be a real problem for its use because S_1 may not be able to afford the large mobilizations that would be necessary to reduce the probability of war. One particularly intriguing scenario can occur when S_1's valuation is moderately low and his opponent is fairly easy to coerce provided he believes S_1's threat. We have already seen, in Example 3.4, that under complete information this sort of situation will lead to mobilizing "overkill" capability. Over-mobilization can also happen here for reasons that are roughly analogous.

Figure 3.4 illustrates this case. Here, the type that is actually resolved at his optimal

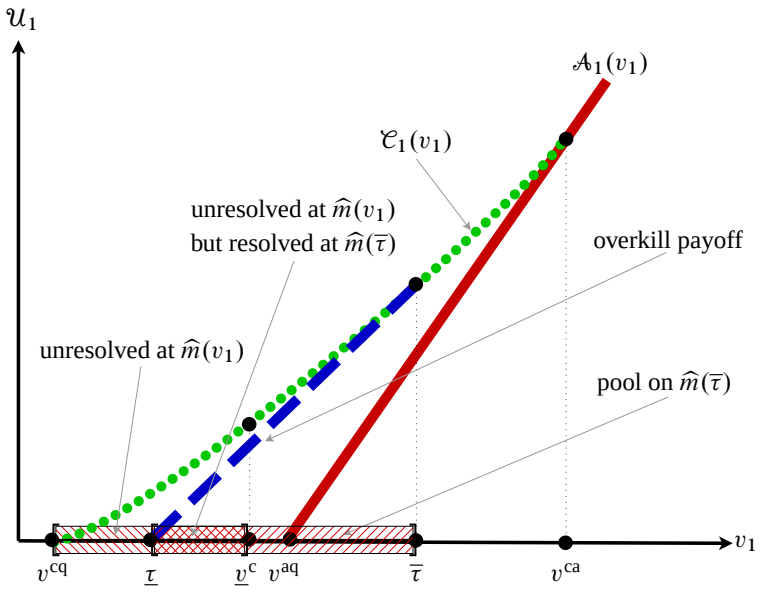


Figure 3.4 Over-mobilization to Establish Credibility.

coercive mobilization $\widehat{m}(v_1)$ is denoted by $\underline{v}^c$. If S_1's valuation is such that $v_1 \in [v^{cq}, \underline{v}^c)$, then he would strictly prefer to mobilize for coercion at $\widehat{m}(v_1)$ than to appease S_2 (assured compellence is out of reach). The problem is that none of these types can credibly threaten to fight when resisted. Because the optimal coercive level is unique for each valuation, if S_2 were to observe it, she will infer S_1's lack of commitment and will resist with certainty as she will know that there is no positive risk of war. Hence, none of these types will want to mobilize at their coercive levels.

Unfortunately, that does not mean that they will be content with appeasement. Instead, they may attempt to bluff. Consider, for instance, what happens if S_1's valuation were smaller than $\underline{v}^c$ but still among the ones that could profit from coercion. As we have just seen, he would not want to mobilize for coercion at his own uniquely optimal level. But what if he mimicked $\underline{v}^c$'s behavior and allocated $\widehat{m}(\underline{v}^c)$? Since this mobilization is credible and profitable for $\underline{v}^c$, $\underline{v}^c$ himself would use it in equilibrium. This means that when S_2 sees it, she would infer that her opponent is surely committed. Suppose the bluffer deviated to that (costlier) mobilization and still remained unresolved at it. He would obtain a strictly positive probability of capitulation by S_2 who erroneously believes the threat to be credible. Often this would be enough to give him a payoff that is strictly better than appeasement. If this happens, than such a type would clearly prefer to pretend he is $\underline{v}^c$. But Lemma 3.1 tells us that there can be no equilibrium in which bluffing occurs when there is a risk of war after the threat, as there would be here because $\widehat{m}(\underline{v}^c)$ is merely coercive and not compellent.

What, then, prevents this bluffer from undermining the credibility of $\underline{v}^c$'s threat? It is $\underline{v}^c$ himself. He does this by engaging in behavior that is sufficiently unattractive for the bluffer to mimic. Note first that $\underline{v}^c$ will be hurt by the potential presence of a bluffer who might imitate his mobilization level: because S_2 is well aware of the incentives such a bluffer

might have, she cannot possibly believe that S_1 is truly committed if he mobilizes $\widehat{m}(\underline{v}^c)$. This ruins the coercive strategy of $\underline{v}^c$ because it increases the probability that S_2 will resist. Suppose now that $\underline{v}^c$ instead mobilized $\widehat{m}' > \widehat{m}(\underline{v}^c)$. Clearly, he will still be committed at this larger allocation. Furthermore, if this over-mobilization is sufficiently costly, the bluffer will no longer find it profitable to mimic (which he would have to do because now $\underline{v}^c$ is not expected to use $\widehat{m}(\underline{v}^c)$, so staying with this mobilization would reveal to S_2 that it is actually the bluffer using it). In any equilibrium with plausible beliefs, when S_2 sees $\widehat{m}'$, she will infer that S_1 is committed for sure, so this over-mobilization will restore the credibility of $\underline{v}^c$'s threat. Although it will reduce his payoff somewhat (because of the higher costs he has to pay), it will still be better than appeasement and better than using his own coercive allocation that S_2 does not believe credible. Consequently, mobilizing "overkill" capability will be preferable for this type than permitting bluffing to diminish his chances for coercing S_2 successfully.

As Figure 3.4 shows, $\underline{v}^c$ would have to go as high as $\overline{\tau}$'s mobilization before all potential bluffers are eliminated. Because larger mobilization are more committing, using $\widehat{m}(\overline{\tau})$ will actually credibly commit types with valuations smaller than $\underline{v}^c$. These cannot use their own coercive allocations, which are smaller than $\underline{v}^c$'s and are certain to be mimicked by bluffers if S_2 were to believe them. If these types want to ensure that S_2 believes their commitment, they must also over-mobilize at $\widehat{m}(\overline{\tau})$. There is no reason to go higher because doing so will not improve credibility and is quite costly. To see which types are willing to over-mobilize, let $\underline{\tau} < \underline{v}^c$ be the type that is just indifferent between over-mobilization with $\widehat{m}(\overline{\tau})$ and appeasement. If S_1's valuation is $v_1 \in [\underline{\tau}, \overline{\tau}]$, he will mobilize "overkill" capability $\widehat{m}(\overline{\tau})$, which signals credibly his commitment and achieves coercion. Clearly, no $v_1 < \underline{\tau}$ is willing to imitate such a high allocation because doing so would give him a payoff that is worse than appeasement. All over-mobilizing types get payoffs that are strictly worse than what they would have obtained had bluffers not ruined the credibility of their optimal coercive mobilizations.

RESULT 3.7 *Uncertainty about an actor may sometimes undermine the credibility of his commitment because his opponent believes that his coercive mobilization is small enough to be profitably imitated even when it would leave him unresolved. To re-establish the credibility of his commitment, this actor will mobilize overkill capability when the resulting credibility gains outweigh the signaling costs.*

By the end of 1990, the American military buildup in the Persian Gulf—especially the controversial doubling of troops in Saudi Arabia—persuaded Saddam Hussein that the US threat to force him out of Kuwait was credible. The US could not have attained credibility if it had relied on the deterrent mobilization only—the forces necessary for offensive operations were about double the size required for defense of Saudi Arabia (Brune, 1993, 60-61). In the event, credibility proved insufficient because Bush expanded US demands (from withdrawal from Kuwait to dismantling of Iraq's military machine) and, more importantly, because Hussein overestimated his expected payoff from war. He clearly hoped for a protracted war that would force the Americans to offer some concessions or, failing that, to extract political dividends even in defeat by resisting a much more powerful Western invader, much like Nasser had done in 1956 when he lost the military confrontation with the British and the French but won the political battle (Freedman and Karsh, 1993, 275-78).

As in the complete information case, lack of credibility at a mobilization too small to

commit one to war can be a serious problem even when it would have been quite sufficient to coerce the opponent if she would only believe the threat it represents. Under uncertainty, the problem is even worse because when S_2 is unsure of S_1's commitment she will attempt to infer it from his behavior. When she knows S_1's coercive threat is not credible for some middling valuations of his, she can prevent him from using it by resisting any such mobilization. Unfortunately, this creates incentives for S_1 to bluff when he has such a valuation. He may exploit the fact that S_2 can only observe his behavior and may attempt to bluff her into believing that he is truly committed by mimicking the allocation of some resolved type. Of course, S_2 is quite aware that S_1 may be tempted to pretend that he is tougher than he actually is, and will not believe that larger threat either.

Hence, asymmetric information about S_1's valuation has created a problem for S_1 himself if his valuation is among the ones that could be mimicked profitably by bluffers. In equilibrium, any type affected by this problem will resolve it by mobilizing "overkill" capability—a level that is too high for lower-valuation types to imitate profitably. The sole purpose of such excessive allocation is to restore the credibility of S_1's commitment. Even though over-mobilization does undermine S_2's resolve more than the optimal coercive mobilization does (and so reduces the probability that she resists), this type would not have chosen it if he were not forced to deal with bluffers—it is too costly to be worth the benefit. It is when credibility is at stake that over-mobilization becomes worthwhile.

As any Great Power, the US frequently faces a problem when it comes to convincing opponents of the seriousness of its intentions. The problem is that while such a state has far-flung interests, its power makes military threats relatively inexpensive. As a result, and perhaps somewhat contrary to one's intuition, these threats become less effective for there is nothing to stop the Great Power from attempting to bluff its opponent into submission—certainly not the (relative lack of) expense associated with gunboat diplomacy. This is why Great Powers may often have to resort to overkill preparations when they are serious in their intent.

This logic was at work in 1994 when the US confronted the military leaders of Haiti. Although both Congress and public opinion in the US appeared hostile to an armed intervention to oust the junta, on September 15 President Clinton publicly threatened to use military force to do just that. The generals led by Raoul Cedras only agreed to step aside when they were provided with evidence that paratroopers were boarding planes in North Carolina for the invasion Clinton had ordered to begin that evening (on the 16th, CINC declared the junta hostile, which altered the Rules of Engagement for the US forces and cleared them for lethal action against Haitian security forces). As the President noted, "This agreement only came because of the credible and imminent threat of the multinational force."[41] The threat was indeed massive: 18 warships, including two aircraft carries, along with an invasion force of 20,000—the same number that landed in Haiti to enforce the Carter-brokered agreement a week later. The planned airborne assault was to be the largest in history since D-Day in 1944 (Ballard, 1998, 85-103). Even then, sixty-one planes were already airborne on the 18th before the junta leadership blinked.[42]

[41] *St Petersburg Times*, Florida, September 18, 1994.
[42] *United Press International*, September 18, 1994.

When Commitments Need Not Be Credible

Bluffing need not always be a problem for S_1, so he may not need to over-mobilize to re-establish the credibility of his commitment. Whereas Lemma 3.1 shows that threats must be credible if they induce a positive risk of war, this need not be the case if they do not. As I have argued before, when a threat carries no risk of war, it must be because it causes S_2 to capitulate for sure or else S_1 would not be make it. But if his threat leads to assured compellence, then S_1 need not worry about his potential lack of credibility. To put it another way, if $\overline{m}$ causes S_2 to give up with certainty even though she suspects S_1 might be bluffing, then there is absolutely no need for him to mobilize more to establish his credibility beyond doubt. Doing so would serve no useful purpose: S_2 can do no more than capitulate, which she already is doing, and the larger mobilization is more expensive.

It is not difficult to see why S_2 would sometimes capitulate for sure even though she knows that S_1 might be bluffing. As (CR_2') makes clear, if $v_2^*(\overline{m})$ exceeds her highest valuation, then she cannot credibly threaten to resist even if $G_1(v_1^*(\overline{m})) > 0$; that is, even if there is a positive probability that S_1 will capitulate if she resists. The risk of resistance is just too great, not only because there is a serious chance of war but also because the payoff from war is quite unattractive given S_1's mobilization.

Bluffing with military threats drastically differs from bluffing in the simple crisis game from Figure 2.1. Proposition A.3 shows that when S_2 remains unsure whether S_1 is truly committed despite his threat, she might resist even if she is unresolved. Because she only resists genuine threats when she is resolved, this implies that the probability that S_2 resists when S_1 might be bluffing is higher than the corresponding probability when his threat is credible. If S_1 happens to be bluffing, this means that he must face a higher probability of capitulation. It is precisely this risk that deters excessive threats by an unresolved S_1: if his valuation is not sufficiently high, then the increased risk of capitulation keeps him from escalating. In contrast, bluffing is without any risk in the military threat game: what keeps low-valuation unresolved types from threatening is the cost of the mobilization itself.

More importantly, however, in the simple crisis game bluffing increases the risk of war for S_1 when he is resolved. This happens because all such types fight when resisted and, as we have just seen, S_2 is more likely to resist. Hence, a resolved S_1 faces a problem analogous to the one faced by $\underline{v}^{\text{c}}$ in the military threat game (and shown in the previous section): the credibility of his threat is undermined by the bluffers and he would like to find a way to re-establish it by engaging in behavior that bluffers are unwilling to imitate. Unfortunately, in the simple escalation game, S_1 is limited in his actions and there is no way for resolved types to separate themselves from the bluffers. A resolved S_1 cannot escalate "more" than a bluffer to signal his commitment credibly to S_2. It is precisely this type of problem that sunk-cost (Corollary A.1) or tying-hands (Proposition 2.5) escalation manage to solve. Military threats can also solve this problem through overkill mobilization. In all these scenarios, credibility is restored by making the threat unattractive to the bluffer: it is either too costly, too risky, or some combination of those, for him to attempt.

In contrast to the simple crisis game, bluffing with a military threat carries no risk of war for S_1 when he is resolved. Consequently, the only types that would normally have an incentive to engage in separating behavior because the credibility of their threat is undermined, have no reason to signal their commitment. Bluffing is possible because it is permitted by

the resolved types who could reveal their commitment but choose not to. Nothing like this can happen with non-military threats because these do not affect S_2's payoffs and cannot undermine her resolve. With the military instrument, however, the situation is very different. A sufficiently large mobilization by S_1 can dissuade even the highest-valuation opponent from resisting despite the chance that such mobilization might be a bluff. Once S_2's commitment is fully undone, there is no need for a resolved type to relieve S_2 of any lingering doubt she might be harboring about his commitment. When the conditions are right, bluffing can occur in the military threat model, and it cannot be eliminated because the only types with a conceivable reason to do so cannot benefit from doing it.

3.5 Conclusion with a Side on Reputation

In previous work I examined military threats in contexts where S_1's valuation was common knowledge and there was uncertainty only about S_2's valuation. I found that bluffing only happens in an assured compellence equilibrium, and I conjectured that "the result of bluffs never being called in equilibrium probably arises from the one-sided incomplete information in the model" (Slantchev, 2005, fn. 13, 541). I then further speculated that "If there were uncertainty about S_1's valuation as well, S_2 could bluff hoping that S_1 will quit." So it came as a surprise to me that bluffs never being called is not an artifact of the informational environment—the same result persists even when there is uncertainty about S_1 valuation.

More importantly, bluffing is not simply deterred by S_2's knowledge of S_1's valuation but by the optimal behavior of resolved types of S_1 whose normal coercive tactic would be most vulnerable to credibility problems—they over-mobilize to re-establish credibility. This militates against the widespread notion that the weak could benefit from having a reputation for strength. This notion is at the heart of the traditional formal approach to reputation based on the seminal work by Milgrom and Roberts (1982), and Kreps and Wilson (1982). The idea is that when costly confrontation is always suboptimal in a one-shot encounter, it may be worth engaging in it if the relationship is long-term. The problem, as Selten (1978) observed a long time ago, is that with finite relationships reputation cannot be built if both players are rational and we assume that conflict in one-shot encounters is suboptimal for the one trying to establish a reputation. Reputation for strength can be possible if there is a positive probability that one may be a type that will fight even in the one-shot setting. If that is the case, then the weak can mimic that type's behavior in a few encounters, which in turn would increase the opponent's belief that he might be facing the strong opponent, which would in turn moderate the opponent's behavior, yielding the benefit from the costly building of reputation by the weak.

One need not model repeated encounters to see the logic. In our setting, mobilization can yield large benefits if it is believed by the opponent. The defender can establish a reputation (credible commitment) by engaging in costly behavior. If an unresolved type somehow managed to persuade the opponent that he is resolved, then this reputation would serve him well in a dispute with that opponent. That is, reputation is your opponent's belief about what you would like to be, in this case a defender with a credible commitment to fight when resisted. The military threat model suggests that one cannot establish such a reputation, and the reason is that the strong type—which the weak will have to emulate—will take into account this possible mimicry and will opt for behavior that the weak cannot possibly profit

from emulating even if doing so would establish the desired reputation. In other words, the costs of building a reputation of strength will outweigh the benefits. The traditional models circumvent this by assuming that the strong has no further action to take to distinguish himself from the weak: it is an escalate or do not escalate decision, which, as we have seen in Chapter 2, can be mimicked with (relative) impunity. But if the strong can engage in more complex behavior—by choosing the intensity of effort in fighting, for instance—then the possibility for mimicry vanishes very fast.

These results support Mailath and Samuelson's (1998) point that reputation is not who you would like to be (a resolved type when you are not resolved) but who you are not (an unresolved type when you are resolved).[43] For reputation to work, it is not enough for the weak type to pretend to be strong by taking actions a strong one would have taken in his place. It must also be the case that the strong type cannot alter his behavior to restore the separation. The strong must have no incentive to maintain his reputation that is being threatened by the potential for mimicry by the weak. If the weak could mimic the strong by taking some action, then this action cannot persuade the opponent that one is strong (after all, there is a positive probability that it is being taken by the weak). But this implies that this action would hurt the reputation of the strong because it will not be persuasive enough, which in turn implies that the strong will have incentives to find another action that will enable him to restore his own reputation. This is exactly what happens in the military threat model: high-valuation types, whose credibility would be undermined by the lower-valuation types bluffing at their optimal coercive mobilization levels, opt for "unnecessarily" intense mobilizations to discourage that behavior.

This suggests that *reputation is a curse of the strong*, who are forced to maintain it, often at great cost, just to differentiate themselves from the weak. While actively doing something may establish one's reputation in the sense that the opponent may believe that one is resolved, reputation does not automatically mean pace—war is still a possibility. But failure to take that action can then be a full-blown signal that one is not resolved. Being proactive may build a reputation while not necessarily leading to peace. Being passive, on the other hand, would either lead to capitulation or to higher risks of war because now a challenge is more likely.[44]

This is an unhappy result: to avoid the severe losses arising from inaction, a strong type must embark on a costly and risky course of action. The strong must always be prepared to prove their strength. This explains Taylor's (1961, xviii) paradox:

> Though the object of being a Great Power is to be able to fight a great war, the only way of remaining a Great Power is not to fight one, or to fight it on a limited scale.

Now, the "object of being a Great Power" is not to fight a great war, it is the ability to fight such a war that makes a state a great power. However, it seems true that many great powers that have engaged in all-out contests with others like them have not done very well. Even allowing for Taylor's typical hyperbole, there is a kernel of truth here but the paradox is only apparent: one must not underestimate the benefits from having the ability to fight a great war.

[43] Mailath and Samuelson (2006) offer a full treatment of reputation in long-term relationships.

[44] This suggests, however, that the strong may pretend to be weak if they estimate there is a high risk of war anyway and they may derive advantages from misleading the opponent into a false sense of security (Slantchev, 2010).

It is this ability that enables one to threaten to use force with some credibility, which is what in turn can deter challenges from other great powers, not to mention minor powers which one can even fight with near-certainty of success. To create the paradox, Taylor ignores the counterfactual: without such an ability, a state would be perceived as weak and would not be able to enjoy the fruits its status as a great power brings. Having to pay for one's reputation may be a curse of the strong, but they are the ones that reap the benefits as well.

4

Comparing the Instruments of Coercion

Those who know when to fight and when not to fight are victorious. Those who discern when to use many or few troops are victorious. Those who face the unprepared with preparation are victorious.

Sun Tzu

We are now in a position to compare and contrast optimal crisis behavior using the military instrument to other escalatory moves, such as sinking costs, running risks, or tying hands. Crisis behavior almost always involves more than one tactic: from diplomatic maneuvering to military threats to small-scale fighting. As discussed in Chapter 2, the instruments represent ideal types that only roughly approximate actual behavior. It is nevertheless useful to establish some benchmark comparisons that will facilitate the exposition of the advantages and disadvantages of the various instruments of coercion. Before this comparative exercise can commence, however, we need two preliminaries. First, we need to be precise about what effects of the various instruments we are interested in. At the most basic level, we would like to know how they affect the probabilities of war, of escalation, of preserving the status quo, and so on. To this end, I will define these quantities of interest more precisely. Second, I introduce the basic setup for the simulations that I will use to explore the behavioral dynamics of the various models.

4.1 Stability and Expected Mobilization

The military threat model (MTM) developed in Chapter 3 assumes that S_2 has made a demand, and so the following discussion is predicated on the existence of a crisis. There are two distinct crisis phases that we might be interested in: one is prior to S_1's escalatory decision—the *crisis* phase—and the other follows it—the *militarized (or acute) crisis* phase. This now allows us to distinguish between several possible concepts. Appendix C provides formal derivations of these quantities.

Risk of War: Crisis and Escalation Stability

Crisis stability refers to the probability that the crisis will end in war. This probability is evaluated prior to S_1's initial move and takes into account the likelihood that he will attempt to appease his opponent. This is an *interim* calculation in the sense that it presupposes the

crisis. As we shall see in Chapter 5, the concept of *ex ante* stability must take into account that S_2 challenges the status quo in the first place.[1]

DEFINITION 4.1 (Crisis Stability). The unconditional probability that a crisis escalates into a war. The higher the probability, the less stable the crisis.

This notion of stability differs from the *ex post* evaluation of the probability of war that occurs right after S_1 decides to militarize the crisis. This likelihood is now conditional on S_1 abandoning appeasement altogether. *Escalation stability* is essentially the probability of war before S_2's last chance to avoid it.

DEFINITION 4.2 (Escalation Stability). The *ex post* probability that a crisis escalates into a war conditional on S_1 choosing to militarize it. The higher the probability, the less stable the militarized crisis.

To illustrate the differences between the two concepts, suppose that $[t, u]$ and m are such that the war preparation equilibrium from Proposition B.1 obtains. For the sake of simplicity, suppose $\overline{v}_1 = 1$, $v^{\text{wq}} = 1/4$, and $v^{\text{wa}} = 3/4$. The probability that S_1 appeases is $\Pr[v_1 \leq v^{\text{wq}}] = 1/4$, and the probability that he escalates is $\Pr[v_1 > v^{\text{wq}}] = 3/4$. The unconditional probability that the crisis would escalate to war equals $\Pr[v^{\text{wq}} < v_1 < v^{\text{wa}}] = 1/2$, and so crisis stability is 50%. Finally, given that S_1 has escalated, the probability that the crisis will end in war is $\Pr[v^{\text{wq}} < v_1 < v^{\text{aq}} | v_1 > v^{\text{wq}}] = 2/3$ by Bayes rule, and so escalation stability is approximately 66%.

To summarize, crisis stability answers the question, "How likely is some crisis to escalate into war?" One can think of this as measuring the danger of war after the status quo is challenged. Escalation stability answers the question, "How likely is a militarized crisis to escalate into war?" One can think of this as measuring the additional danger that failure to appease introduces into the crisis, turning it into an acute confrontation.

Peaceful Resolution: Appeasement and Capitulation

The other quantity of interest is the likelihood that the crisis is peacefully resolved in one way or another. If S_1 opts for appeasement, the crisis is defuzed and the status quo is peacefully revised in S_2's favor. If S_1 escalates, the *ex post* probability of a peaceful resolution reduces to evaluating the likelihood that S_2 will capitulate. (This is because S_1 never capitulates in equilibrium when S_2 stands firm.) In this case, the crisis is defuzed with the maintenance of the status quo in S_1's favor.

Mobilization Levels

By analogy with crisis and escalation stability, there are two points at which S_1's mobilization can be estimated: prior to his decision to escalate, and then after it, conditional on such an escalation having occurred. To isolate the size of mobilization from the prior probability of escalation, we shall always use the conditional estimate. This does not mean that we can avoid taking expectations: after all, there is residual uncertainty about S_1's valuation

[1] See Powell (1990, 58) for the concepts of crisis and situational stability, which correspond to the ones I develop in this book, and so I retain their names. The third concept, that of *ex post* escalation stability is new to MTM.

and different types use different war or coercion mobilizations. (This is not the case in the non-military models where everyone uses the same signal.) For example, to compute the conditional mobilization level when the continuation game's equilibrium is from Proposition B.2, we need to account for the three ranges of escalating types.

4.2 The Basic Simulation Setup

With so many moving components, visualizing the results can be quite difficult. To assist with the discussion, I will provide numerous graphs to isolate particular effects. Of these, we are especially interested in how S_1 will react given some distribution of power and the resulting probability that war will break out (crisis stability). Both of these quantities are heavily dependent on the residual uncertainty about S_2's valuation, the relative balance of the costs of fighting and of the audience costs, as well as the relationship between the valuations and these costs.

We have something of an embarrassment of riches in the sense that it is possible to run any imaginable scenario through the MTM to see the impact of varying any or all of its parameters. Because all of the variables involved are continuous, we are essentially facing an infinite number of potential combinations to explore. We can go well beyond the usual comparative statics which isolate the effect of a single variable while holding everything else static. This flexibility, however, presents a serious problem for deciding what to include here. To reduce this complexity, I will focus on several qualitatively distinct situations.[2] For the simulations that follow, I shall assume that $\overline{v}_i = 25$. I shall vary the costs of fighting, with $c_i \in \{\overline{v}_i/10, \overline{v}_i/5, \overline{v}_i/2\}$. (That is, set at 10%, 20%, and 50% of the maximum valuation.) For convenience, I shall refer to these values as *low*, *medium*, and *high*, respectively. For all simulations except ones that are specific to audience costs, I shall keep $a_i = \overline{v}_i/50$, which ensures that $a_i < c_i$, as required by Assumption 3.1, for all c_i from the set.

In Chapter 3, we discussed the definition of the distribution of power and the fact that the marginal impact of mobilization is stronger when opponents are lightly armed. For the simulations that follow, I will use the *baseline system militarization* in which M_1 is 10% of $\overline{v}_1$ and then vary M_2 to produce the entire range of possible values for the status quo distribution of power, $p(M_1, M_2) \in (0, 1)$. I will defer the investigation of system militarization itself for Chapter 5.

We would like to investigate how the intensity of interest in the disputed issue affects crisis stability. Intuitively, an actor's interest is peripheral if his opponent believes that he does not value the issue too much. Conversely, the interest is vital if the opponent believes that he values it highly. More formally, define S_i's interest as *peripheral* if the opponent believes v_i is distributed uniformly on $[0, \overline{v}_i/2]$, and as *vital* if it is uniform on $[\overline{v}_i/2, \overline{v}_i]$. Matching the two categories for each actor yields four scenarios to explore. When both players have peripheral interests, there is a *minor dispute*; when both have vital interests, there is an *acute crisis*. When one of the players has a vital interest and the opponent only a peripheral one,

[2] The programs (written in Aptech's Gauss) that run the simulations and generate the graphs are available from the author's website along with instructions on how to use them to explore the model beyond the presentation in this book.

the crisis has *high stakes* for that player. The baseline case, of course, assumes that v_i is distributed uniformly on $[0, \overline{v}_i]$.[3]

4.3 Comparing Threat Instruments

Before turning to specifics, let me note some important differences between the model of military threats and non-military mechanisms as they relate to escalation stability.

4.3.1 Functions of Coercive Instruments

We can distinguish among four distinct but related functions the military instrument can have: *communicative*, *committing*, *subverting*, and *preparatory*. In its communicative role, it can credibly reveal whether one is committed; in its committing role, it can create a commitment by rearranging one's own incentives to fight; its subverting role, it can undermine the opponent's commitment by rearranging his incentives to fight; and its preparatory role, it is just a prelude to war.

The coercive value of the instrument is a combination of these functions, and the overall impact depends on their interdependent effects. For example, the probability that a crisis will erupt in a war will be determined by the extent to which the two opponents are separately committed to fighting (each partially a function of S_1's mobilization), the probability that S_2 will resist (a function of her beliefs about S_1's resolve and her own commitment), the extent to which S_1 is willing to minimize this probability (a function of his commitment and the degree to which he offsets S_2's), and the credibility of his communication (a function of the willingness of low-resolve types to mimic S_1's mobilization). In other words, the use of the military instrument by any player affects the commitments of both adversaries as well as their beliefs, making the overall effect very hard to grasp without the aid of formal analysis.

Which of these functions will dominate S_1's decision-making depends on the particulars of the situation he is facing. The subverting role is most obviously paramount in the bluffing equilibrium. When S_2 has exposed herself to the possibility of assured compellence, S_1's primary task is to make resistance as unpalatable as necessary to induce the highest-valuation type of opponent to capitulate. As we have seen, the optimal compellent allocation is such that some types who use it are bluffers. That is to say, they are not committed to fighting at that level, and S_2 knows that this is the case. However, despite the positive probability that S_1 will actually capitulate if resisted, S_2 is unwilling to chance it because the risk is too high

[3] The cases with $\underline{v}_1 > 0$ require some extra care because the derivation assumed $\underline{v}_1 = 0$. In particular, the conditions (CC) and (NB) must be revised to handle the possibility that $\underline{v}^{\mathrm{c}} \leq \underline{v}_1$ and $\underline{v}^{\mathrm{a}} \leq \underline{v}_1$, respectively. In both cases, the conditions are satisfied when the inequalities hold because credibility is not a problem for the existing types. Further, the calculation of $\underline{\tau}$ and $\overline{\tau}$ must take into account the following scenario: $\underline{v}_1$ obtains a positive payoff from coercion using $\underline{m}(\underline{v}_1)$, his minimum credible mobilization level. This implies that it is this level that types would pool on because there is no need to go above it to induce credibility; that is, there is no need to ensure that $\underline{v}_1$'s payoff is zero in equilibrium. Hence, $\underline{\tau} = \underline{v}_1$ and we only need to find $\overline{\tau}$ such that $\widehat{m}(\overline{\tau}) = \underline{m}(\underline{v}_1)$, a straightforward calculation. It is quite possible that $\overline{\tau} > \overline{v}_1$, which means that S_1 pools on a common level regardless of valuation. Clearly there is no incentive for any type to increase mobilization because S_2 is already convinced that he is resolved (since $\underline{v}_1$ is resolved at $\underline{m}(\underline{v}_1)$ by definition). Any attempt to reduce spending causes S_2 to revise beliefs all the way down to being certain that S_1 is not resolved, which makes any such attempt unprofitable.

and because her payoff from war is too low. Neither commitment nor communication are as important for S_1 here as his capability to manipulate S_2's payoff from war.

Whereas the communicative role is not all that important in the bluffing equilibrium, it is clearly of primary significance in the pooling region that occurs in the overkill mobilization equilibrium. The optimal mobilization for any $v_1 \in [\underline{\tau}, \overline{\tau}]$ is $\widehat{m}(\overline{\tau})$. It exceeds both his minimum credible allocation (and so overshoots the committing purpose) and the optimal coercive allocation (and so overshoots the subverting purpose). Any such type overpays significantly to ensure that S_2 is properly impressed; that is, he can credibly communicate his established commitment at a higher cost by making it unprofitable for unresolved types to mimic this behavior.

Whereas the committing role is only peripheral for the bluffing equilibrium, it is the focus of coercive tactics which can occur in the coercive equilibrium. If S_1 happens to be among the types who mobilize $\widehat{m}(v_1)$, he is clearly exceeding the minimum credible allocation level. Unlike the overkill equilibrium, overshooting the committing purpose does not involve paying a premium to communicate the resolve in a believable way. Furthermore, the subverting role is balanced against the costs of ensuring S_2's capitulation, and S_1 settles for a strictly positive risk of war instead of an extremely costly certain peace. In this sense, the committing role dominates both the communicative and subverting functions.

Finally, the preparatory function is most evident in the war fighting equilibria. When S_1 chooses $m^*(v_1)$, he is not interested in coercing his opponent, be it at a positive or at no risk of war, for such tactics are too costly. Instead, he opts for a mobilization that does signal his resolve to fight but not at a level anywhere near what is necessary to make even the least resolved type of opponent to consider capitulating. In fact, this allocation represents the optimum for waging war, and as such the preparatory function subordinates everything else.

	Communicative	Committing	Subverting	Preparatory
Basic	limited	no	no	no
Sinking Costs	yes	no	no	no
Risk	limited	yes	no	no
Tying Hands	yes	yes	no	no
Military	yes	yes	yes	yes

Table 4.1 *Functionality of Coercive Instruments.*

Table 4.1 summarizes the functionality of the several instruments we have examined. The basic escalation and risk-generation models have only limited signaling functionality because the lack of flexibility of the instrument does not permit to signal commitment with certainty, and so bluffing cannot be eliminated. As an interesting exercise, let us ask what functions would another coercive instrument, say economic sanctions, have? Since they are costly to impose, they will have some communicative role. However, because these costs are sunk to the sender, they will not have a committing function, and since they do not affect the payoff from war, they will play no preparatory role. However, unlike pure sunk costs, sanctions do affect the payoffs of the target directly, and will therefore have a subverting role.

4.3.2 Escalation and Resistance

With either sunk costs or tied hands, S_1's escalation resolves all uncertainty about his valuation, and since that reveals his commitment, the probability of war turns entirely on S_2's resolve: if she happens to be prepared to fight, she resists and war is the inevitable outcome. All S_1 has to do is gauge this risk assuming that he can persuade S_2 of his commitment and then decide whether it is worth taking his chances. With the sunk costs signal, he would also have to subtract the payment to convey credibly the information about his resolve (tying hands is essentially free).

Using military threats is a lot more involved. Consider any non-bluffing situation where S_1's action does signal his commitment to fight if resisted. On the surface, it would appear that the result is identical to the two cases we just discussed: the *ex post* probability of war is wholly determined by whether S_2 is herself resolved at the new distribution of power. She resists, causing war, if she is, and capitulates otherwise. The crucial distinction here is that S_2's level of commitment depends on S_1's mobilization too, not just S_2's own preparedness for war. Since S_1's decision *affects the credibility of his opponent's commitment, he can choose the level of risk he wants to expose himself to.* A more aggressive military stance reduces this risk because it lowers the probability that his opponent's valuation will be sufficiently high to cause her to resist.

This manipulation of risk is very different from the randomized threat mechanism. Recall that in that case, escalation saddles S_2 with the choice between certain war and capitulation. That is, the probability of war is absolutely the same as under the other two mechanisms and depends entirely on S_2's valuation, which is outside S_1's control. In contrast, the residual risk where military threats are involved can be manipulated by S_1, even up to the point of its complete elimination. If he really wanted to, S_1 could make war so unpalatable that even the highest-valuation opponent would just give up rather than fight it.

The problem, of course, is that S_1 may be unwilling to go to this extreme because the military instrument does not come without costs of its own. If this were not the case, S_1 would go all the way up in his mobilization and ensure that his opponent capitulates. Even though S_1 can potentially manipulate S_2's resolve drastically, his effective ability to do so is limited by two factors: the costs of the military relative to his valuation of the disputed issue (which puts downward pressure on his allocations) and the necessity to establish a credible commitment (which maintains the upward pressure).

Recall from the discussion in Chapter 2 that tying-hands is equivalent to randomized threats in the sense that the escalation threshold for S_1 and the resistance threshold for S_2 are the same in both models. Because of this, I shall treat them as interchangeable in the rest of the discussion. I shall state all results in terms of tying-hands with the understanding that they hold for the randomized threats as well.

Only in the basic model does S_1 capitulate with positive probability after threatening in equilibrium, and it happens whenever a bluff is called. In the sinking-costs and tying-hands models, bluffing never happens in equilibrium at all. In the randomized-threats model bluffers escalate in a way that removes the option to capitulate entirely (and so they end in inadvertent war when S_2 resists). Finally, in the military coercion model bluffing only happens when S_2 is sure to capitulate. This means that the only outcomes of comparative interest are war, capitulation by S_2, and appeasement.

In the two non-military signaling models, the probability that S_2 capitulates conditional on a threat is the probability that she is not resolved for war at the existing distribution of power. This implies that these models all share the same escalation stability. In other words, once S_1 escalates, the risk of war is the same regardless of the signaling mechanism. Crisis stability, on the other hand, will vary because the threshold for escalation is different depending on the method of signaling. As we have seen in our discussion of tying hands, the risk of war with that mechanism is higher than with sinking costs. The reason for that is the higher probability that S_1 would escalate if he could tie his hands instead of sinking costs.

Unfortunately, there is not much more than we can do without making these models directly comparable to the MTM. In particular, we have to make conformable the assumptions about the support of S_i valuations. We shall assume that F_1 is the uniform distribution with support $[0, \overline{v}_1]$ and F_2 is the uniform with support $[t, u]$.[4]

4.4 Threats and Stability

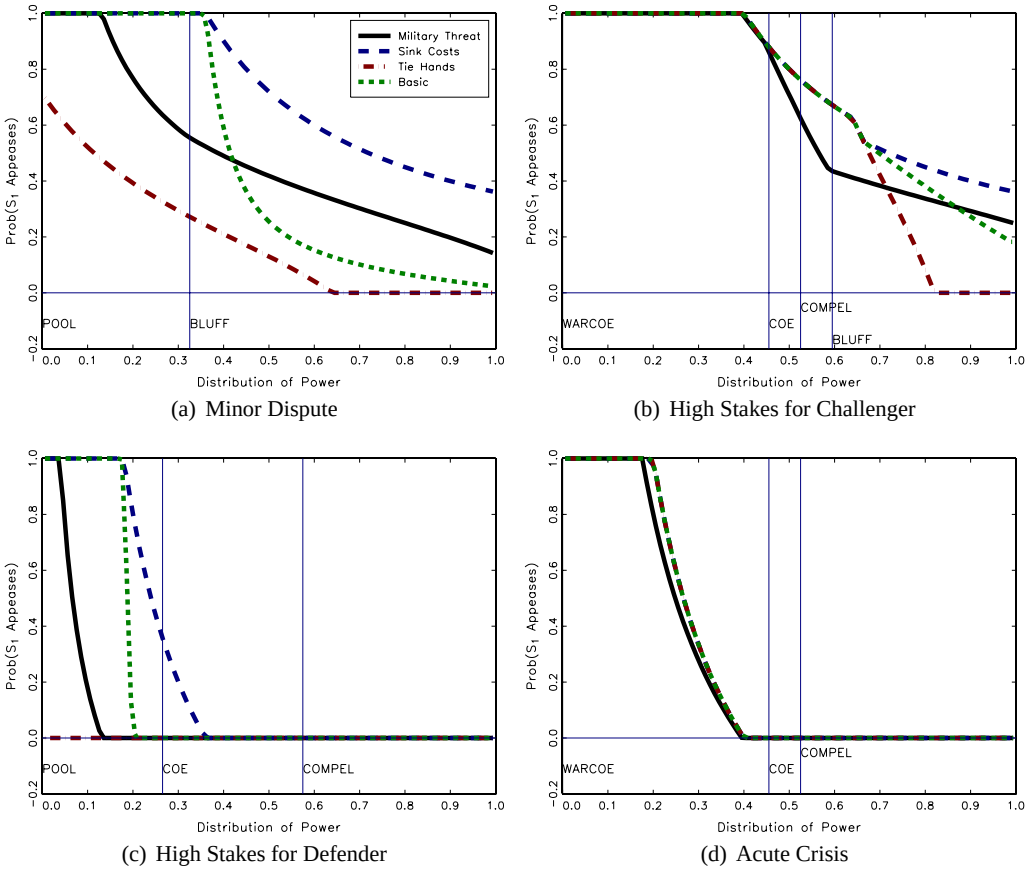
4.4.1 Appeasement

We begin by analyzing S_1's willingness to escalate the crisis. That is, when are defenders more difficult to deter? Figure 4.1 shows the probability that S_1 will appease in equilibrium for each of the four coercive instruments.[5] It is immediately obvious (and can be easily shown analytically) that the ability to tie hands makes the defender more aggressive than if he could only sink costs. This is, of course, precisely what Fearon (1997) finds, and the result follows directly from the fact that sinking costs involves a signal whose costs must be paid immediately whereas tying hands involves a costless (in equilibrium) signal where the only relevant issue is the defender's willingness to risk a war by committing himself to it with the signal. There are always types who are willing to run this risk but unwilling to pay sunk costs.

The basic escalation model is "in between" these two. The probability of appeasement would tend to be lower than under the sunk costs scenario because escalation does not incur direct costs, only a risk of having to back down or fight when S_2 resists. Hence, it is more attractive to lower-valuation types. However, it is not quite as attractive as tying hands for them because the ones who bluff would have to face the audience costs of backing down unlike the ones who have tied their hands by making war the better option. All of this is moot when the lowest-valuation type who can profit from credible escalation is resolved:

[4] This requires some care because the signaling results in the original models assume that (GT) is not satisfied. This was sufficient to rule out the genuine-threat equilibrium from Proposition A.1 because $v_2^* > 0 \Rightarrow F_2(v_2^*) > 0$, where the second inequality follows from v_2 distributed on $[0, \overline{v}_2]$. This is no longer true because when $0 < v_2^* < t \Rightarrow F_2(v_2^*) = 0$ and (GT) is not even defined. That is, depending on the distribution of types $[t, u]$ it may be the case that S_2 would resist for sure (the least-committed type's valuation is smaller than t). Equilibrium threats must be genuine because escalation would lead either to war or S_1's capitulation. Since bluffing is not an issue here, signaling is pointless.

[5] The informational pane at the bottom of each graph shows the equilibrium that obtains for the relevant range of distribution of power (DOP) values. The mnemonic designations are self-explanatory: WAR (Proposition B.1), WARCOE (Proposition B.2), COE (Proposition B.3), POOL (Proposition B.4), COMPEL (Proposition B.5), BLUFF (Proposition B.6), VE (Lemma B.4), and APPEASE (Proposition B.7).



(a) Minor Dispute (b) High Stakes for Challenger

(c) High Stakes for Defender (d) Acute Crisis

Figure 4.1 Coercive Instruments and Appeasement (baseline militarization, medium costs).

in this case the three mechanisms are equivalent because signaling is unnecessary for only resolved types find it profitable to escalate.

From our perspective, the interesting comparison is between the non-military signaling models and the MTM. Given the logic of military escalation, we should expect that the defender should be more aggressive than what he would have been if he could only sink costs—the military instrument functions as a commitment device, and its benefits beyond credible signaling make it attractive to lower valuation types. However, unlike tying hands, the costs of achieving these benefits must be paid regardless of outcome, which does provide a disincentive to the lowest-valuation types among those who would like to use the pure tying-hands mechanism if they could. In other words, the probability of appeasement in the MTM should generally be no higher than the sunk-costs one and no lower than the tying-hands one, which is precisely what the plots in Figure 4.1(a) and Figure 4.1(c) reveal.

However, as the two other figures, 4.1(b) and 4.1(d), show, this is not universally the case. In particular, when the stakes are high for the challenger, the ability to make militarized threats may lead to more aggressive behavior by the defender. It should not be too hard to see

why this is so: in these situations the defender is in a less advantageous position on account of facing a high-valuation opponent. Since the military instrument allows undermining of her resolve, a benefit over and above the informational role of the others, it becomes attractive when DOP is less in favor of S_1 than what it has to be for the informational role to have a bite. In other words, the defender will threaten with military escalation when the distribution of power favors him less than what it has to if only the informational instruments were available.

RESULT 4.1 *The defender is more likely to appease the challenger when escalation is militarized than when it involves only sunk costs. Compared to tying hands, however, appeasement when escalation is militarized can be more likely (when the challenger's stakes are low), less likely (when the crisis is acute), or dependent on the distribution of power (when the challenger's stakes are high).*

Although the trend is clear, it is worth noting that the differences among the threat mechanisms are most pronounced when the stakes are low, and least pronounced when they are high. Figure 4.1(d) is particularly striking in that respect: there is virtually no difference in S_1's propensity to escalate regardless of the instrument he uses to do so. This is a result of the combination of high stakes and moderate costs of war for the defender, which makes him very likely to press his advantage even at relatively unfavorable distributions of power, a tendency that is strengthened by the moderate costs of S_2, which make her a plum target for coercion. As we shall see later, however, the dynamics of the threat itself are almost as strikingly different as these probabilities are similar.

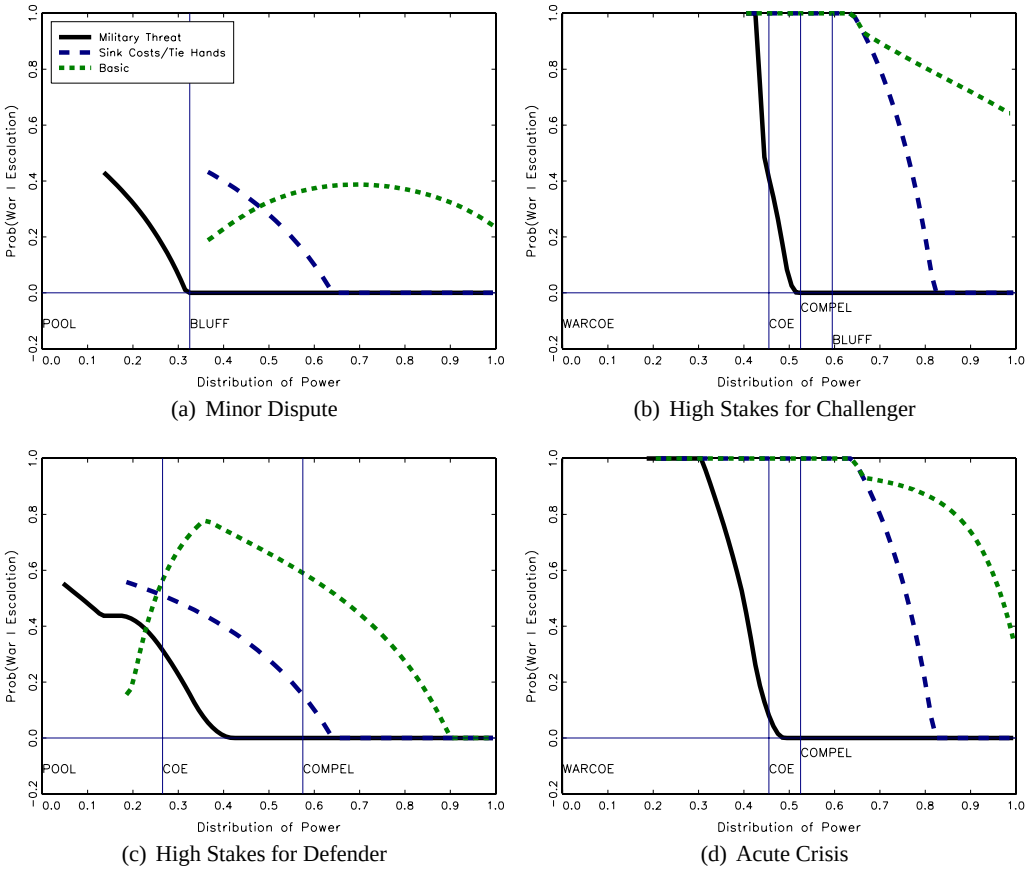
4.4.2 Escalation Stability

Which crises are more likely to end in war once the defender attempts to coerce the challenger? Figure 4.2 helps answer this very question. In sunk-costs and tied-hands models this conditional probability is the same because the signal fully reveals S_1's credible commitment to fight, all types who prefer war to capitulation resist. Because neither signal affects S_2's own commitment, the sets of types are equivalent for both types of signal. Even though basic escalation does not affect that commitment either, the conditional probability is different because bluffers would back down when resisted, and S_2 is more likely to resists given that there are bluffers.

Turning now to the MTM, we have our first qualitatively different results: regardless of the balance of interests, escalation, when it occurs, is more stable when it is militarized. This may appear quite surprising until we consider the logic involved. Non-military escalation reveals S_1's credible commitment to fight and escalation stability reduces to the probability that S_2 resists. Since in all non-bluffing equilibria of the MTM escalation is also credible, the stability also reduces to this likelihood. However, unlike the other two instruments, military escalation undercuts S_2's commitment, and makes her more likely to capitulate, which leads to a lower probability of war conditional on such escalation. In other words, the subverting function of the military threat has reduced the danger of war once appeasement fails.

RESULT 4.2 *Militarized escalation is at least as stable as non-militarized signaling when the distribution of power favors the defender, and is generally much more stable otherwise.*

The magnitude of the effect can be staggering. For example, consider the acute crisis



(a) Minor Dispute (b) High Stakes for Challenger

(c) High Stakes for Defender (d) Acute Crisis

Figure 4.2 Coercive Instruments and Escalation Stability (baseline militarization, medium costs).

scenario in Figure 4.2(d) and suppose the distribution of power is at 55%. *Using any of the three non-militarized threats leads to certain war conditional on escalation whereas the military threat leads to certain peace!*

Why is this the case? It is so because the expected mobilization here is quite large. This convinces S_2 that the threat is credible and manages to undo her commitment completely, and S_1 is willing to pay this cost because at this relatively advantageous DOP, it is possible to compel S_2 given her medium costs of war. Because non-militarized escalation cannot affect her commitment and the DOP is not by itself sufficient to make the highly resolved types of S_2 participating in this crisis willing to capitulate, the defender must face certain war. The interests here are so vital that S_1 does not even have to signal his resolve: the equilibrium signal under any of the two pure signaling models is zero, so credibility is not even an issue.

If we maintain high stakes for the challenger but assume peripheral stakes for the defender, the dynamic is very similar. The moderate costs of fighting for S_1 make it possible to commit credibly to war when DOP is high enough, and at 55% escalation certainly leads to war. Under the MTM, aggressive mobilization subverts S_2's commitment completely, and

the result is assured compellence. Of course, because S_1 is now in a less advantageous position on account of his peripheral interest in the issue, the mobilization required to achieve that compellence will be quite a bit larger (more than double). This leads to the following counter-intuitive conclusion:

RESULT 4.3 *When the defender uses the military instrument for coercion or compellence, he will be more aggressive when his interests are peripheral than when they are vital.*

In contrast to the acute crisis scenario where bluffing is never an issue, once DOP exceeds about 65% here, the defender must incur positive signaling costs even with non-military instruments to convince S_2 of his resolve. To see why this must be so, note that at DOP below 65%, S_2 is certain to resist even though she knows escalation to be credible. Because this implies that escalation automatically leads to war, this deters some low-valuation types from escalating when stakes are low for the defender but not when stakes are high. Of course, all else equal, an improvement in DOP can only make previously unresolved types committed. Hence, if S_1's escalation is credible at a lower DOP, it will also be credible at higher levels. This makes the positive signaling costs somewhat puzzling for it appears that S_1 should have nothing to signal about. Figures 4.2(b) and 4.1(b) provide the key to the answer. Note that once DOP exceeds 65%, S_2 begins capitulating with an increasing probability, leading to the more stable escalation we observe in the former plot. This naturally causes a decrease in S_1's propensity to appease, as seen in the latter plot. This now implies that some low-valuation types who previously were deterred from escalating by the certainty of war now find it profitable to threaten. When the increase in DOP is not sufficient to ensure that these types are resolved, bluffing becomes a possibility, so S_1 must now pay the price to convince S_2 of his resolve.

This problem does not arise in the acute crisis scenario even though S_2 begins capitulating with increasing probability there as well: the smallest valuation for the defender is so high that once he is resolved for fighting he escalates no matter what S_2's action is going to be; since there are no types with valuation lower than that, her increasing propensity to capitulate changes nothing from S_1's perspective (there are no low-valuation types to be tempted here as opposed to the case where S_1's stakes are low so such types always exist). This then implies that S_1's incentive to escalate is not strengthened in the sense of making escalation attractive to unresolved types, hence there is no need to incur any signaling costs.

These two cases further suggest that the military instrument's effect is most pronounced when S_2 has serious interests at stake, which makes her a difficult opponent to contend with. In particular, if the instruments only allow signaling, escalation stability will be extremely low because such S_2 is very likely to be highly resolved. But if military escalation is on the table, the very resolve of S_2 makes S_1 quite willing to attempt coercion and even compellence provided it is feasible. Either of these courses of action would stabilize escalation despite the apparent aggressiveness of the threat.

RESULT 4.4 *When the challenger's interests at stake are vital, signaling has a very limited role and non-military escalation will tend to be extremely unstable. Militarized escalation will tend to seek coercion or compellence, and will thus tend to be very aggressive but also very stabilizing.*

When the challenger's interests are peripheral, signaling regains some of its importance as an instrument of coercion. Because the types of S_2 involved in the crisis have relatively

low valuations, she will be reluctant to resist an escalation when doing so is certain to lead to war. Hence, escalation will tend to be much more stable, as shown by Figures 4.2(a) and 4.2(c). Of course, this also means that S_1 is less likely to appease, as demonstrated by Figures 4.1(a) and 4.1(c). Given S_2's propensity to capitulate, S_1's escalation is likely to run into credibility problems, which means that the equilibrium signal must be quite substantial.

RESULT 4.5 *Signaling has a coercive effect only when the challenger's interests are peripheral. Although all instruments enable credible revelation of information, military escalation is safest and the required mobilization will often be cheaper than a sunk-costs signal.*

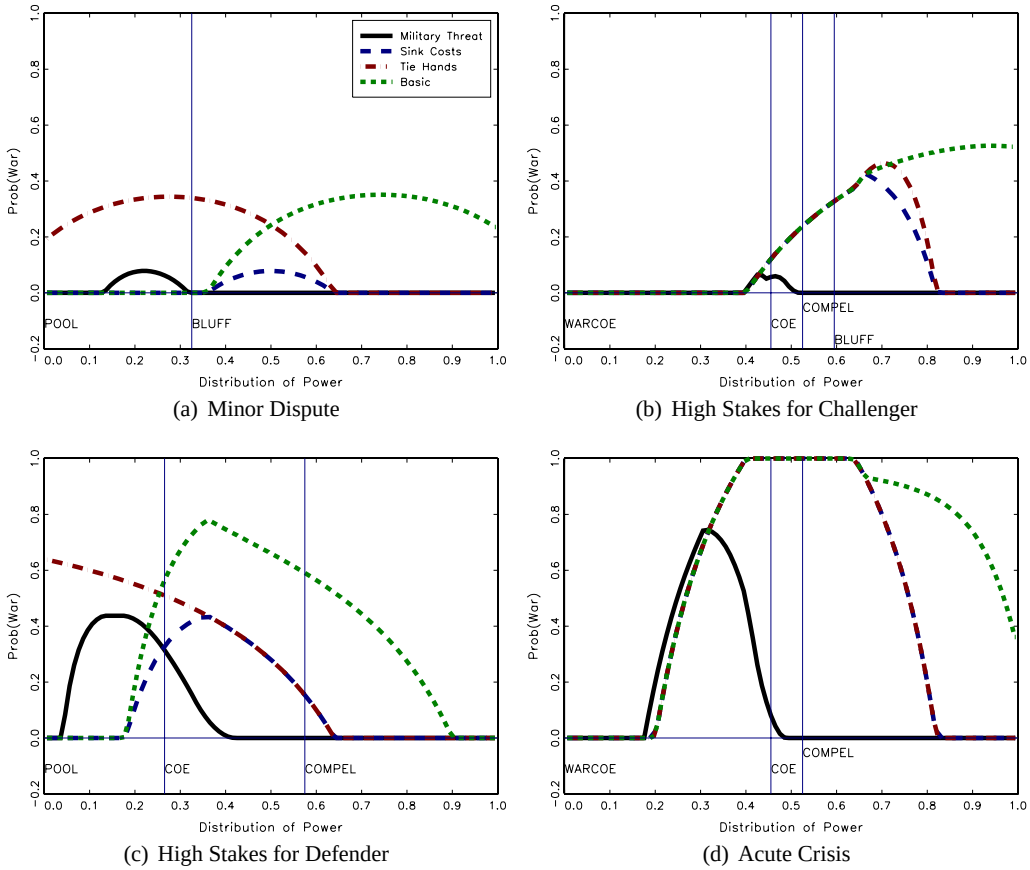
In general, then, when equilibrium mobilization is very large, the difference in escalation stability will be most pronounced because S_1 is in effect paying to increase the likelihood that S_2 will capitulate beyond what credible communication can accomplish. The spikes in mobilization are associated with rapid stabilization of escalation, as in Figure 4.2(b), because they increase the probability of S_2's capitulation.

4.4.3 Crisis Stability

We are now ready to answer the more general question: are crises more or less stable in the MTM than in the non-military escalation models? Figure 4.3 gives the four balance of interests scenarios we have been considering. The answer is a qualified "yes, mostly." Let us first look at the situations where the military instrument's effect is most pronounced. As we have already seen, these are the cases where S_2's interests are vital, as in Figures 4.3(b) and 4.3(d). The effect here is quite strong: crises will tend to be much more stable in the MTM than under in of the alternative models. Although this may come as a surprise to many (who have not studied the MTM dynamics as we have), the logic is actually quite straightforward. Because of her high interests in the issue at stake, S_2 is likely to resist absent any way to undermine her commitment. The military instrument does just that and this stabilizes the crisis overall even though it may not make appeasement more likely. The magnitude of the effect is astounding: between DOP of 40% and about 65%, an acute crisis will certainly end in war under any non-military escalation, whereas the probability of war declines rapidly (from 40% to zero) when the defender can militarize the crisis. A similar, albeit less dramatic, improvement in crisis stability can be observed when the defender's interests are peripheral. Even though in this case crisis stability is better overall, militarizing a crisis can drop the probability of war from over 35% to 0% (at DOP around 60%). Note further that militarization can destabilize an acute crisis relative to the other instruments only in the narrow band where DOP is between 20% and 35%, and the deterioration everywhere is quite small. Nowhere else does militarizing a crisis destabilize it. As expected, the worst-case scenario for the defender in both situations is to have neither the subverting power of the military instrument nor the signaling capability of the sunk-costs or tying-hands mechanisms: when straight escalation is the only option, the probability of war is highest.

RESULT 4.6 *When the challenger's interests are vital, crises that can be militarized will tend to be most stable and crises where signaling is impossible will be least stable.*

Turning now to the crises in which the challenger only has peripheral interests at stake, we can see that the stabilizing impact of military threats persists except when the distri-



(a) Minor Dispute (b) High Stakes for Challenger

(c) High Stakes for Defender (d) Acute Crisis

Figure 4.3 Coercive Instruments and Crisis Stability (baseline militarization, medium costs).

bution of power heavily favors the challenger, and even then stability will be better with a military threat than with a tying-hands one. In this region of DOP, sinking costs is quite unattractive because the required amount is excessively large. Hence, for very low values of DOP, S_1 will have to appease with very high probability, as seen in Figure 4.1(a), for example. Since the military instrument is more efficient, he is more willing to use it, which decreases his propensity to appease and destabilizes the crisis. However, once the distribution of power makes coercion or compellence possible, the advantages of utilizing the military threat quickly make themselves noted and the increasing probability of S_2's capitulation outweighs the higher likelihood of escalation, leading to an overall stabilization of the crisis (e.g., from DOP over 30%). An analogous dynamic occurs when the defender's stakes are high, as in Figure 4.3(c): militarized crises will be the most stable ones at DOP higher than about 25%. Tying hands is almost universally the most destabilizing tactic among the three signaling instruments. First, it always produces a probability of war that it at least as high as sinking costs and usually much higher. Second, it also almost always causes a higher probability of war than militarization as well (the only exception is in the acute crisis

scenario described above). Only straightforward escalation in the basic model can be worse than tying-hands in terms of crisis stability.

RESULT 4.7 *Militarizing a crisis reduces the probability of war relative to any other form of escalation except when the distribution of power is skewed in favor of the challenger. Even there, militarizing a crisis will generally be less destabilizing than tying hands.*

Thus, we have found that tying hands is the most destabilizing form of any escalation with signaling. There is hardly anything worse than managing to commit oneself to war in an essentially costless way while simultaneously failing to undo the commitment of one's opponent. In general, militarization is more conducive to peaceful crisis resolution. The advantage comes from the subverting role of the instrument which increases the odds that S_2 will capitulate above and beyond what mere credible information of S_1's commitment can. The stabilizing impact of military threats is most pronounced when the stakes are high even though these crises are least stable.

RESULT 4.8 *Increasing the stakes produces crises that tend to be less stable overall, but in which the stabilizing effect of militarization is most noticeable.*

To summarize, we found that militarization does dampen S_1's propensity to appease relative to sunk-cost signals but not relative to tying-hands signals. However, escalation stability is always higher with military threats than any of the other two non-military signals. Even though signal size (and therefore, its costliness) may be substantially larger in the MTM, overall crisis stability is much improved except perhaps when the distribution of power disproportionately favors the challenger and her interests happen to be peripheral. This is worth exploring in more detail.

4.5 Gaining More by Risking Less

We have now established that under very general conditions militarizing a crisis stabilizes escalation, often dramatically, relative to using the non-military coercive instruments. It might not be apparent just how discrepant this finding is from our existing theories of coercion. There is a long tradition in international relations theory which holds that the higher an actor's expected payoff from war (the stronger he is militarily), the better deals he should be able to command but the higher the risks of war he would have to run. Although this willingness to run serious risks started out as a behavioral assumption, it was later provided with game-theoretic microfoundations: stronger types *had* to run risks so as to discourage potential bluffers from mimicking their strategy. Since the results thus far are aggregates (e.g., the probability of war is obtained by integrating over all types that escalate towards a positive risk of war), it may not be obvious that the MTM dynamics violate the venerable tradition in international relations theory.

4.5.1 Power, Risk, and Gain in Crisis Bargaining

To understand the direct relationship between the expected payoff from war and the willingness to run risks, consider any of the signaling games in Chapter 2. In any non-military model, each type's expected payoff from war remains constant but varies by valuation, which

means that we can think of types in terms of their expected war payoffs: a type whose expected war payoff is high is stronger than a type whose war payoff is lower. Since players use type-dependent strategies, each type faces a risk of war generated by its equilibrium strategy and since we can identify types with their war payoffs, we can represent this equilibrium risk of war as a function of the type's expected payoff from war. The direct relationship can be expressed simply as follows: stronger types use strategies that generate non-decreasing probabilities of war.

This direct relationship has a long and distinguished history in crisis models in general. It started out as a behavioral assumption that players with higher payoffs from fighting are more likely to go to war (Bueno de Mesquita and Lalman, 1986). This *assumption* could be succinctly expressed as follows: "the probability of a violent escalation by an actor (nation) rises and falls as a monotonic function of its own expected utility from challenging another" (Lalman, 1988, 596). Although this relationship appeared intuitively appealing, it also seemed too strong to be asserted; one would want to see it emerge as consequence of equilibrium behavior from more basic underlying primitives. Morrow (1989b) demonstrated that it would indeed do so, at least in his formalization. Finally, Banks (1990) showed that the monotonic relationship between expected payoff from war and willingness to run risks in equilibrium must hold in a very large class of crisis bargaining models irrespective of the specification of their extensive form. Furthermore, he was able to demonstrate that this also implied that higher types would also obtain better peaceful settlements.

4.5.2 Why Stronger Types Must Run Higher Risks

Banks (1990) assumes a general crisis bargaining scenario, in which one of the players knows his own expected payoff from war whereas the other is uncertain about it.[6] The crisis may end either in war or in some peaceful outcome with a negotiated deal (there may be any number of those). Banks shows that regardless of the extensive-form game one chooses, incentive-compatibility requirements impose certain constraints on optimal behavior, and as a result "in equilibrium the probability of war is an increasing function of [the informed player's] expected benefits from war" (605). In other words, no matter what sequence of moves one uses to describe the interaction, if the model belongs to the class Banks studies, any equilibrium will exhibit this monotonic relationship between the strength of the player and his willingness to risk war. As the author succinctly puts it,

> while higher types go to war at least as often as lower types, they also receive at least as high expected benefits if no war is fought. [. . .] Therefore, in crisis bargaining situations, equilibrium analysis predicts the following trade-off between the gains from settling the dispute and the probability of war: as the expected benefits of war increase, the informed player receives a better negotiated settlement but in addition runs a greater risk of war (606).

This fundamental result follows from basic incentive-compatibility properties that equilibrium strategies must have. Since the very notion of equilibrium is that no type should be willing to deviate from its supposed strategy, it follows that in any equilibrium all types

[6] Even though Banks (1990) restricts himself to one-sided incomplete information, his results can be extended to environments with two-sided incomplete information, such as the one we have been using all along. For more on mechanism-design, see Fudenberg and Tirole (1991a); in particular, see Myerson (1979) and Myerson and Satterthwaite (1983) on the revelation principle and its application to bargaining problems.

should have the proper incentive to use their own strategies; that is, there should be some disincentive to mimic the behavior of other types. In practical terms, a weak type should not be willing to play the strategy of a strong type: mimicking the strong type's behavior should leave the weak type no better off than using its own strategy. (Of course, the strong type should not profit from pretending it is weak either. However, this is not an issue in any of the crisis bargaining models under consideration.) As it turns out, in this traditional environment, the only way to provide an appropriate disincentive for the weak types requires the strong ones to run larger risks of war.

To see how this logic works, imagine a crisis situation in which one player has private information about his expected payoff from war. The bargaining environment is such that actions do not alter the payoffs associated with any particular peaceful outcome in a type-dependent way. This is absolutely crucial to the result, so it is worth belaboring what it means. In general, we can reduce the outcomes in any crisis game to war and peace. That is, any strategies players use will end the game in one of these two ways only. The expected payoff from war is straightforward: in this environment, this payoff remains fixed and independent of the actions players take (which is why it is possible to index players types with it). The peace outcome is actually a probability distribution over all peaceful outcomes where the probabilities are conditional on no war occurring. The expected payoff from peace is then calculated using this probability distribution and the payoffs from each outcome in the usual way. To say that the peace payoff is not type-dependent means that if some type t's strategy produces a peaceful outcome with an expected payoff of x, then another type, t', can obtain this same x by adopting t's strategy. For example, suppose t's strategy produced a conditional probability q of concessions worth x, yielding an expected peace payoff of qx. Then t' could use t's strategy to obtain qx himself. Obviously, none of the signaling models satisfy this requirement: since types differ in their valuations, a concession worth x to type t is worth $x' \neq x$ to type t'. Mimicking t's strategy would then give t' an expected payoff of qx' instead. However, this is not to say that the class is not very general.

In any game, types can mimic each other's strategies. Since the peace outcomes are not type-dependent, a weak type can obtain the same expected peace payoff as a strong type by adopting that type's strategy. If this payoff is better than what the weak type would obtain from its own strategy, then it will have an incentive to change strategies, which should not happen in equilibrium. If the weak type is to stick with its own strategy, something must be lowering the expected payoff from adopting the strong type's strategy. Since weak types by definition have a lower payoff from war and because the peace outcome is not type-dependent, the only way a strong type could make its strategy unattractive to a weak type and simultaneously keep it relatively profitable for itself is to increase the risk of war. Should the weak type adopt such a strategy, it can benefit from the peace gain of the strong type but must run a larger risk of war, an outcome which is quite unattractive for it. This "balances out" the expected utility calculation and renders the weak type unwilling to mimic the behavior of the strong one even though doing so could yield a larger peace benefit: the risk of war it has to run by adopting the strategy wipes out the potential gain. Therefore, in such an environment, incentive-compatibility (the requirement that types stick with their own strategies) implies that equilibrium behavior must exhibit a monotonic relationship between expected payoff from war and risk of war.

To paraphrase this result in a manner more convenient for subsequent use, in equilibrium,

a stronger type cannot obtain a better expected peace payoff at a lower risk of war than a weaker type. Lebow (1981, 264), who is far from sympathetic to these types of models, agrees with the conclusion:

> In crisis, nations may have to demonstrate willingness to go to war in order to prevent war. This fundamental axiom of crisis is also the most paradoxical axiom of crisis, because the very actions designed to convey willingness to risk war can also make war more likely by courting loss of control over policy. Escalation, mobilization of public opinion, and the dispatch of ultimata are all cases in point. They are credible indicators of willingness to fight precisely because they court loss of control.

As we shall see, this contradicts the most fundamental finding of the MTM, which shows quite unequivocally that equilibrium strategies of higher types may involve lower risks of war. The immediate reaction would be to say that the general results do not apply to the MTM because it does not belong to the class of games satisfying the assumptions. However, neither do the other signaling models we have studied but it turns out that they, too, exhibit the monotonicity. Hence, there is more going on in the MTM than just violating the type-independent peace outcome assumption. Before seeing precisely what, it is worth examining why monotonicity obtains in the non-military models.

4.5.3 Risk and Gain with Non-Military Threats

We have already noted why Banks's (1990) environment does not include the non-military models, so we need to examine how strategies maintain incentive-compatibility in them. If escalation convinces S_2 that S_1 is resolved, then S_2's probability of capitulation must increase. This now makes escalation attractive to some unresolved types who will bluff in attempt to take advantage of that. In equilibrium, these types must be discouraged from escalating. Except for the risk-generation game, S_1 controls the decision to go to war, which means that the disincentive cannot come from running higher risks: weaker types will opt to capitulate at the final node rather than fight. But in the environment studied by Banks (1990), this risk provides the only disincentive that keeps weaker types from mimicking the strategy of the stronger ones: without it, using that strategy is costless and since it yields better expected outcomes, weaker types have an incentive to use it.

As we would expect at this point, the costly signaling models provide strong types with additional tools at their disposal. In particular, unlike the models that comprise the general class in Banks's (1990) study, the signaling games do not allow for costless mimicry even when the risk of war generated by the strategy remains the same. Strong types exploit their other ways of communicating commitment and manage to overcome some of the informational disadvantage.

In the sinking-costs model, the necessary disincentive in the intuitive equilibrium is provided through paying costs associated with escalation that are just large enough to make the least resolved type indifferent between escalating and staying with the status quo. Any weaker type that mimics this strategy would have to capitulate if S_2 resists. However, the expected payoff from such a bluff is strictly worse than the status quo because the costs that the weak type would have to pay are too large to make it profitable. When the least resolved type is indifferent, any weaker type will necessarily be worse off if it escalates because of its

smaller capitulation payoff. In other words, strong types choose a strategy that is too costly for weaker types to mimic given the benefits they would reap from doing so.

In the tying-hands model, the necessary disincentive is provided through large audience costs which the unresolved types would have to pay if they bluff. Since resolved types fight when resisted, the size of these costs is irrelevant for them. Consequently, they incur audience costs that are prohibitively high for weak types to mimic, which prevents these types from escalating. The stronger types exploit their willingness to fight at the final node to drive a wedge between themselves and weaker ones. Doing so is a bit more complicated than the sunk-costs model because the very act of incurring audience costs makes some types resolved, but as the model shows, it can be done.

The risk-generation model is closest to the traditional environment. In equilibrium, resolved types pick strategies that are too risky for most, but not all, unresolved types to mimic. Since the probability of war is partially determined by S_2's willingness to resist, some types may bluff hoping that she would capitulate. (This would not be so if the randomized threat "kicked in" before her decision; then only resolved types would make it.) Resolved types have no way of distinguishing themselves from these bluffers through their actions because there are no additional costs, for example, that they could incur. As it turns out, however, in equilibrium they do not have to pay a price for this inability: since their strategy involves a sure commitment to war should S_2 resist, she behaves as if any type who escalated were genuinely resolved to fight. Consequently, her capitulation probability stays the same regardless of the possibility that she might be facing an unresolved opponent: that type has no choice to act on its temptation to back down if resisted.

In all of the non-military signaling models, stronger types will run the risk of war but without additional options at their disposal, the risk involved is not sufficient to prevent weaker types from mimicking their behavior. As a result, all of the action in these models is toward strong types improving their payoff by obtaining higher capitulation probabilities of S_2, which they can only do through convincing her that escalation is genuine. Sometimes this eliminates all bluffing altogether, as it does in the sinking-costs and tying-hands models, and other times it does not because the strategy renders providing the disincentive moot, as it does in the risk-generation model.

The interesting point in all this, however, is that even though the signaling models are not in the general class analyzed by Banks (1990), the equilibrium strategies produce, yet again, the familiar monotonic relationships. Observe that in all three non-military signaling games, the equilibrium risk of war conditional on S_2's resistance is 1, and the probability that S_2 resists conditional on escalation is the same (which is not surprising given that resistance leads to certain war). In other words, stronger types in all these models face the same, strictly positive, probability of war. Weaker types, on the other hand, either do not escalate at all—meaning zero risk of war—or face the same risks as the higher types (randomized threats model). This means that higher types do obtain better peace outcomes conditional on no war (S_2 capitulates with strictly positive probability, an outcome that weak types do not obtain because they appease), but do so at a higher risk of war. The distinction is sharpest when drawn between appeasing and escalating types, but even within the latter set monotonicity obtains: stronger types among the escalating ones get better payoffs and the risk they run is no lower than the risk run by any other escalating type. That is, no strong type can obtain a better payoff by running smaller risks. In other words, the monotonicity results obtain even

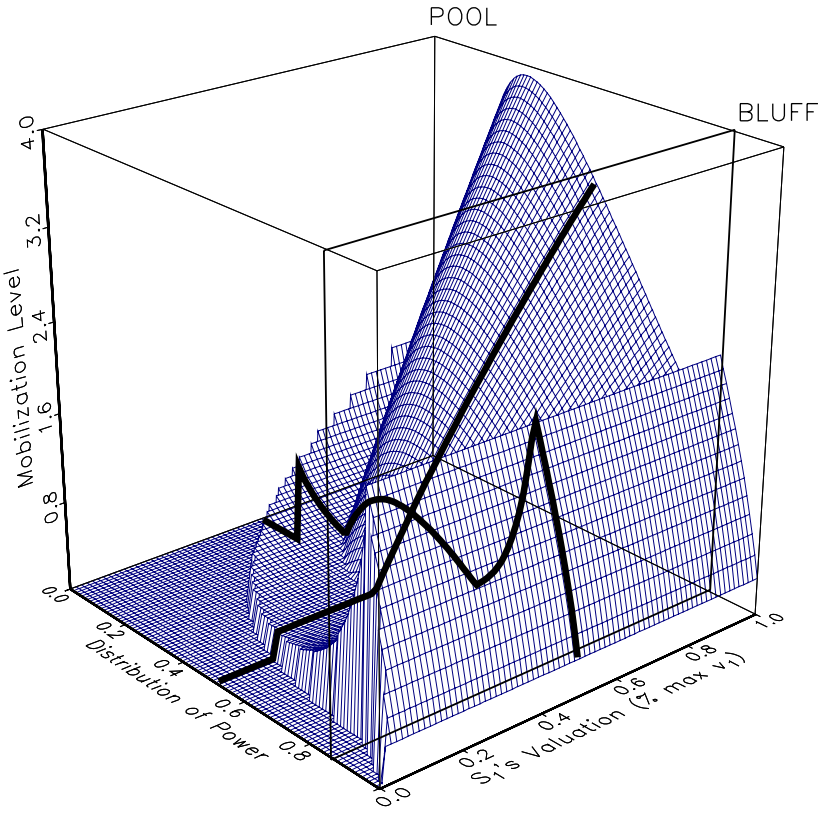


Figure 4.4 Type-Dependent Mobilization and the Distribution of Power (baseline system militarization, baseline balance of interests, low costs).

in these signaling games. This means that the discrepancy between all of these models and the MTM is not simply due to the type-dependence of peace outcomes. As we shall see now, it is the unique dual character of a military threat that is responsible for MTM's strikingly different implications.

4.5.4 Risk and Gain with Military Threats

To see how the MTM differs from the ones enumerated above, we must look at the optimal behavior of S_1, which depends on his (privately known) valuation. To this end, I compute the type-dependent expected mobilization levels for all possible distributions of power (DOP), as shown in Figure 4.4. The configuration of parameters results in the pooling equilibrium from Proposition B.4 for DOP up to about 85%, and the bluffing equilibrium from Proposition B.6 over the rest of the range.

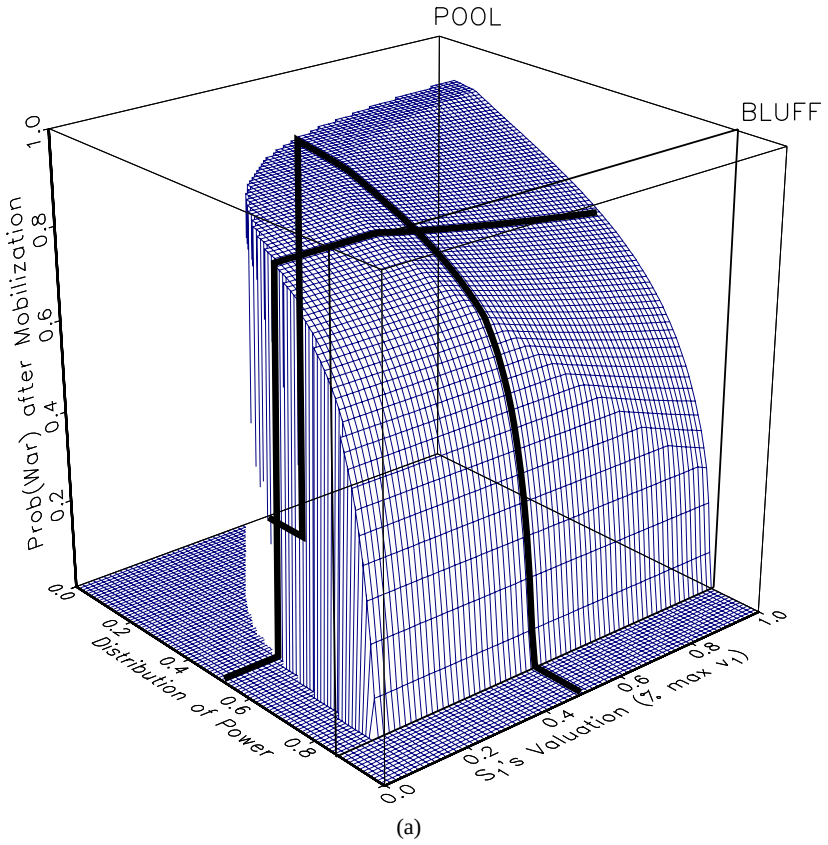
Consider first S_1's behavior as a function of the distribution of power. I will illustrate it for a type with valuation set at 50% of $\overline{v}_1$ whose mobilization level is indicated with a bolded curve. The DOP determines how effective any particular level of mobilization will be in undermining S_2's payoff from war, which in turn determines the probability of war S_1

would face (shown in Figure 4.5). Very low DOP below 15% are quite disadvantageous for S_1 and require high (and costly) mobilizations to coerce S_2. They are so unattractive that this type would not even mobilize for war; appeasement is the only rational outcome, so the probability of war is zero. An improvement in DOP permits higher valuation types to engage in credible coercion, which in turn "pulls in" some types with moderate valuation (ours among them) in the pooling region. As DOP improves further the coercive pooling level decreases because more types become resolved at their optimal coercive mobilizations. The risk of war in this region is very high (around 90%). Once DOP reaches a moderate level of about 30%, our type can credibly commit with his own coercive mobilization, and he is no longer among the pooling types. As DOP improves to about 45%, he increases his mobilization to lower the risk of war (to about 75%). When DOP improves beyond about 45%, he can attain additional reductions even with smaller mobilizations. Unfortunately, somewhere around DOP of 70% this increasing advantage opens up the possibility for lower-valuation types to bluff, which this type must discourage by mobilizing overkill capability—hence the gradually increasing levels and the resulting dramatic drop in the risk of war. From DOP of around 85% he can finally compel S_2 to capitulate with certainty, and the required mobilization declines as DOP improves further. The probability of war is zero throughout.

We can now address the question of gains and risks. Consider the bolded curve at DOP of 50% in Figure 4.4: it traces the mobilization of S_1 as a function of his valuation when the distribution of power is fixed. In Figure 4.5 the analogous curve traces the risk he must run, and in Figure 4.6 it traces his expected payoff. Observe that *equilibrium mobilization levels are non-decreasing in S_1's valuation:* higher-valuation types never mobilize less than lower-valuation types. In this case, types with valuations less than about 15% of $\overline{v}_1$ do not escalate at all: appeasement is riskless and yields a payoff of zero. Types with valuations that are higher than 15% but lower than about 40% of $\overline{v}_1$ pool on a common coercive mobilization. They run a substantial risk of war (about 85%) but their payoffs are higher, and increasing in type. Types with valuations higher than that mobilize at their optimal levels for coercion. Since these are increasing, the probability that S_2 will capitulate increases as well, so the probability of war decreases. This unequivocally benefits these types: the expected payoff is strictly increasing in their valuations.

Now that I have explained the logic, all of this should be straightforward. It may therefore come as a surprise that these results contradict the fundamental conclusion from nearly all existing crisis bargaining models. In particular, even though the pooling types must pay for their higher gains by running larger risks than appeasers (which hews to the traditional logic), the coercing types obtain even higher gains but run substantially lower risks than the pooling types (which flatly contradicts tradition). Since this is such a departure from received wisdom, it is incumbent upon me to investigate the reason for this discrepancy and explain why this should be so.

This first feature of the MTM that makes it very different from the other models is the *endogenous distribution of power*: the expected payoff from war depends, at least partially, on actions the players take during the crisis. Observe that because optimal mobilization levels are non-decreasing in type and the expected war payoff is strictly increasing, we can say that a type is stronger than another if his valuation is higher. The traditional logic is that higher-valuation types obtain better expected peace payoffs but at risks of war no lower than those incurred by lower-valuation types.

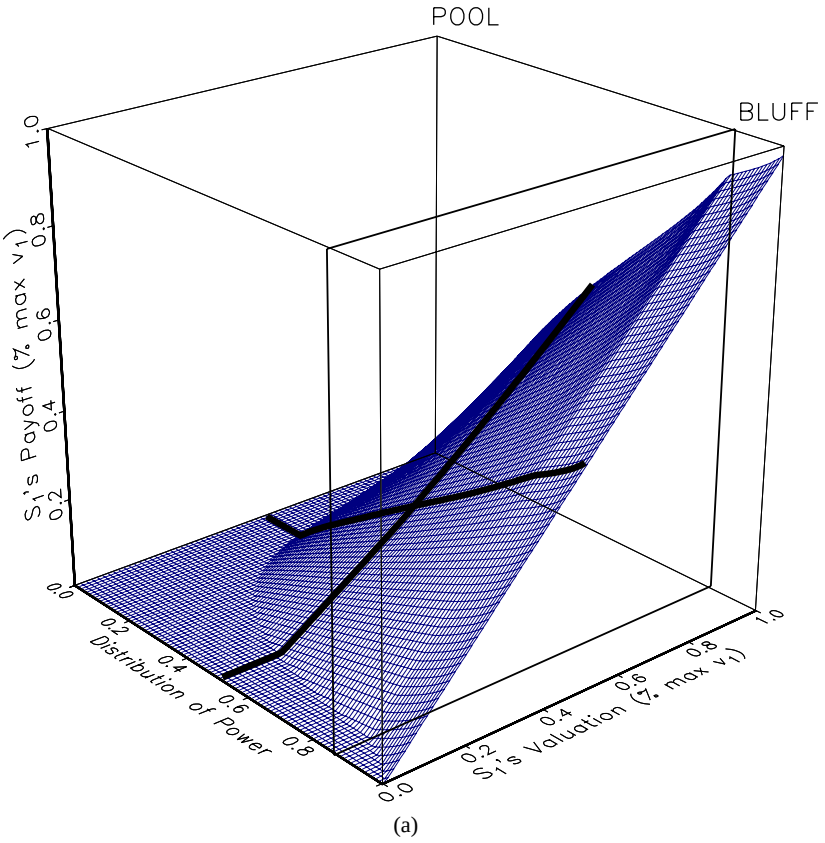


(a)

Figure 4.5 Type-Dependent Expected Probability of War (baseline system militarization, baseline balance of interests, low costs).

Although we can clearly see that this is not the case in Figure 4.5, it should be apparent from a simple examination of the equilibrium coercive mobilization levels. Ignoring for the moment the pooling types, recall that $\widehat{m}(v_1)$ is strictly increasing in type, which implies that higher-valuation types who mobilize for coercion do so at higher levels. Take now two types from among these such that $v_1 < \hat{v}_1$. Observe that if $\hat{v}_1$ mobilized at $\widehat{m}(v_1)$ rather than his own optimal level, he would obtain a probability of S_2's capitulation associated with the weaker type's allocation. Since $\widehat{m}(\hat{v}_1) > \widehat{m}(v_1)$ and both mobilizations are credible, using his own optimal level ensures a bonus bump in that probability: even though there is no additional information gain to be had, the higher mobilization undercuts S_2's expected payoff from war, which reduces the range of S_2 types willing to resist. Consequently, the probability of capitulation increases, and the risk of war goes down. In other words, *the expected payoff from peace conditional on no war is higher for the stronger type and it is obtained at a lower risk of war.*

RESULT 4.9 *Unlike all other crisis bargaining models, in the MTM an escalating type with a larger expected payoff from war will run a lower risk of war than an escalating type with a smaller expected payoff from war.*



(a)

Figure 4.6 Type-Dependent Expected Payoff (baseline system militarization, baseline balance of interests, low costs).

This relationship is even more evident if we compare compelling types to, say, coercing ones. The stronger types who can afford to mobilize for compellence obtain higher expected payoffs from peace at zero risk. The only sets of types for which monotonicity obtains involve the extremely weak ones who appease: any of the stronger types who escalate for coercion or war must run larger risks in order to secure their better peace payoffs. However, even here monotonicity breaks down: very strong types who opt for compellence can obtain their peaceful returns at no risk whatsoever. Of course, in the bluffing equilibrium, the risk of war is zero and so a trivial version of monotonicity obtains.

In sum, the MTM does not yield a neat monotonic relationship: the risk of war is neither increasing nor decreasing in the expected payoff from war. Instead, if the risk of war is positive in some equilibrium, then it is zero for the weakest types (which appease), then strictly positive and constant (for certain war types) or decreasing (for coercive types), then zero again for all compelling types.

RESULT 4.10 *In the MTM, very weak types run no risks, stronger types run strictly positive, but generally decreasing, risks, and the strongest types also run no risks.*

Why do these results obtain? Recall that in the non-military models, the risk of war is almost solely a function of S_2's beliefs about S_1's resolve: the best S_1 could hope for is persuading his opponent that he would fight if resisted. Once this information is credibly communicated, the probability of war turns on S_2's own level of resolve, which determines if she is willing to fight given that resistance would lead to certain war. Since there are types who prefer war to capitulation, this risk is irreducible. The only role of costly signals is to, well, signal; that is, beyond revealing information in a credible way, S_1 had no influence on S_2's decisions. In that setup, it is perhaps more appropriate to talk about *persuasion* rather than *coercion*.

The crucial difference between these models and the MTM is in S_1's ability to affect S_2's payoff from war through his own mobilization. On one hand, escalation still serves the signaling role it has in other models: that is, by the time S_2 gets to respond, she is generally certain that if she resists, war is sure to follow.[7] On the other hand, militarization goes beyond persuasion in that it can effectively discourage from resisting some types that did prefer fighting to capitulation under the pre-crisis DOP but no longer do so once the crisis develops. This now is true coercion: any type of S_2 that is resolved under the original distribution of power but that capitulates in equilibrium has been forced to do so by S_1's military preparations. Since no resolved type would capitulate in any of the other models, this provides the clue to the uniqueness of MTM's dynamics: *Militarization decreases the resolve of the opponent, so it can lower the likelihood of resistance beyond what revelation of one's own resolve can even when it is credible.*

This implies that mobilizing more should, all else equal, lower the risk of war because it should increase the probability that S_2 capitulates. The *all else equal* qualifier is necessary because it could be the case that stronger militarization perversely leads S_2 to update her beliefs in the opposite direction: if she comes to think that S_1 is more likely to capitulate, she would become more likely to resist even though her own resolve is lowered. The benefits of military coercion that S_1 obtains by lowering S_2's resolve seem to conflict with the credible communication of his own resolve: mobilization is so decisive that even unresolved types may attempt it, which, perversely, may undermine its effectiveness.

As we have seen, however, this does not happen, and the key to the higher-valuation types' ability to separate themselves from potential bluffers is in the costliness of mobilization. Whereas mimicking their high mobilization levels can benefit weaker types, they cannot reap all the benefits from pretending they are strong because even though the strategies are a lot less risky, they are also a lot more costly. It is their willingness to incur these costs that permits the higher types to distinguish themselves from bluffers. This then renders their escalation credible. The best example of this is provided by the pooling types in the "overkill" scenario: they pay extra costs to mobilize above and beyond what would have been otherwise optimal in order to make their escalation credible by placing it out of the reach of weaker types.

To obtain the non-monotonic relationship between expected payoff from war and risk of war in equilibrium, it is necessary that the coercive instrument is costly to use and that it

[7] Except, of course, in the bluffing equilibrium, where she capitulates even if she is not quite certain whether escalation is genuine. However, as we have already seen, the only reason resolved types do not separate themselves from unresolved ones is that there are no gains to be had from revealing the information.

can affect the opponent's payoff from war. To understand why this should be so, consider what would happen if one of the features were missing. If mobilization did not affect S_2's resolve, then it would be a pure sunk cost, and we have already seen the results in that case. If it were costless, then any type would have an incentive to mobilize at least up to the level that achieves assured compellence. In either case, monotonicity would be restored.

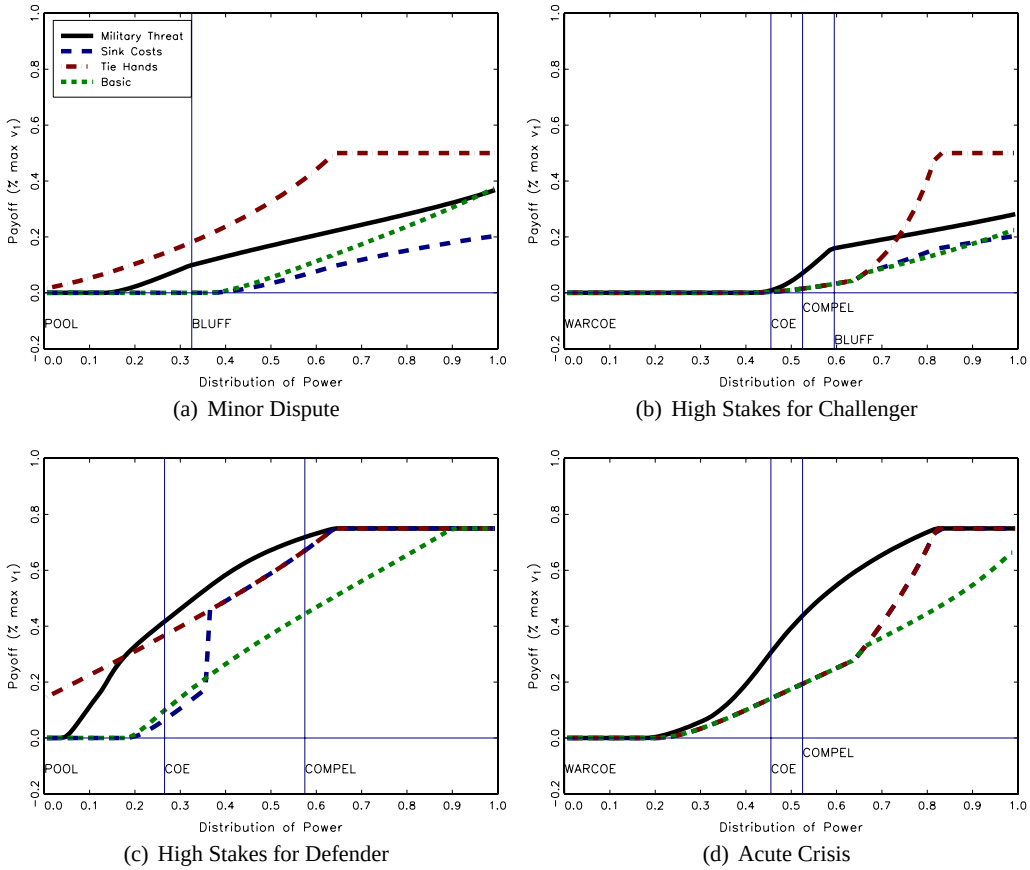
It is worth noting that higher types can enjoy smaller risks of war in equilibrium only when their behavior is truly coercive; that is, when their mobilization undermines sufficiently the resolve of the least-resolved type of opponent. Take, for example, a scenario in which both war preparation and mobilization for compellence are possible, and observe that the types who escalate toward certain war do not mobilize in a coercively meaningful way: the change in S_2's expected payoff from war that they produce is minor and too small to undo the resolve of even her least-committed type. Even though escalation signals their resolve credibly, the probability of S_2's resistance remains fixed at one. Again, the costliness of the instrument is critical because it prevents these types from upping the ante sufficiently to lower the risk of war.

4.6 Choosing the Instrument of Coercion

When it comes to deciding which instrument of coercion to employ, we have somewhat conflicting findings. On one hand, we found that escalation stability (and in most cases crisis stability as well) is better when the defender can use military threats. On the other hand, the equilibrium mobilization levels that he must use to accomplish this can sometimes be quite high. The first effect makes the military instrument more attractive but the second suggests that perhaps it is too costly to use profitably. To decide which of the available instruments S_1 would prefer, we need to look at the payoff he expects in equilibrium in the three models. Figure 4.7 plots the *ex ante* expected payoffs; that is, the payoff computed from S_1's perspective before he learns his own type. This way the plots incorporate the probability of appeasement. (Appendix C shows how these payoffs are computed.)

As we know already from Fearon (1997, Proposition 3), the expected payoff will be higher when the defender ties his hands than when he sinks costs. This holds regardless of the distribution of power or the balance of interests and follows directly from the fact that in equilibrium S_1 does not have to pay the audience costs.

Turning now to the MTM, we see that despite the stabilizing benefits of military threats, the defender may or may not want to use them if other options are available. In fact, when the benefit conferred by making these threats is relatively negligible (as in a minor dispute), mobilization costs will provide a large enough disincentive compared to tying hands. Figure 4.7(a) shows that S_1 does strictly better by tying hands for all distributions of power. The problem here is that in such a dispute there are strong incentives to bluff (the presence of very low valuation types is the culprit). To overcome this with the military instrument requires compellent allocations that can be quite costly. Since there is no restriction on the costs one can generate by tying hands (they are never paid in equilibrium), S_1 can signal his resolve much more effectively using that instrument. Clearly, the least attractive instrument is sinking costs, which are both expensive and do not yield the compellent effect of military threats. Potential bluffing is also a problem when DOP favors the defender in Figure 4.7(b)



(a) Minor Dispute (b) High Stakes for Challenger

(c) High Stakes for Defender (d) Acute Crisis

Figure 4.7 Coercive Instruments and Expected Payoff (baseline militarization, medium costs).

where the stakes are high only for the challenger. Again, tying hands is the more attractive method of signaling resolve than mobilization.

In contrast, when mobilization has a huge impact in terms of improving escalation stability, the defender will be willing to incur the costs to induce S_2 to capitulate with higher probability. As we have seen in Figure 4.2(d), the effect is most pronounced when the crisis is acute, and as Figure 4.7(d) shows, this is precisely when the defender will find military threats most attractive. Observe further that the analogous dynamic over a compressed range of DOP occurs when the stakes are high only for one of the actors. This follows from the fact that the benefit is attainable over that smaller range, and that it costs more to get it, which accounts for the reduction in the expected payoff.

RESULT 4.11 *Military threats are only useful to the defender when their coercive effect is large enough to justify the significant costs he has to pay to make them.*

We can actually go further than this general point. In fact, we know that the defender should always be able to get at least as much with military threats as he can with sinking

costs whenever he has to incur positive costs. This is because mobilization is a sunk cost with the added benefit of undermining S_2's willingness to resist. Hence, whatever S_1 can achieve by sinking costs, he will be able to achieve with military threats, and usually he will be able to get much more.

Similarly, military escalation is always at least as good as basic escalation. Although this is most evident in the verbal escalation equilibrium where the two are equivalent, it is also the case for any other type. When S_1 cannot even signal, he must cope with a higher probability that S_2 will resist (because she believes he might be bluffing), which automatically implies a higher probability of war. Since he cannot affect his payoff from war either, this gives the military instrument a double advantage.

Hence, the only instrument that is possibly more advantageous to the defender than the military is the tying-hands one. However, as the plots reveal, this will be so under specific circumstances only. For example, when the actors are involved in a minor dispute, credible revelation of information has a lot of coercive power, which implies that the least costly credible signal should be the most attractive. The military instrument is strictly less useful than tying hands for all distributions of power. The advantages of tying hands decrease substantially once the subverting impact of the military instrument has a bite.

RESULT 4.12 *Generally, the defender will prefer to tie his hands in minor crises or when the distribution of power is highly skewed. The higher the stakes and the less skewed the distribution of power, the more attractive military threats become. The defender will always prefer military threats to either sinking costs or escalation without additional preparations.*

Recall that militarization stabilizes crises and although the defender does not unequivocally prefer to minimize the risk of war, military threats are the most attractive instrument except in minor disputes. Most importantly, he will always want to militarize acute crises.

RESULT 4.13 *The more intense the conflict of interest, the more likely is the defender to militarize the crisis for reasonable distributions of power.*

This is point worth repeating: Our definition of an acute crisis is *not* based on whether the defender employed military threats but on whether both sides value the issue highly. The findings explain why the defender should be expected to use military threats in this scenario. The fact that the defender becomes more likely to resort to military threats the more intense the conflict of interest becomes provides a rationale for identifying serious crises with those in which we empirically observe such escalation. In other words, it is perhaps not tautological to conclude that a crisis is serious on the basis of observable behavior: if interaction is militarized, then we can infer the crisis is not a minor dispute.

4.7 Tying the Knot of War: A Conjecture

In this chapter, we found that military threats lead to dynamics in the MTM that can often be strikingly different from any of the other signaling mechanisms. Whereas the military instrument can be thought of as being an intermediate case between the two pure signaling mechanisms, a combination of sinking costs and tying hands, its implications for crisis stability, among other things, cannot be obtained by thinking of it as being "in between" the

other two. The dynamics of military threats we uncovered also suggests another way to look at the causes of war.

Blainey (1988) argues that war must be explained in terms of deliberate choices by state leaders. In the MTM, war can come in two ways: either S_1 simultaneously finds coercion too costly and appeasement too unpalatable and prepares for war without any attempt to reduce its risk, or else he mobilizes to force S_2's capitulation and fights when it does not work. From our perspective, it is the road to war through coercion that is most interesting: it is here that S_1 would genuinely like to avoid fighting and would actually try to but may fail, and in the process may end up creating an environment where war is certain. The MTM, therefore, suggests a cause of war under uncertainty that operates in a two-step fashion: *physical actions states take in order to coerce their adversaries and communicate their own resolve may, perversely, create an environment in which war becomes inevitable even when the uncertainty is resolved.*

Asymmetric information causes actors to risk committing too much (so they would not want to back down if resisted) but not quite enough to force their opponent to back down (and so the opponent resists). While the lock-in occurs because actors have private information and incentives to misrepresent, war occurs because actors find it the better option in the new environment even after all information has been revealed. The tragedy of crisis bargaining in the shadow of power is that actors may end up creating the circumstances that make war the best choice, circumstances they would have loved to avoid, and ones they would have avoided had they possessed complete information from the very beginning. The following example illustrates the logic.

EXAMPLE 4.1. Let $\overline{v}_i = 25$, assume baseline balance of interests, baseline system militarization, high costs for S_1 and low costs for S_2.[8] The solution is the coercive equilibrium with pooling: all types with $v_1 < \underline{\tau} \approx 16.19$ appease, all types with $v_1 \in (16.19, 22.39)$ will pool on a common mobilization level, $\widehat{m}(\overline{\tau}) \approx 4.66$, and all others mobilize at their unique coercive levels, $\widehat{m}_2(v_2)$. The threshold valuation for compellence exceeds $\overline{v}_1$, so no existing type will manage to mobilize for assured compellence.

Suppose now that the true valuations are $v_1 = 0.75 \times \overline{v}_1 = 18.75$, and $v_2 = 0.6 \times \overline{v}_2 = 15$. Given his valuation, S_1 will mobilize credibly for coercion with $\widehat{m} \approx 4.66$. Since S_2 capitulates if $v_2 < v_2^*(m) \approx 7.73$, she will resist given her valuation. At the last node, S_1 will fight if $v_1 > v_1^*(m) = 16.19$, so he attacks when resisted (in fact, S_1 will be resolved to fight for any $m \gtrapprox 1.94$). In other words, these two types of opponents will certainly end up fighting in equilibrium. Observe that at the time of his final decision, S_1 faces an awful choice: his expected payoff from war at this point is -3.26, which is worse than appeasement before escalation. On the other hand, capitulating now is worst: -5.16, so he prefers to fight. In other words, S_1's own actions have produced a situation where he has firmly committed to fight even though he would much rather have conceded at the outset. The reason he does not appease, of course, is that in expectation escalation is profitable. Since S_2 is expected to capitulate with about 31% probability, S_1's expected payoff from escalation is about 3.54, strictly better than appeasement. Therefore, the equilibrium outcome under uncertainty is war.

What would happen if these opponents had complete information about each other? As before, at the final node S_1 would fight only if $m \gtrapprox 1.94 \equiv \underline{m}$. If S_1 mobilizes at least $\underline{m}$, then war is certain if S_2 resists. Since S_2's payoff from fighting is $37.5/(m+5) - 2.5$ and from capitulating -0.5, she would fight if $m < 13.75 \equiv \overline{m}$ and would capitulate otherwise. Hence, if S_1 mobilizes $m < \underline{m}$, S_2 would resist and he will capitulate, earning a negative payoff, so appeasement is preferable. If he mobilizes $m \in [\underline{m}, \overline{m}]$, S_2 would resist but S_1 would fight anyway. The optimal war allocation for S_1 is at $m^* \approx 1.85$ but his payoff in this case is -2.44, so appeasement is preferable. Finally, if he mobilizes $m > \overline{m}$, he can guarantee S_2's capitulation, in which case his payoff would be $v_1 - \overline{m} = 5$, which is strictly better than appeasement.

[8] That is, v_i is uniformly distributed on $[0, \overline{v}_i]$, $M_1 = 0.10 \times \overline{v}_1$ and $p_1(0, M_1, M_2) = 0.5$, which implies $M_1 = M_2 = 2.5$. The costs are $c_1 = 12.5$, $a_1 = 0.5$, $c_2 = 2.5$, and $a_2 = 0.5$.

In the unique subgame-perfect equilibrium, S_1 mobilizes at $\overline{m} = 13.75$ and S_2 capitulates. Therefore, the equilibrium outcome with complete information is successful compellence by S_1.

What is especially striking about the result under complete information is that S_1 achieves compellence even though his best war fighting payoff (-2.44) is worse than appeasement (0). Why does this work? Because sinking the mobilization cost makes capitulation costlier than before: if S_2 resists, the new choice S_1 has is between a payoff of -16.25 from capitulating and a payoff of -10 from fighting at $\overline{m}$. Although war at such a costly mobilization level is much worse than the optimal war payoff, it is better than capitulating. Thus, S_1 has tied his hands by sinking the mobilization costs at the outset, and he will certainly fight if challenged now even though he would have appeased rather than fought even under optimal conditions at the outset. Because of S_1's rather high mobilization level, fighting becomes too painful for S_2 and she capitulates.

Contrast this with the results under asymmetric information, where S_1 allocates $\widehat{m} \approx 4.66$. First, this is less than what is required to get S_2 with valuation $v_2 = 15$ to capitulate ($m \geq 13.75$). Second, it is more than the maximum mobilization at which S_1 is willing to capitulate himself ($m \leq 0.36$). In other words, S_1's mobilization level is too high for him to backtrack once S_2's valuation is revealed, but it is too low to get S_2 to capitulate either. The outcome is war: S_1's actions have now created a situation where neither opponent is prepared to back down. This situation arises because of uncertainty and would not have occurred had S_1 known his opponent's valuation from the beginning.[9]

It is precisely this dynamic of simultaneous commitments that so worried Khrushchev during the tense days in 1962. As he wrote in his October 26 message to Kennedy warning the American President of the dangers of continuing escalation,

> If... you have not lost your self-control and sensibly conceive what this might lead to, then, Mr. President, we and you ought not now to pull on the ends of the rope in which you have tied the knot of war, because the more the two of us pull, the tighter that knot will be tied. And a moment may come when that knot will be tied so tight that even he who tied it will not have the strength to untie it, and then it will be necessary to cut that knot, and what that would mean is not for me to explain to you, because you yourself understand perfectly of what terrible forces our countries dispose.[10]

The immediate reaction to this conclusion would be to ask the original question once again, this time applying it to the final stage prior to the outbreak of war: after all information has been revealed, shouldn't the actors strike a bargain? There are three ways to approach this. First, one can argue that certain situations involve threats to use force if one oversteps some boundary or fails to comply with a particular demand, and as such may not be open to negotiations about distribution of benefits.

For example, in 1797, Venice recruited and armed peasants to deal with pro-French insurrections in her territories. However, when these ragtag forces began harassing the French rather than the rebels, Napoleon got incensed and ordered his *aide-de-camp* General Junot to deliver an ultimatum to Venice. After coercing the Collegio to meet on a Saturday despite

[9] One may ask what would happen if S_2 could counter S_1's mobilization with one of her own. I explore a model that permits this in Slantchev (2005) under one-sided uncertainty. The results are absolutely the same although some additional work is necessary to show that given S_1's initial high mobilization level, S_2 will not be willing to counter-mobilize at a level that is sufficiently high to get S_1 to capitulate in the endgame.

[10] The letter is reprinted in US Department of State (1996) as Document #65.

the religious holiday, Junot read Napoleon's letter in which Bonaparte denounced Venice and bluntly asked its government:

> Is it to be war, or peace? If you do not take immediate measures to disperse these militias, if you do not arrest and deliver up to me those responsible for the recent murders, war is declared.

To underscore the fact that these terms were non-negotiable, Junot threw the letter on the table and then walked out before anyone could say anything. In the event, Venice capitulated with a "cringing apology" (Norwich, 2003, 619-20).

This sort of behavior is not confined to the eighteenth century. Following US mobilization to eject Saddam Hussein from Kuwait, there were some last-minute attempts to compel Iraq to withdraw without a war. One of them was a proposed meeting Foreign Minister Tariq 'Aziz of Iraq and Secretary of State James A. Baker III. President Bush described the intent as follows:

> This offer is being made subject to the same conditions as my previous attempt: no negotiations, no compromises, no attempts at face-saving, and no rewards for aggression. What there will be if Iraq accepts this offer is simply and importantly an opportunity to resolve this crisis peacefully.[11]

While it is possible that the President was making this claim for strategic purposes, the events that followed demonstrated that in January, the US was in no mood to negotiate anything but the unconditional liberation of Kuwait and, the expanded aim, the destruction of Iraq's military capability. The decision to cross the 38th parallel in Korea, discussed in more detail in Chapter 6, was also about overstepping a limit set by the opponent. Hence, such a model can apply in certain situations but perhaps not in others.[12]

Second, one can argue along the lines in Chapter 3 that eleventh-hour negotiations may be impossible either because of risks of preemptive attack or because of inability to maintain combat readiness for too long. For example, since mobilization cannot be maintained indefinitely, there is a risk that if one fails to strike and has to disengage, the process of demobilization would leave one vulnerable to attack. A combination of mobilization pressure and fear of surprise attack was the main contributing factor to Israel's decision to strike Egypt preemptively in 1967 even against the vociferous opposition of the Americans (Oren, 2002).

The third, and perhaps best, option would be to resolve this theoretically by incorporating a richer bargaining framework into the model. The literature on bargaining breakdown under complete information suggests that this would indeed be possible.[13] Powell (2004) provides a general mechanism which guarantees that all equilibria in a large class of games will be inefficient and Powell (2006) applies its logic to other explanations of war. This mechanism relies on large rapid shifts of power between the players which create dynamic commitment problems when a previously weak actor no longer has incentives to fulfill the terms of his promises after he becomes the stronger one. As Leventoğlu and Slantchev (2007) note, however, despite its generality, this mechanism leaves a lot to be desired. In particular, the

[11] Statement of January 3, 1991. http://bushlibrary.tamu.edu/research/papers/1991/91010300.html. Accessed September 10, 2004. See Brune (1993, 105–06) for more on Bush's refusal to negotiate.

[12] For more on ultimata, see Lauren (1994).

[13] See, *inter alia*, Fearon (1995), Garfinkel and Skaperdas (2000), Slantchev (2003a), and Langlois and Langlois (2005) on the possibility of conflict without uncertainty in this framework.

crucial power shifts are exogenous to the models; that is, they are beyond the players' control. However, since such shifts can arise only from deliberate decisions (e.g., investing in armaments), it is more appropriate to assume that players can influence the rate of change. If that is so, one has to explain why actors still make decisions that risk producing power shifts that are large enough and rapid enough to guarantee bargaining breakdown.

The MTM's two-step logic seems especially apposite here: it is not hard to envision how the costliness of mobilization can prevent a player from purchasing sufficient armaments to offset the advantage of his opponent, thereby failing to decrease the size of the power shift and ending up in a situation where war becomes inevitable even in a rich bargaining framework and with complete information. The first step is to make the arming decisions under uncertainty with coercive and informative purposes, and the second step is to resolve the crisis through bargaining, which may turn out to be impossible to do peacefully. However, even though this seems plausible, I have yet to formalize the mechanism. Hence, I can only offer the two-step logic as a conjecture at this point.

Part III

Elements of Militarized Deterrence

5

Militarization and the Distribution of Power and Interests

Nothing can be a more amazing folly than for two great countries like India and China to go into a major conflict and war for the possession of a few mountain peaks, however beautiful the mountain peaks might be, or some area which is more or less uninhabited.
Jawaharlal Nehru, 1959

I now outline some elements of a theory of militarized deterrence that arise from embedding the military threat model of coercion into the wider context of a conflict of interest encounter. To do this, I develop a fuller theory of military coercion that allows S_2 to choose whether to challenge the status quo at the existing distribution of power. This extension brings the model into close correspondence with the informal one used by Huth (1988, 20-23, esp. Figure 2) and the formal one used by Fearon (1994b), and facilitates the comparison of its predictions with their hypotheses.

To extend the MTM to allow for a prior move by the challenger, assume that she can either choose to live with the status quo (in which case her payoff is zero, and the defender's payoff is v_1), or choose to make a public demand. If she does make a demand, the game continues as in the MTM with the defender's escalation choice. I assume that S_2's valuation is distributed uniformly on $[0, \overline{v}_2]$ and she is privately informed of it before she makes the first move while the distribution from which it is randomly drawn is common knowledge.

The MTM results from Chapter 3 are defined in terms of arbitrary beliefs about S_2's valuation; all S_1 knows when he escalates is that her valuation is uniformly distributed on the interval $[t, u]$. Since S_2's initial choice does not affect any variables other than (potentially) S_1's beliefs, it follows that the entire effect of crisis initiation would be reduced to influencing these beliefs in a way that would be most beneficial to the challenger; that is, initiation will be a form of signaling. Since there are only two actions available to S_2, in equilibrium her types would partition themselves into at most two sets: those who prefer to live with the status quo, and those who prefer to start a crisis. Furthermore, these sets will be represented by continuous intervals because it is not difficult to see that if in equilibrium some type prefers to initiate, then all higher ones would strictly prefer to do so as well. In other words, S_2's equilibrium behavior would necessarily define the interval $[t, u]$ with $t \geq 0$ being the type indifferent between the status quo and a crisis, and $u = \overline{v}_2$ being the highest valuation type. Of course, if S_2 chooses to live with the status quo regardless of valuation, it will be up to the analyst to specify reasonable off the path beliefs necessary to sustain this behavior in equilibrium. The bottom line is that in equilibrium only types $v_2 \geq t$ will initiate, which determines $[t, \overline{v}_2]$, which in turn determines the equilibrium of MTM that S_2 should expect, which itself rationalizes the choice of t in turn. The rest of the analysis reduces to identify-

ing possible values of t with the corresponding MTM equilibrium behavior they induce and showing that there is at most one such pair that can be supported in equilibrium. In other words, if the initial choice results in a unique MTM equilibrium, then the combination of an initial choice and MTM equilibrium is itself unique. The full analysis can be found in Appendix D.

5.1 Stability and Deterrence Failure

Extending the model to include the challenger's initial decision introduces a new concept of stability. Recall that in Chapter 4, we defined two closely related concepts that capture different aspects of the risk of war: *crisis stability*, which is the probability that a crisis will end in war, and *escalation stability*, which is the probability that a crisis will end in war provided that S_1 militarizes it. *Situational stability* refers to the probability that the status quo is disrupted by a crisis which then escalates to war. In other words, it is a measure of how prone a status quo distribution is to producing fighting.

DEFINITION 5.1 (Situational Stability). The unconditional probability that the status quo is challenged and the crisis escalates into a war. The higher the probability, the less stable the status quo.

Recall that crisis stability is an *interim* calculation because it presupposes an existing crisis, and escalation stability is an *ex post* calculation because it presupposes a militarized response. In contrast, situational stability is an *ex ante* calculation because it takes into account S_2's initial decision to challenge the status quo.

The other quantities of interest involve peace, which now can be had in one of three ways: if S_2 never challenges the status quo (preservation under an implied threat), if she does but S_1 appeases (redistribution), or if she does, S_1 escalates, and she capitulates (preservation under direct military threat). The first is a measure of the proneness of the status quo to crises, and answers the question "how likely is some distribution of benefits to be challenged?" It is the probability that S_2's valuation is too low to make initiating a crisis worthwhile. The second is a measure of the potential for the status quo to be peacefully revised in favor of the challenger. It is the probability that S_2 initiates a crisis and S_1 appeases her by voluntarily relinquishing the disputed good. The third is a measure of S_1's ability to protect the status quo with a vigorous militarized response to a challenge. It is the probability that S_2 initiates a crisis but is then forced to back down after S_1 responds by mobilizing his forces.

To relate the notions of stability developed in this book to those of deterrence that are in widespread use, we must disentangle two concepts that are frequently conflated in the literature—those of deterrence failure and war. When talking about the efficacy of threats, we need to be more specific and more careful in saying exactly what it is that we are measuring as success. The effect on avoiding war? On securing one's objectives? Both? Traditionally, deterrence success is identified with cases where a defender's threat restrained the potential revisionist from using force (Huth and Russett, 1984; Lebow and Stein, 1987). That is, peace is deemed equivalent to deterrence success. As Huth and Russett (1988) succinctly state it, "Failure [of deterrence] is defined as an attack on the protégé by regular military forces." As this model clearly demonstrates, Danilovic (2001, 103) gets it right when she disputes this equivalence. As she puts it, "Although war undoubtedly represents deterrence failure,

peaceful outcomes imply three possibilities in terms of perceived successes: a deterrer's perceived success if the challenger peacefully acquiesces..., the challenger's success if the deterrer acquiesces without fighting,..., or both might compromise." Even though the MTM does not allow for the compromise outcome, it does support the notion that crisis stability should not be conflated with deterrence success.

This now allows us to link more precisely our notions of stability to the traditional concepts of general and immediate deterrence (Morgan, 1977). *General* deterrence refers to the initial choice to challenge the status quo. As such it would appear to be equivalent to the probability of the challenger initiating a crisis. However, since classical deterrence theory conflates deterrence failure with the use of force, it may well correspond to situational stability instead. Avoidance of war does not tell us whether changes in the distribution of benefits have occurred. Matters can be settled peacefully if S_2 does not challenge the status quo at all (in which case S_1 retains his possession of the good), if she does press her demands and S_1 yields without resisting (in which case the good is transferred to S_2), or if she presses, S_1 resists, and she yields (in which case S_1 keeps the good but pays the price). In other words, the status quo can be peaceful either under (possibly costly) *preservation* or under *redistribution*. Although the original notion of deterrence failure is vague about these instances, it seems to me that its spirit is to suggest that cases of peacefulness under redistribution would fall well within its scope. On the other hand, preservation would be construed as deterrence success regardless of cost. As a result, it appears more natural to identify general deterrence with the probability of crisis initiation rather than situational stability.

DEFINITION 5.2. General deterrence failure is measured by the probability that the challenger initiates a crisis.

Immediate deterrence refers to the defender's ability to coerce the challenger to capitulate short of war once a crisis has begun, and as such corresponds directly to escalation stability. Note now that in all cases where S_1 fails to coerce S_2 his credible escalation leads to war. This means that even though in principle deterrence failure is not equivalent to the actual use of force, in equilibrium of the MTM, it is. When it comes to immediate deterrence failure, then, the original insight appears to have been entirely on the mark. The fact that we can derive an original axiomatic definition from equilibrium behavior should be encouraging about the usefulness of the MTM.

DEFINITION 5.3. Immediate deterrence failure is equivalent to escalation instability and is measured by the probability that an escalation ends in war. Because equilibrium escalation is credible, immediate deterrence in the MTM is the same as the traditional definition.

We should not conflate crisis instability with immediate deterrence failure. Crisis stability incorporates the probability of escalation and the conditional risk of war. This does answer the question of how likely a crisis is to end in fighting but this is not the same as asking how likely a threat is to lead to war. In other words, whereas immediate deterrence failure presupposes that a threat is made (and its failure to coerce S_2 leads to war), crisis instability includes the probability that S_1 makes the threat. This means that as S_1 becomes more likely to appease, crisis stability will increase as well. Obviously, this is a far cry from saying that immediate deterrence will improve.

This now leaves a curious gap: what type of failure is measured by the probability that

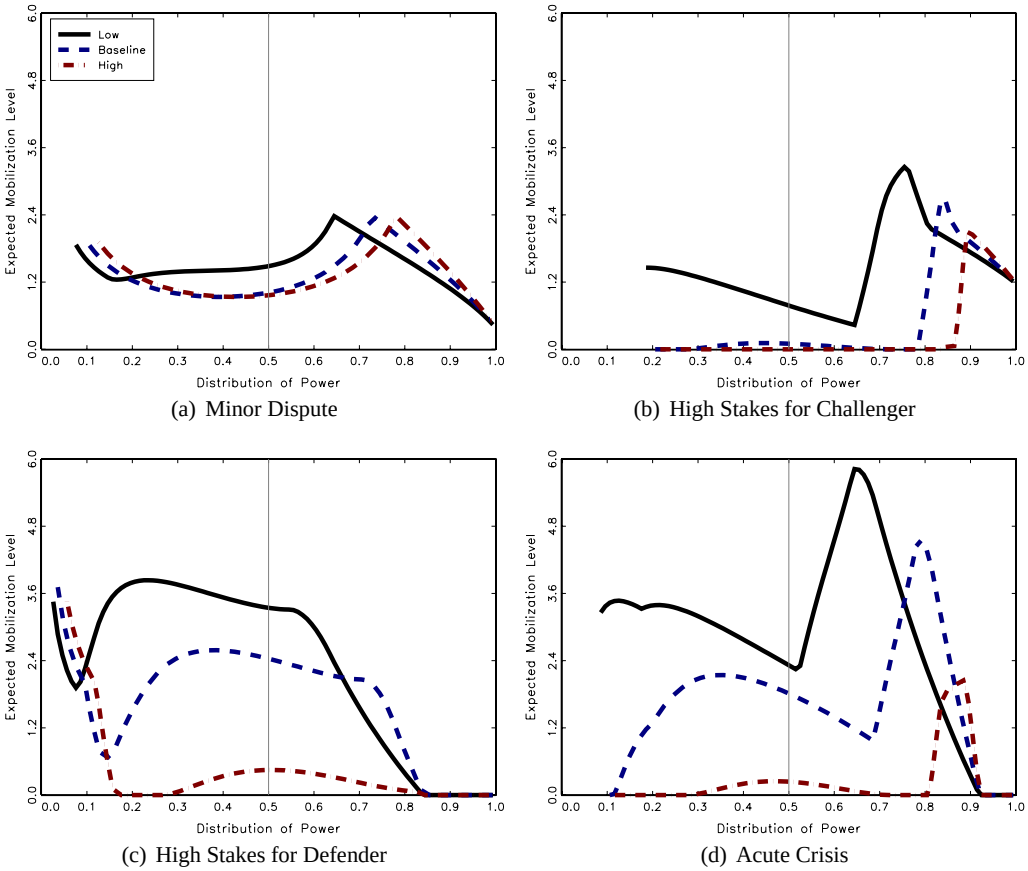
the defender will not meet a challenge with an escalatory threat? Since the status quo gets revised in favor of the challenger, this must be some sort of failure from the perspective of the defender. Appeasement must involve failure of general deterrence because the challenger has initiated a crisis. But since immediate deterrence is only defined conditional on the defender escalating (a threat cannot fail if it was never made), the interim step remains in limbo. What about the decision to actually make the immediate deterrent threat? That this decision is absolutely crucial cannot be doubted: the credibility of the general deterrent threat rests entirely upon it because in equilibrium all escalation that involves any possible resistance by S_2 is credible. This implies that it is never an issue whether S_1 will follow through on an escalatory threat, the only relevant uncertainty then from S_2's *ex ante* perspective is whether S_1 will be willing to make that threat at all. This suggests that whereas credibility of immediate deterrence is correctly considered an important component of general deterrence, the other essential ingredient, the probability of escalation, has been unduly neglected. Moreover, as one of the outcomes identified by Danilovic (2001), appeasement by the defender should be counted as deterrence failure despite being peaceful.

Rather than invent yet another category of deterrence failures, I will just stick to the concepts of stability developed in this book and be mindful of the deficiencies of our existing nomenclature. These will become especially important when we consider the selection models that attempt to evaluate the risk of war conditional on general deterrence failure. In effect, these models are estimating crisis stability, not immediate deterrence failure, even though they often mislabel the findings. Since the two are distinct concepts and the explanatory variables often relate differently to each, we shall need to do some careful re-evaluation of the findings of such studies when it comes to immediate deterrence.

5.2 System Militarization and Military Threats

Recall that in the MTM, the marginal effect of additional mobilization depends on the existing system militarization levels. We shall explore three scenarios: (i) baseline militarization: M_1 is 10% of $\overline{v}_1$; (ii) low militarization: M_1 is half the baseline (5% of $\overline{v}_1$); and (iii) high militarization: M_1 is double the baseline (20% of $\overline{v}_1$). For each of these scenarios, we shall change M_2 such that the distribution of power varies over the entire range.

As we know, the marginal effect of mobilization is higher in under-militarized systems (more bang for the buck). Figure 5.1 shows the expected mobilizations under each of the three militarization levels in four balance of interest scenarios. As the preceding discussion should have suggested, equilibrium behavior is heavily conditioned on existing military forces. On one hand, this may appear surprising: after all, these forces are summarized by the contest success function, which produces the same probabilities for equivalent distributions of power. Hence, one may think that the relative "balance of power" is the relevant metric. Indeed, this is what usually ends up as an explanatory variable in many statistical models. On the other hand, the MTM shows clearly that because the marginal effect of committing additional forces is decisive, one cannot ignore the overall militarization of the system. Bennett and Stam (1996) get it exactly right when they include (for a different reason) the absolute size of forces in addition to the customary balance of power variable in their statistical model.



(a) Minor Dispute (b) High Stakes for Challenger

(c) High Stakes for Defender (d) Acute Crisis

Figure 5.1 System Militarization and Mobilization Levels (low costs).

RESULT 5.1 *Crisis behavior does not depend simply on the relative pre-crisis power of the opponents but also on the absolute levels of their military forces.*

This result has some intriguing implications for the debate on whether arms races cause war. The usual story is that arms races contribute to the outbreak of war.[1] Wallace (1982) found that disputes preceded by arms races were more likely to escalate to war, although the strength of the relationship he found has been disputed (Altfeld, 1983). A perennial confounding issue is the unresolved and very difficult question of whether countries arm because they anticipate the conflict that eventually breaks out (in which case the arms races cannot be said to have contributed to war even though their incidence will be highly correlated with it) or whether military buildups create a dynamic which makes violent crisis outcomes more likely (Glaser, 2004). Morrow (1989c), for instance, argues that temporary advantages created by swings in the military superiority during the buildup may provide

[1] This is implicit in Richardson's (1960) formulation, but is characteristic of the literature in general, especially in its early stages.

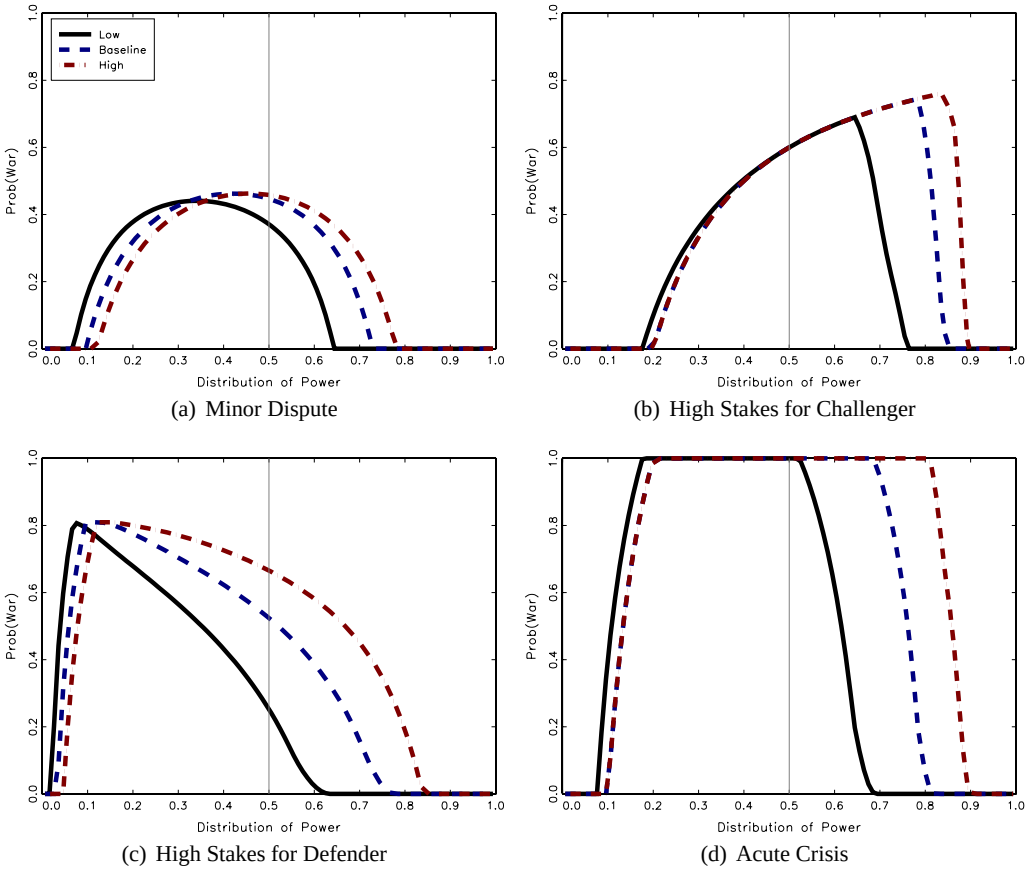
powerful incentives for war. Kydd (2000), on the other hand, argues that arms races may actually reduce the risk of war because they clarify the relative power of the opponents.

The findings here suggest that militarization affects how crises are resolved because it changes the relative importance of intra-crisis mobilization. In other words, the causes of a military buildup (which alters the status quo distribution of power) may have very little to do with the causes that lead to a dispute, crisis, or even war. Werner and Kugler (1996) make this argument as well in their criticism of traditional approaches. The nuance added by the MTM, however, appears quite important because here military buildups *affect stability indirectly anyway* because they change the marginal usefulness of the military instrument. In other words, although countries may engage in an arms race for reasons very different from the ones that eventually propel them into a dispute, the arms race affects how the crisis will be settled because of the context it creates for the use of military threats. It is worth emphasizing this point: *even if military buildups are not simple preparations for conflict, they should still be associated with crisis instability.* As Maoz (1990, 60) aptly put it, "the path from arms race to war goes through crises." The reason buildups may be causally related to the outbreak of war has to do with the effect they have on behavior during crises.

Figure 5.1 shows several intriguing findings. First, under-militarized systems will tend to exhibit more aggressive mobilizations under all but very skewed distributions of power. In other words, regardless of the balance of interests, defenders will use more cautious mobilizations when the crisis is between highly armed states. This should not come as a big surprise in the light of Claim 3.1: since mobilization is more effective against lightly-armed opponents, defenders will generally tend to find it more useful, either coercively or in preparation for war. Seeing this is a bit complicated in the aggregate plots because changing system militarization also shifts the cut-points, and thus the ranges of the various equilibrium types. However, it is not hard to understand it logically.

For example, in Figure 5.1(d), the expected mobilization levels sport two distinctive humps; in fact, the first of these occurs because in that range, the equilibrium is from Proposition B.2 and escalating types prepare for war. The expected mobilization starts at zero when DOP favors S_2 because war is quite unprofitable for S_1 and he appeases. However, as the DOP improves, more and more types of S_1 are able to benefit from war. Consequently, mobilization levels increase while DOP still favors S_2 and then decline again as the DOP shifts toward S_1 (because the improved pre-crisis DOP makes additional forces less needed). At the trough, mobilization hits a very low level before leaping dramatically upward. This surge occurs because the DOP has become favorable enough to permit coercion and some types begin taking advantage of it. This shift from war to coercion becomes evident when we look that crisis stability for the same set of parameters, as shown in Figure 5.2(d). When the DOP favors S_1 sufficiently, compellence becomes possible and further improvements enable S_1 to achieve it at even lower mobilization levels. Because of the relative efficiency of mobilization in under-militarized systems, these various strategies become available at lower DOP, hence the leftward shift of the graphs as system militarization decreases.

Second, the shift is not just to the left (coercion becoming more attractive at lower DOP in under-militarized systems), but in most cases also upward. That is, coercion involves higher mobilizations in these systems as well. This immediately implies that such crises should be more stable because S_2 will be much more likely to concede. Indeed, Figure 5.2 demonstrates that this is the case almost everywhere. The probability of war tends to be lower,



(a) Minor Dispute (b) High Stakes for Challenger

(c) High Stakes for Defender (d) Acute Crisis

Figure 5.2 System Militarization and Crisis Stability (low costs).

sometimes dramatically so, in under-militarized systems despite the seemingly aggressive mobilizations. This leads to a somewhat counter-intuitive conclusion:

RESULT 5.2 *Except at very skewed distributions of power, crises between heavily armed opponents will tend to involve less aggressive behavior but the risk of war will be greater than ex ante probability-equivalent crises between lightly armed states.*

The third finding is that even though mobilizations in under-militarized systems are lower when DOP disproportionately favors the defender, crises remain more stable. Although this is evident in all plots, Figure 5.1(b) and Figure 5.2(b) are perhaps most illustrative: when DOP becomes so high that S_1 can start compelling S_2 to capitulate with certainty, the mobilization levels begin dropping quite dramatically. Again, compellence becomes possible under somewhat less favorable DOP in under-militarized systems, so the decline begins sooner there. However, because equilibrium play involves increasing probabilities of S_2's capitulation, crisis stability improves and even exceeds that of militarized systems where S_1 still has to risk coercion given the same DOP. Hence, except when the distribution of power

is extremely unfavorable to the defender, crises will be more stable in under-militarized systems irrespective of the relative mobilization levels.

The last thing to note here is the inconsistent prediction when DOP heavily favors S_2: at these distributions, under-militarized systems will exhibit lower mobilization levels and higher risks of war. However, if we recall the reason for the left-ward shift in the plots, this discrepancy becomes intuitive. In this range of DOP, equilibrium escalation leads to certain war under Proposition B.2 because no types find even coercion profitable, let alone compellence. At extremely unfavorable DOP, no type would even escalate, which is why the probability of war will be zero. Because of the relative effectiveness of the military instrument in under-militarized systems, war becomes profitable at somewhat less favorable DOP than in more militarized ones. Consequently, higher types of S_1 begin taking advantage of it at lower DOP, and since escalation leads to war, crisis stability quickly begins to deteriorate. In other words, crises are less stable in under-militarized systems here because at these values of DOP, the set of types who escalate to certain war is larger. It is not until coercion becomes possible that the advantages of these systems translate into better prospects for peaceful crisis resolutions.

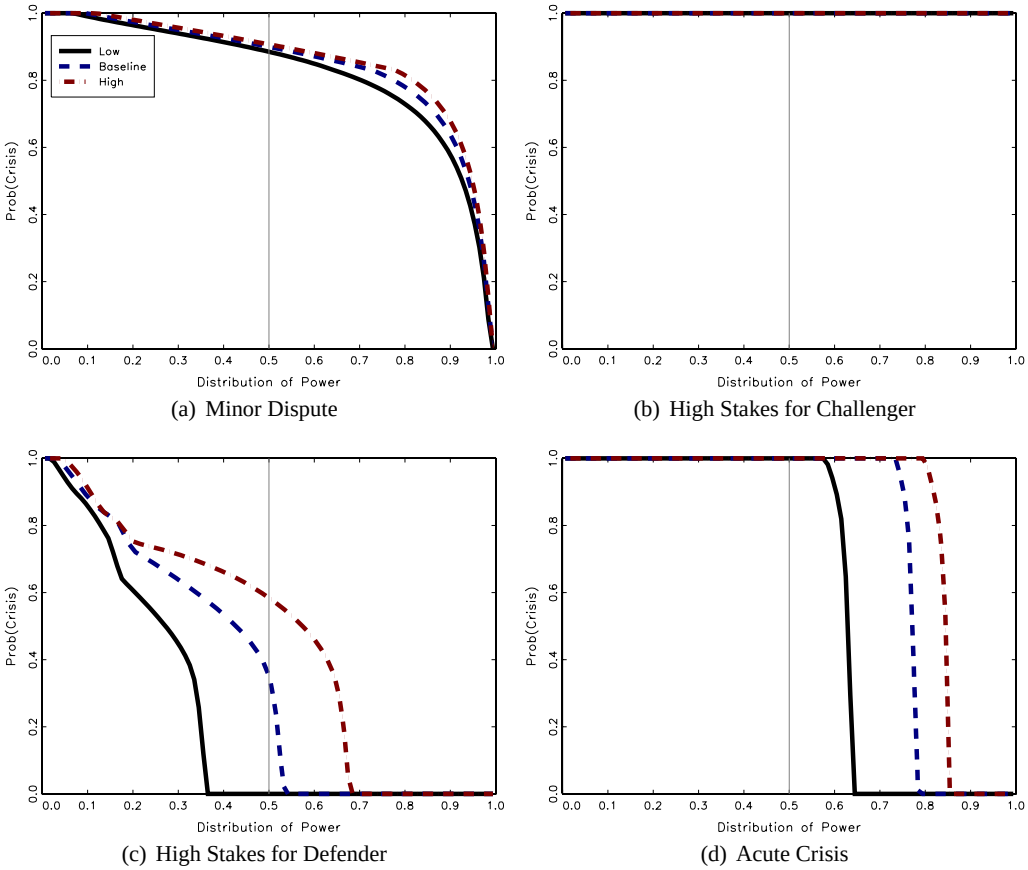
RESULT 5.3 *Crises in under-militarized systems will be more stable unless the distribution of power highly favors the challenger, in which case they will be less stable than crises in heavily militarized systems.*

Hence, the fact that the attractiveness of the military instrument is conditional on the existing distribution of power is a two-edged sword. On one hand, it will mean more stable crises in under-militarized systems when the DOP is not extremely favorable to the challenger because the coercive function increases the probability of her capitulation. On the other hand, it will also mean less stable crises in these systems when the DOP is extremely favorable to her because when coercion is not attractive but war is, more types will escalate to fight one. Somewhat perversely, the very features that make the military instrument more attractive for coercion (and are stabilizing) also make it more attractive for war preparation (and are destabilizing).

5.3 System Militarization and Deterrence

As in the previous section, I begin by looking at how existing levels of militarization affect the probabilities that a crisis occurs (S_2 issues a challenge) and that the status quo is disrupted by war (situation stability). To make comparisons with intra-crisis behavior, the simulations use the same parameters as in Figure 5.2. The crucial difference, of course, is that the balance of interests now describes the status quo rather than the crisis itself: the distribution of possible valuations of S_2 in the MTM that follows depends on the lowest-valuation type that would challenge the existing distribution of benefits. Note that since S_2's challenge can only improve S_1's information in our equilibrium, it does not change the type of dispute (because the underlying upper limit on her valuations remains the same).

Figure 5.3 shows the probability that S_2 challenges the status quo in each of our four balance of interests scenarios as a function of the existing distribution of power. The first thing to note is that S_2 is always at least as likely to start a crisis in a highly militarized system as she is in a less militarized one. That is, increasing the overall level of armaments



(a) Minor Dispute (b) High Stakes for Challenger

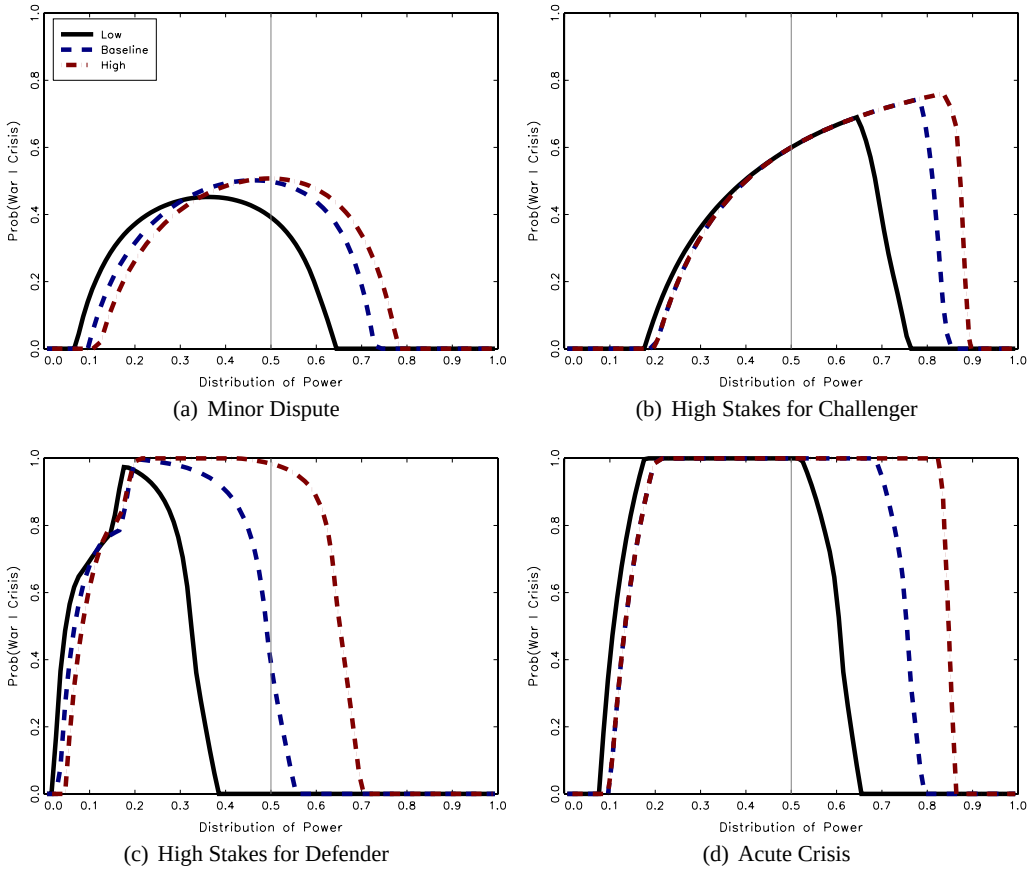
(c) High Stakes for Defender (d) Acute Crisis

Figure 5.3 System Militarization and General Deterrence Failure (low costs).

is unlikely to lead to a crisis-free coexistence between the opponents. In fact, it is much more likely to do just the opposite, especially when the stakes are high for the defender.

RESULT 5.4 *General deterrence is more likely to fail in highly militarized systems.*

Contrast now Figure 5.3(b) and Figure 5.3(c). In the first scenario, the defender's interest is peripheral and the challenger who values the issue highly is certain to initiate a crisis regardless of the level of militarization. It is not difficult to see why: she expects to do relatively well in the ensuing confrontation and cannot be deterred by the feeble threat the defender can make given his disinterest in the issue at stake. When the balance of interests is reversed, as in the second scenario, the level of system militarization has a dramatic impact. Take, for example, DOP at about 45%: S_2 will stay with the status quo for sure if the system is under-militarized; will challenge with probability around 50% if it is militarized at the baseline level; and will challenge with almost 70% if is over-militarized. This appears counter-intuitive: from Result 5.2, we know that crises will tend to be less stable in more militarized systems. In fact, crisis stability is even worse once S_2's initial choice is taken



(a) Minor Dispute (b) High Stakes for Challenger

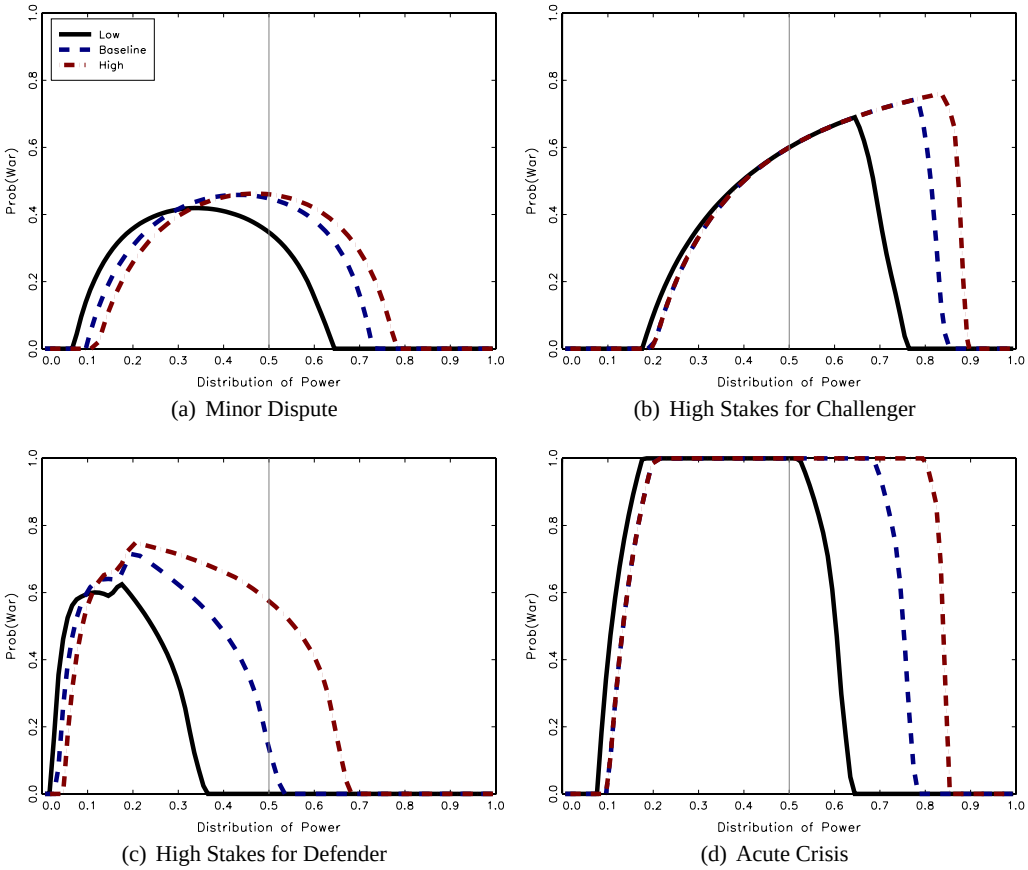
(c) High Stakes for Defender (d) Acute Crisis

Figure 5.4 System Militarization and Crisis Stability Conditional on S_2's Challenge (low costs).

into account, as Figure 5.4 shows. To see that, compare this figure to the plots in Figure 5.2, especially Figure 5.2(c) where the difference is most striking because the probability of a challenge leads the the most significant revision of beliefs, as seen in Figure 5.3(c).[2]

That the probability of a challenge increases with system militarization appears quite strange because the probability of war given a crisis is so much higher, which should make S_2 less likely to initiate! In fact, that is precisely what would happen in any traditional model where S_1's compellent threat does not affect S_2's payoffs directly like it does in the MTM. To understand the seemingly odd behavior, observe that war occurs in equilibrium in the MTM only when S_2 stands firm at her final decision node. This means that all else equal, lower mobilization by S_1 would make her more likely to resist even if doing so is sure to lead to war because her expected war payoff will be higher. From Result 5.2, we

[2] The plots in Figure 5.2(b) and Figure 5.4(b) are, of course, identical because S_2 challenges regardless of her valuation, as seen in Figure 5.3(b), and hence S_1's posterior belief in the full deterrence model and the MTM will be the same.



(a) Minor Dispute (b) High Stakes for Challenger

(c) High Stakes for Defender (d) Acute Crisis

Figure 5.5 System Militarization and Situation Stability (low costs).

also know that S_1's mobilization will be less aggressive in more militarized systems, which means that S_2's expected payoff from war will be higher. Since the probability that S_1 would escalate is never higher (and usually much lower) in more militarized systems, all of this means that S_2's *expected payoff from a challenge actually increases with militarization.* This then explains why she is more likely to initiate a crisis. Observe that this prediction exactly reverses the seemingly intuitive expectations from non-MTM models. Here, *the probability of a challenge increases with system militarization despite the fact that crises will tend to be less stable.*

This "double whammy" with respect to stability is most evident when the defender's interests at stake are high (otherwise the effect a challenge would have on his beliefs will be negligible because almost every type of S_2 initiates it). Putting everything together, as done in Figure 5.5, leads to the conclusion that situational stability will generally be lower in more militarized systems, and that the magnitude of the effect will increase as the balance of power increasingly favors the defender.

RESULT 5.5 *The risk of war is generally higher in more militarized systems (situational stability is lower) except when the distribution of power seriously favors the challenger.*

It is worth emphasizing that this result would be counter-intuitive for any theory that does not account for the function of the military instrument. The common recipe for deterrence calls for a credible threat by the defender to fight, and the increased risk of war is supposed to deter the challenger from initiating a crisis in the first place. What this result shows is that matters are not that simple: here, the probability of a challenge increases even though the defender's threat is credible and the resulting risk of war is higher as well. What the military deterrence model shows, then, is that what matters is not only *whether* the defender will fight, but also *how* he intends to do it. If he will not mobilize enough forces to make the challenger's expected payoff from war really small, then a commitment to fight simply leads to war. The defender's threat is not capable of exercising a deterrent effect on the challenger. Of course, the distinction between a *capable* and *credible* threat is not new. What generally has gone unrecognized is the fact that the defender may have incentives that produce less capable threats, *especially when he thinks that war is very likely*. As we have seen before, the defender cannot get coercion on the cheap: to achieve compellence, he has to mobilize at least as much as he would for a real war. Hence, the more convinced he is that war will follow, the less likely is he to go to the extra expense of convincing S_2 of his resolve. This, in turn, makes S_2 more likely to challenge him in the first place, closing the vicious circle. This rationale provides yet another resolution to the "para bellum paradox" that arms races that are designed to promote peace may actually contribute to the outbreak of war (Maoz, 1990, 32-64).

Since military buildups essentially correspond to system militarization, this result can be tested empirically. Sample (1998) finds that even when other factors (e.g., type of issue being disputed, history of disputes, relative power capabilities, and so on) are taken into account, crises in more heavily militarized dyads are more likely to escalate. This supports earlier, and disputed, findings that military buildups are strongly correlated with the outbreak of war (Wallace, 1982). In a more nuanced study, Sample (2002) finds that crisis instability in militarized dyads is higher when two major powers confront each other or two minor powers do, but that there is no effect when a major power confronts a minor power. This can be taken as evidence that supports the second part of the above result: militarization will not be destabilizing if the distribution of power seriously favors one of the actors.

Now that we have established how high militarization undermines both S_1's deterrence posture and crisis stability, we shall carefully look at the relationship between the distribution of power and situation stability. All simulations that follow use the baseline system militarization.[3]

[3] This is because system militarization affects crisis stability in a consistent manner: the dynamics as a function of DOP remain roughly similar when we vary militarization—all that changes are the precise values of DOP where the peaks and troughs occur. Consequently, we shall use the baseline system militarization for the rest of the simulations to focus on the other variables of interest.

5.4 The Distribution of Power and Interests

The distribution of power (DOP) is among the most widely used concepts in international relations regardless of the school of thought.[4] Perhaps not surprisingly, there is a lot of controversy over precisely how it should affect the likelihood of war. Without going into an exhaustive literature review, it is fairly easy to identify at least four mutually contradictory hypothesized relationships between the distribution of power and crisis stability:

1 *Balance of Power:* the probability of war is lowest when power is approximately evenly distributed among opponents.[5]
2 *Preponderance of Power:* the probability of war is lowest when the distribution of power disproportionately favors one of the actors.[6]
3 *Power and Status Quo Benefits:* the probability of war is a function of the disparity between the existing distribution of benefits and the distribution of power. The probability of war is lowest when the benefits actors can obtain from fighting are relatively close to the benefits from the status quo. Conversely, as the disparity grows, so does the risk that an actor will resort to arms.[7]
4 *No Direct Influence of Power:* whereas the distribution of power affects the terms of a negotiated settlement, it will generally have no effect on the likelihood of war.[8]

Before I continue, two caveats are in order. First, I should note that I am considering these claims in a context limited to two actors only. While the third and fourth hypotheses come from such models, the first two (especially the balance-of-power) are usually stated for a world with more than two actors, and their focus is on alliance patterns (Waltz, 1979), and the probability of war among coalitions (Kim, 1989). Second, I should point out that it is not impossible to reconcile some of these relationships, as for example Wagner (1994) does when he argues that each may arise logically from a start with different sets of premises. A model with multiple equilibria, for instance, would produce different conclusions seemingly from the same set of assumptions unless one explicitly includes expectations that support equilibrium selection in that set.

Since a sufficient enlargement of the set of assumptions would either unify the theories or expose their fundamental contradictions, it may be useful to consider the empirical record for a preliminary prediction at how successful this approach may be. Unfortunately, efforts to resolve the supposed theoretical impasse have thus far yielded what we call "mixed results," i.e., contradictory findings that can be used to support or reject none of the hypothesized relationships. We have studies that support the balance-of-power hypothesis;[9] others that

[4] The classical realist tradition stretching back to Thucydides (1996) through Carr (1939), Morgenthau (1948), and Gilpin (1981) is most explicit, although Waltz's (1979) neorealism and the more recent defensive (Van Evera, 1999) and offensive (Mearsheimer, 2001) strands of realism also make very extensive use of the concept. Its overwhelming popularity also accounts for schools explicitly in opposition to it, usually with modest claim that the distribution of power is not the sole (or even the primary) explanatory variable (e.g., liberalism and neoliberalism).

[5] Morgenthau (1948); Claude (1962); Wright (1965); Kissinger (1979); Mearsheimer (1990).

[6] Organski (1968); Organski and Kugler (1980); Blainey (1988).

[7] Powell (1996b). Bueno de Mesquita and Lalman (1992, 188-190) argue, to the contrary, that dissatisfaction with the status quo is unrelated to the risk of war.

[8] Wittman (1979); Fearon (1992).

[9] Ferris (1973); Siverson and Sullivan (1983); Siverson and Tennefoss (1984).

support the preponderance-of-power hypothesis;[10] and yet others that find no consistent relationship one way or another,[11] even on the dyadic level (Kim, 1989), which may be taken as supporting the no-influence hypothesis. The findings are so confusing that Levy (1989a) has despaired about ever uncovering the effect of power in dyadic-level relationships.

The usual remedy is to suggest, *inter alia*, omitted variables like alliance capabilities (Kim, 1989), considering systemic or regional, as opposed to dyadic, levels of analysis,[12] or changing the focus on type of conflict.[13] Signorino (1999) shows that statistical analyses that use improper estimation techniques that ignore the fundamental strategic nature of the data-generating process can produce seriously misleading results. This implies that the instability of findings may be due to their sensitivity to the model specification. Until we acquire statistical methods capable of dealing with complex strategic models, we have to strive to resolve theoretically as many issues as possible.

It is mildly disturbing if internally consistent theories should yield diametrically opposed hypotheses from the same assumptions. In keeping with the general thrust of this book, we should at least make an effort to identify assumptions in these theories that are responsible for these results. Should these assumptions prove untenable because they fail to capture essential features of the interaction (and hence are likely to be distorting), the models that employ them would be seriously undermined on theoretical grounds. This is the approach I propose to use here: in the absence of compelling empirical findings and statistical models that can adequately capture the complex strategic interaction implied by the MTM, we shall look at the controversy from a primarily theoretical viewpoint.

Looking back to Figure 5.5, it would appear that the model's prediction about situation stability hews to the *preponderance of power* school of thought: the risk of war is higher when the distribution of power becomes more even. However, a closer look casts significant doubt on that conclusion because situation instability can peak just about everywhere (looking at the baseline militarization case):

- when there is approximate power parity, as in Figure 5.5(a);
- when there is serious asymmetry in favor of the defender: S_1's forces are about 80% of the dyadic total, as in Figure 5.5(b);
- when there is serious asymmetry in favor of the challenger: S_2's forces are about 80% of the dyadic total, as in Figure 5.5(c);
- for most values of DOP except at the two extremes, as in Figure 5.5(d).

Thus, on one hand it is striking that DOP tends to affect crisis stability in roughly similar way: probability of war is lowest when either the defender or the challenger enjoy an extremely pronounced advantage. On the other hand, the differences for values of DOP outside these extremes are so pronounced that they render the generalization almost meaningless. Take, for example, DOP set at 70% and observe that war is certain when the conflict of interests is acute, but its likelihood declines to about 70% if the balance of stakes favors the challenger, then further down to about 20% if both have peripheral interest, and then drops

[10] Garnham (1976); Organski and Kugler (1980); Moul (1988); Bueno de Mesquita and Lalman (1992); Geller (1993).

[11] Bueno de Mesquita (1981); Maoz (1983); Karsten et al. (1984).

[12] Singer et al. (1972); Gochman (1990); Lemke and Werner (1996).

[13] Houweling and Siccama (1988); Kim and Morrow (1992).

to zero when the balance of stakes favors the defender. In other words, just how destabilizing an asymmetrical distribution of power is depends on the balance of interests, which echoes the findings of the bargaining model of war.

RESULT 5.6 *The relationship between the distribution of power and the risk of war depends crucially on the distribution of interests.*

This may seem obvious at this point but it is surprising how many of the hypotheses tend to assume this away by postulating a serious conflict of interest. Even the bargaining model, which comes closest in its claims, is actually subtly different.

Observe first that what matters in the bargaining model is how closely the distribution of benefits mirrors the distribution of power (which determines the size of the discrepancy between life with the status quo and attempting to secure a better deal by force). In the MTM, the status quo is constant across the four scenarios, what does change is the value the actors attach to possessing it. In a minor dispute, S_1 has the good but neither he nor S_2 cares all that much about it. In contrast, in an acute crisis, both value it highly, making S_2 extremely dissatisfied with the status quo and S_1 quite happy. Our assumption that S_1 possesses the benefit essentially renders him the satisfied actor and S_2 the dissatisfied one across all scenarios. These labels are equivalent to Powell's (1999) where the distinction is between preferring to live with the status quo and using force. To see that, observe that in the MTM, $w_1(v_1) < v_1$ for all v_1—that is, S_1 would prefer to stick with the status quo rather than fight regardless of valuation; and $w_2(v_2) > v_2$ for high-valuation types—that is, S_2 sometimes prefers to fight rather than live with the status quo. In other words, S_1 is the satisfied state and S_2 is the potentially dissatisfied challenger. Thus, the MTM imposes by assumption that only one actor can be potentially dissatisfied. That the bargaining model reaches the same conclusion should give us some confidence that this assumption is not distorting (Powell, 1996a).

Turning now to the conflict of interest, note that in the bargaining model one actor becoming more satisfied with the status quo cannot make the other actor more satisfied and will usually render him less so. That is because the actors are assumed to have a zero-sum conflict of interest over the distribution of the benefit, so more for one automatically means less for the other, and their utility is non-decreasing in the amount they possess. Since valuations in the MTM are independent, it is quite possible for one actor to care much more intensely about the issue without affecting the preferences of the other. For example, the defender may have a relatively high valuation but the conflict of interest can be mild if the challenger does not care that much about the issue or acute if she does care a lot. Whereas satisfaction in the bargaining model can approximate this type of distinction, it cannot capture it precisely.

The real bite of the nuance comes when we compare a minor dispute to a high-stakes for defender only crisis. It is not at all clear what the bargaining model would have to say about this for S_1 is satisfied in either case and S_2's valuation is the same, so her dissatisfaction is constant. As we have seen from Figures 5.5(a) and 5.5(c), however, situational stability is very different: the MTM would predict that the likelihood of war is highest at parity in the former case and at severe imbalance in favor of the challenger in the latter; furthermore, once the DOP exceeds 50%, the probability of war drops to zero when the stakes are high for the defender only but remains positive (but declining) in a minor dispute all the way up to DOP exceeding 70%. In other words, there is a real difference here even though the

distribution of benefits has ostensibly remained the same. These results should be seen as a plea for taking intensity of interests a lot more seriously in research than we have generally done.[14]

RESULT 5.7 *Situational instability peaks when power is relatively evenly distributed if interests are approximately balanced, or when power disproportionately favors the actor with less intense interests at stake.*

This result contradicts the *no direct influence* school and subsumes the arguments of the other three schools with the appropriate qualifications. It links the distribution of interests to instability, as the bargaining theory does but draws attention to the important distinction between the status quo distribution of benefits and the intensity of preferences with respect to these benefits. In a way, the fact that the probability of war is maximized when the DOP favors the actor who cares *less* about the issue runs counter to our intuition but is readily explicable in the context of the model: these are situations where an actor (who is not keenly interested in the issue) is unwilling to expend the resources necessary to coerce his opponent (who is very interested, which is why he is difficult to coerce) but the advantage in the distribution of power is so large that he is willing to risk war anyway.

The results are in agreement with the *preponderance of power* school that the most dangerous crises can occur when power is evenly distributed but qualifies this with the observation that this only holds when interests at stake are approximately balanced as well. The MTM also shows that instability when the conflict of interest is acute is much more serious than in a minor dispute, which further sharpens the precision of the relationship and implies that not all power transitions must be dangerous. Finally, the MTM is in agreement with the *balance of power* school that blames power asymmetries for instability, although this has to be qualified with the observation that the hypothesized relationship only holds when there is a serious discrepancy in the balance of interests *and* when this asymmetry is exactly contrary to the distribution of power. Hence, the MTM predicts that power imbalances are most dangerous either when a very powerful challenger confronts a weak defender over an issue about which the defender cares deeply but the challenger does not (the expansion of the Korean War in 1950); or when a weak challenger confronts a powerful defender over an issue about which the defender does not care much but the challenger cares a lot (Sino-Indian War of 1962). I shall discuss the Chinese intervention in Korea at some length in the next chapter because there are several competing explanations that I would like to address. Here, I shall limit myself to a brief example of the crisis of 1962 between China and India.

5.5 The Sino-Indian War of 1962

As the quote by Nehru in the epigraph for this chapter suggests, nobody really believed that the People's Republic of China (PRC) and India would go to war over an obscure border dispute in remote territories in the Karakoram and the Himalayas. Almost to the day they attacked, the Chinese maintained that the disagreement about the delineation can be resolved by negotiations. India's entire policy of refusing to negotiate was predicated on the

[14] This plea is not new. In fact, critics of rational deterrence theory have commonly faulted it for its exclusive focus on power and relative neglect of motivations (George and Smoke, 1974; Maoz, 1983; Karsten et al., 1984; Levy, 1988). I shall have a lot more to say about this in Chapter 7.

assumption that the Chinese would not use their superior military power to settle the dispute by force. I will not trace the origins of this complex territorial dispute. Suffice to say that upon independence India inherited from the British a northern boundary that was murky and not well defined. The two main areas over which the war with China would be fought were the approximately 30,000 square kilometer Aksai Chin in the west and the 90,000 square kilometer North-East Frontier Agency (NEFA, now the state of Arunachal-Pradesh).[15]

The value of the territories themselves as land was minimal. Maxwell (1970, 26) describes Aksai Chin ("desert of white stones") as "high and desolate plateau, 17,000 feet above sea level, where nothing grows and no one lives," and whose sole potential importance lay in the trade route that was only accessible during the summer months.[16] Most of the area is to the north of the Karakoram mountain range, which means that it is easily accessible from China and well-nigh impossible to defend from India because access requires negotiating inhospitable passes at very high altitudes. As part of the Ladakh kingdom, Aksai Chin was first conquered by the Dogras and annexed to Kashmir, which itself was absorbed into British India in 1904 through a treaty with Tibet. China, which had never recognized Tibetan sovereignty, denied the legitimacy of any agreements concluded by the Tibetans to which the Chinese were not officially signatories.

NEFA was more populated than Aksai Chin, mostly by various tribes with cultural and ethnic connections to India, Tibet, and Burma. India considered the NEFA area its own as part of the inherited agreement Britain had negotiated with Tibet at the Simla Convention in 1914. As with the other Tibetan agreements, China did not recognize this claim: the thirty or so years of Tibet's *de facto* independence from China were simply a temporary lapse of central authority between the 1911 collapse of the Qing Dynasty and the 1949 reassertion of central power under Mao's communists. Being under Chinese suzerainty, Tibet had no right to conclude any agreements on its own. Therefore, the McMahon Line claimed by India as its boundary, was illegitimate.[17]

The strategic value of the lands was another matter entirely. After China reconquered Tibet in 1950, the only practical way of connecting Tibet and the nearby province of Xinjiang necessitated the construction of a road through Aksai Chin (the alternative was to build it in the even more forbidding Takla Makan desert). Since the People's Republic of China (PRC) did not recognize this territory as part of India, it began construction of the highway in March 1956 without informing anyone. So remote is this area that the Indian government would not learn about the construction until September 1957.[18] Although according to their subsequent claims the building of this unauthorized road constituted an intrusion into Indian territory,

[15] There were also about 20,000 square kilometers of disputed territory on either side of the Karakoram watershed and the passes through that range; the so-called middle sector.

[16] Unless otherwise indicated, all page references are to Maxwell's (1970) seminal study of the war. Although some claim it presents a pro-Chinese gloss on the events, it is hard to find a more objective account of the dispute and the war itself.

[17] Gurtov and Hwang (1980, 115–16). The legality of the dispute is a complex and contentious matter. Lamb (1964) presents a well-balanced account that acknowledges the problems with the Chinese claims without accepting the official Indian line that the PRC was an aggressor. Hoffmann (1990, 9–30) also discusses the border ambiguity but also emphasizes the psychology behind India's approach to their delineation.

[18] Maxwell (1970, 87); Hoffmann (1990, 35–36). Mullik (1971, 196–201) claims that the Intelligence Bureau had been aware of Chinese activity in the Aksai Chin since 1951 but that the army had not considered it a threat it could deal with.

the government did not raise the issue with the Chinese at the time. India only made a formal protest in October 1958 when official Chinese maps suddenly showed the whole of Aksai Chin, previously marked as indeterminate, as being in China.[19] The PRC rejected India's claim but Prime Minister Zhou Enlai suggested that the two countries should negotiate the demarcation of the borders, most likely on the basis of the present status quo.[20]

The status quo in 1958 when the dispute began was straightforward: the PRC had established itself in Aksai Chin, and India had moved into the NEFA. Nehru, however, articulated what would become India's line throughout the crisis: "There can be no question of these large parts of India being anything but India and there is no dispute about them."[21] Zhou pointed out that "border disputes do exist between China and India." Recognizing India's interest in NEFA, Zhou further replied that the PRC's interest in Aksai Chin was firm but that in NEFA, China would be prepared to "take a more or less realistic attitude towards the McMahon Line."[22] In other words, the Chinese, while denying the legitimacy of any agreements between the British and the Tibetans, were proposing to recognize India's claim to NEFA in exchange for a reciprocal recognition of their claim in Aksai Chin.

This pragmatic approach to the issue was realistic because it reflected the situation on the ground. Furthermore, Sino-Indian relations were very warm and friendly with Nehru's *Hindi-Chini bhai-bhai* policy in full swing. It was an opportune time to resolve a potentially serious issue while nobody in India cared about the region that would go to China. Maxwell (1970, 90) summarizes it as well as anyone:

> The two governments were on the best of terms, each country had filled out into the no-man's-land of importance to itself, and all that was needed was an agreement to give binding diplomatic expression to what by all appearances was a mutually satisfactory *status quo*. If both sides were in fact satisfied there would be no Sino-Indian boundary problem at all; if on the other hand both—or either— stood by map claims to territory occupied by the other, the problem would be insoluble.

At the end of 1958, it would appear that the PRC had low interest in NEFA and moderately strong interest in Aksai Chin. India, had a much stronger interest in the NEFA and a correspondingly weaker one in Aksai Chin. However, it was at this point that Nehru chose to interpret India's claims as absolute and nonnegotiable: as inviolable part of India no territory in either sector could conceivably be admitted to being in dispute, let alone belonging to China.

This approach had to confront an unpleasant reality: China was ensconced in Aksai Chin and militarily much more powerful. If Nehru were to enforce India's claims, he had to find a way to dislodge an opponent of superior strength without provoking war. Having decided

[19] Informal Note from the Foreign Secretary to the Chinese Ambassador, October 18, 1958, Ministry of External Affairs, Government of India, *Notes, Memoranda, and Letters Exchanged and Agreements Signed Between the Governments of India and China: White Paper* (New Delhi, 1959–63), 1: 26–27. Thereafter cited as *White Paper*. Note given by the Ministry of External Affairs to the Counselor of China in India, August 21, 1958, *White Paper,* 1: 46.

[20] Memorandum from the Foreign Office of China to the Counselor of India, November 3, 1958, *White Paper,* 1: 47.

[21] Letter from Nehru to Zhou, December 14, 1958, *White Paper,* 1: 51. See also Maxwell (1970, 97) and Hoffmann (1990, 36–37).

[22] Letter from Zhou to Nehru, January 23, 1959, *White Paper,* 1: 53; Gurtov and Hwang (1980, 116), Maxwell (1970, 98), Hoffmann (1990, 38–39).

to secure all the territories under dispute, Nehru at first kept his exchanges with Zhou secret from Parliament. However, the Lhasa Rebelion that erupted in Tibet in March 1959 soured the relations between the two countries and compelled China to take a closer to look at Tibetan affairs. While the Dalai Lama (who had been given political asylum in India) was agitating publicly for resistance to the Chinese and receiving much sympathy and moral support in India, the Chinese were becoming concerned that the rebels were being supported with supplies coming from India. The Chinese attempt to reassert control of the region to prevent that caused an increase in clashes with the Indians who were trying to enforce their rival claims.[23] Public criticism of Nehru's friendly policy with China were met with pleas from the PRC for India to remain non-hostile while China was reeling under pressure in the aftermath of the second Taiwan Straits crisis. Nehru was at a crossroads already when the Longju incident in August—a clash at an Indian outpost in NEFA that ended with the Chinese taking an Indian prisoner—became public knowledge in India and inflamed anti-Chinese opinion.[24]

Three days after the Longju incident, information about it appeared in the newspapers along with revelations about the highway in Aksai Chin. Nehru had to disclose to Parliament his private correspondence with Zhou regarding the border. The chickens from the unyielding policy had come home to roost: the Opposition, supported by widespread aversion to the Chinese, now clamored for a more aggressive action to secure India's borders. Initially, Nehru attempted to draw a distinction between NEFA and Aksai Chin: while maintaining that PRC's claims are "totally and manifestly unacceptable," he also hinted that the situation in Aksai Chin is much more vague.[25] Hawks in Parliament demanded that India bomb the Chinese road and accused Nehru of appeasement. The nuance soon vanished as Nehru was forced to take a firm public stand.

By the fall of 1959, the eminently solvable question of "possession of a few mountain peaks" had mutated into a matter of national honor and prestige. The insignificant peaks and patches of uninhabited territory had become "the crown of India" and even part of her "culture, blood and veins." Rather than offering a workable resolution to the dispute, China was acting from "the pride and arrogance of might" in advancing preposterous demands that almost no Indian would ever agree to.[26] It was in this way that India's valuation of the territory went from moderate to extremely high: once publicly framed as a matter of preserving the integrity and honor of India—which the government had to do for political expediency—the problem was bound to become intractable. On March 14, 1960, the Supreme Court of India made it illegal for the government to cede or acquire territory or modify boundaries without an amendment to the constitution. Although this particular ruling was prompted by a minor cession to Pakistan in 1958, it solidified India's position on the border dispute with China.[27]

[23] Gurtov and Hwang (1980, 117); Hoffmann (1990, 61).

[24] Maxwell (1970, 107–9).

[25] Maxwell (1970, 116–19); Hoffmann (1990, 66–68).

[26] Maxwell (1970, 121); Hoffmann (1990, 55).

[27] It is important to note that while public pressure did make it nearly impossible for Nehru to change course once the crisis erupted in full force, it was not responsible for his initial unyielding policy. The white papers would not be made available to Parliament until September 7, 1959, at which point Nehru had been refusing to negotiate with the Chinese for about a year.

Despite the bluster, Nehru must have realized that India could not stand up to China militarily. It is telling that he adopted the line of the weak confronting the strong: he accused China of bullying India even while pretending to be its friend. Echoing the vacuous *might does not make right* argument, he insisted that "Natural friendship does not exist if you are weak and if you are looked down upon as a weak country. Friendship cannot exist between the weak and the strong, between a country that is trying to bully and the other who accepts to be bullied… China was not fulfilling that prescription of friendship, but was on the contrary using the boundary question to assert superiority, even perhaps dominance, over India".[28] It is difficult to resist drawing a parallel with the famously unsuccessful remonstrance of the weak Melians to the powerful Athenians. While effectively siding with the Athenian position that "right, as the world goes, is only in question between equals in power, while the strong do what they can and the weak suffer what they must," Nehru contrived to argue that since India and China were friends and friendship was only possible among equals, India must be as powerful as China and China should not treat it as a weakling.[29] That was a *non sequitur*: either India was as powerful as China, in which case China could not be a bully, or India was not, in which case they could not be friends. But if they were not friends, then it would be futile to appeal to friendship to prevent the Chinese from exercising their power.

But the Chinese were not making any threats, not yet anyway.[30] Instead, they had reasonably concluded that given the hostile climate of public opinion in India, getting Nehru to agree to anything that would smack of concession would be well-nigh impossible. They proposed that the two countries maintain the status quo, by which they meant the situation on the ground: *de facto* PRC control of Aksai Chin and Indian control of NEFA.[31] At this point, the best thing for the PRC would be to keep the status quo while ensuring that nothing that could be construed as provocation would further aggravate the situation. Later, when people had forgotten about the territories, the governments could settle the issues quietly. It was quite obvious to them that India could not afford to settle the question through military means. China therefore attempted to defuze the situation by proposing a 20-kilometer withdrawal from the McMahon Line in the east and the line of actual control in the west. This would leave the strategic road in Chinese hands but will minimize contact between the border patrols.

To much public acclaim, Nehru rebuffed this offer and essentially refused to negotiate with the Chinese until they vacated all of the disputed territories. As he put it, "an agreement about the observance of the *status quo* would… be meaningless as the facts concerning the *status quo* are themselves disputed." Nehru reiterated his demand that the PRC vacate all of Aksai Chin.[32] Zhou offered to bargain over the depth of withdrawal, and although he rejected Nehru's treatment of Aksai Chin and NEFA as separate issues, he still insisted that the dispute had to be resolved through negotiations.[33] Nehru refused to negotiate until China

[28] Maxwell (1970, 120).
[29] See Thucydides (1996, Ch. XVII) for the Melian dialogue.
[30] Gurtov and Hwang (1980, 118), Hoffmann (1990, 72–73).
[31] Letter from Zhou to Nehru, November 7, 1959, *White Paper,* 3: 44.
[32] Nehru, November 16, 1959, *White Paper,* 3: 49. See also Maxwell (1970, 138); Hinton (1966, 291); Hoffmann (1990, 80–81).
[33] Letter from Zhou to Nehru, December 17, 1959, *White Paper,* 3: 51–55.

agreed to the facts as the Indians saw them.[34] He, however, professed himself ready to talk to the Chinese but only to persuade them of the validity of India's position.[35] The Chinese stopped their patrols anyway and proposed to leave the border undelimited until some later time. Nehru would not even agree that the border was not delimited—"On that basis there can be no negotiations"—and then reaffirmed his previous refusal to meet unless China accepted his pre-conditions.[36]

The sincerity of the Chinese desire to settle the issue by negotiations could be doubted but their actions suggests that there was no ulterior motive. Right before coming to India in the spring of 1960, Zhou Enlai went to Burma and successfully negotiated a treaty delimiting the common border. It is significant that the Burmese had clashed with the Chinese around the McMahon Line's extension into Burma and that they were quite pessimistic about their chances of extracting a good deal from the PRC should it decide to rest its claims on the threat of force. Instead, the PRC accepted the reality of the McMahon Line without ever admitting to its legality (in fact, they would not even mention it in the treaty). In other words, China's insistence that the line was inadmissible as a vestige of European imperialism had nothing to do with their pragmatic approach to the actual delimitation. Furthermore, they had given the Burmese a deal that would be puzzling in its generosity if one assumed the China harbored expansionist desires. The PRC could have certainly extracted more had it chosen to threaten Burma with force.[37] This should have been a clear signal of their intent to honor the Line in NEFA in practice without compromising their stand on its legal validity. The Chinese would proceed to negotiate fair border treaties with Nepal and Mongolia, and, shortly after the outbreak of hostilities with India, with Pakistan. Only India and the Soviet Union would remain unwilling to settle their border disputes with the PRC by negotiations.

At any rate, by late 1959 India's valuation of the territories was as high as it could be, while China maintained its moderate valuation of Aksai Chin and the very low valuation of NEFA. The status quo favored their position, and it would be up to India to challenge it. The only way to dislodge the Chinese was through military pressure, and here India was indisputably the weaker of the two. Calvin (1984, Ch. 3) gives a cogent summary:

> The Indian army was in a poor state, especially in their readiness for alpine warfare. Their fire power, supply system, training, and readiness for mountain operations were all quite lacking. They had significant personnel shortages, and would often be outnumbered by the Chinese by 5 : 1.2. To pit troops in such circumstances against an enemy superior in ever detail of military strength would be absurd… But this is what India did.

It is true that the People's Liberation Army (PLA) had its own problems. The fifties had been a decade of considerable disagreement over the proper way to modernize the armed forces and as the decade drew to a close, the Chinese could not expect Soviet assistance either. With the economic disaster of the Great Leap Forward policies, domestic resistance

[34] Letter from Nehru to Zhou, December 21, 1959, *White Paper,* 3: 56–57.

[35] Hoffmann (1990, 86–87). This was the occasion of Nehru's infamous remark that he will not negotiate on borders but that everything is negotiable. When pressed to explain the paradoxical statement, Nehru made a distinction between talks and negotiations (Maxwell, 1970, 140–41).

[36] Note from the Ministry of Foreign Affairs of China to the Embassy of India in China, December 26, 1959, *White Paper,* 3: 58–79; Letter from Nehru to Zhou, February 5, 1960, *White Paper,* 3: 80; Maxwell (1970, 155).

[37] Maxwell (1970, 160–61), Gurtov and Hwang (1980, 118), Hoffmann (1990, 85–86).

increased and occasionally burst into armed rebelion. The Party itself was rent from internal challenges to Mao's policies. It stood to reason that the Chinese would be wary of exercising force to meet external threats. The leadership carefully instructed its border troops against unauthorized actions that could precipitate armed conflict with China's restless neighbors, especially India and the USSR.[38] Despite these difficulties, the Chinese had serious military advantages over the Indians, especially in the disputed territories.

In the Ladakh sector by early 1962, "the Indians were over all outnumbered by more than five to one; but the effective disparity between their strength and that of the Chinese was far greater. It was not only that the Chinese were concentrated where the Indians were scattered, or that they were able to move in trucks where the Indians had to trek on foot; the Chinese had all regular supporting arms for their troops, while the Indian 114 Brigade had nothing beyond one platoon of medium machine guns."[39]

The Indian military was not unaware of the situation. The Chief of Army Staff General Thapar warned the government that India could not match the Chinese in the disputed area: the PRC had more troops and a much better developed infrastructure for supplying them. The reality was just as grim as he suggested. North of the Karakoram range, China could easily deploy troops and resupply them whereas India would have to airlift everything (it was later estimated that only about 30% of drops were recovered). The soldiers were not prepared for the high altitudes: they needed proper training and acclimatization. Communications were bad, medical facilities poor, and all transportation of heavy items that could not be airlifted was by mules and porters.[40] Nehru and Minister of Defense Krishna Menon simply professed belief that China would never start a war over these territories. Therefore, India's obvious military weakness was not an issue.

This was an odd logic: India was pursuing a military policy which hoped to somehow force the Chinese out of areas where the Chinese were militarily superior without provoking the Chinese into resisting force with force. But the logic makes more sense if we consider it in the light of the MTM. India was challenging China in the hope that the latter would not use its overwhelming strength to settle the dispute by force. Since China's interest in these areas was only moderate, there was a high probability that perhaps the Chinese would acquiesce without mobilizing for a costly confrontation. There were risks associated with Nehru's policies and he knew that. The cautious behavior of the Chinese in late 1961 and early 1962 only encouraged him to run higher risks.

In a sense, Nehru had little choice. By branding Chinese presence in Aksai Chin an act of aggression, he had committed the government to not negotiate with the Chinese until they left, and, if they failed to do so, to remove them somehow. Since neither sitting still nor war were options, Nehru tried a middle course of coercion, the so-called "forward policy." Instead of openly challenging the Chinese to a military showdown, the Indian military would play a game of chess with them: it would place outposts into forward positions, in close proximity to the Chinese. Since the Chinese were assumed unwilling to provoke a military confrontation, simply emplacing Indian troops in their lines of supply would force them to

[38] Gurtov and Hwang (1980, 114).
[39] Maxwell (1970, 236); Hoffmann (1990, 103).
[40] Hoffmann (1990, 98–99).

abandon their own posts and retreat. There was no disguising the fact that the forward policy was "a military challenge to a militarily far superior neighbour."[41]

The Chinese government—"exercising its sacred right to defend China's territory and maintain the tranquility on the border"—resumed border patrols in the Karakoram in April and itself began more vigorous deployments investing the Indian outposts on several occasions.[42] In one incident in the Chip Chap valley, they made a show of advancing on an outpost but retreated without attacking it. This reinforced India's view that the Chinese maneuvers were a bluff, which seemed to receive further confirmation when the Chinese did not attack the outpost Galwan despite repeated threats to do so. At this point, the inexorable logic of escalating risks persuaded the emboldened Indians to change the rules of engagement for their troops from "fire only if fired upon" to "fire if the Chinese press dangerously close to your positions."[43]

The Chinese, for their part, had become increasingly pessimistic about the possibility of resolving the dispute through negotiations. They could not help but get even more so when India, flush with military equipment from the United States and the Soviet Union, invaded Goa in total contradiction to Ghandi's peaceful philosophy to which Nehru claimed steadfast adherence. The easy success of this mission—the Portuguese put up no resistance—and the support from both superpowers, which also sided with India in its dispute with China, encouraged an even riskier policy with respect to the PRC.

The Chinese tried to impress upon Nehru that his forward policy was very dangerous. India's deployments came perilously close to the strategic roads in Aksai Chin, not to mention that they were in territories claimed by the PRC. Skirmishing had become a regular occurrence along the line of control. In the spring of 1962, China notified that India's policy is "most dangerous and may lead to grave consequences."[44] If this were not clear enough, the PRC warned India that "to refuse to maintain the *status quo* and reject negotiations is to reject a peaceful settlement."[45] By September, the threats were becoming explicit. After enumerating further incidents, the Chinese warned that "the Indian Government should be aware that shooting and shelling are no child's play; and he who plays with fire will eventually be consumed by fire."[46] Should this prove too metaphorical, the Chinese rephrased it in plain language: after calling India's insistence on preconditions "utterly absurd," they bluntly stated that "if India should continue to nibble Chinese territory, it will certainly meet with China's resistance."[47]

But Nehru pressed on, against the advice of his own military. When confronted with the

41 Maxwell (1970, 179); Gurtov and Hwang (1980, 119–20); Hoffmann (1990, 92–103).

42 Note from the Ministry of Foreign Affairs to the Embassy of India, April 30, 1962, *White Paper,* 6: 39.

43 Maxwell (1970, 236–39). Gurtov and Hwang (1980, 121–22) discern a brief period of hesitation on Nehru's part in late July. Even if he did waver, Nehru soon regained his confidence and on August 13 he reiterated his position that there would be no negotiations without the Chinese vacating the territories. Hoffmann (1990, 104–06) notes the public perception of the Galwan victory.

44 Note from the Ministry of Foreign Affairs to the Embassy of India, March 1, 1962, *White Paper,* 6: 14.

45 Note from the Ministry of Foreign Affairs to the Embassy of India, March 22, 1962, *White Paper,* 6: 24.

46 Note from the Ministry of Foreign Affairs to the Embassy of India, September 13, 1962, *White Paper,* 7: 68.

47 Note from the Ministry of Foreign Affairs to the Embassy of India, September 13, 1962, *White Paper,* 7: 72. See also Maxwell (1970, 255) and (Whiting, 1975, 94–95). Gurtov and Hwang (1980, 121-22) similarly argue that despite the Chinese warnings, the Indian government persisted in its escalatory tactics encouraged by the failure of the Chinese to respond with force.

unpleasant reality about the balance of forces, he would dismiss the concerns as irrelevant because the Chinese would not fight. And when the military insisted that the forward policy was bound to provoke them to do something along the disputed border, Nehru would replace the disagreeable soldiers with more pliable ones (Hoffmann, 1990, 143–49). The reality of India's weakness, however, could not be wished away. Since it was a carefully kept secret, the Opposition and the Indians interested in the dispute could not understand why the government was not pursuing a more assertive policy. Surely the superior Indian Army would make short work of the Chinese. Why is Nehru not ordering it to do so? Perhaps he is still trying to revive *Hindi-Chini bhai-bhai* and appease the aggressors? It was this public sentiment coupled with the apparent failure of the Chinese to follow through on their threats that goaded Nehru into escalating the risks. He had maneuvered the government into a position from which while "pursuing a policy of the utmost recklessness, it was being blamed for excessive, even craven, forbearance".[48]

For the Chinese, responding to Indian maneuvers was proving too costly. Maxwell (1970, 347) sums up precisely what the MTM would lead us to expect:

> Easy as the Indian pressure was to ward off when it came to the issue, it still kept the whole great sweep of the Sino-Indian borders alive, requiring troops to be kept in battle-readiness, creating a heavy logistical demand and complicating the problem of pacifying Tibet. For the army to be put and kept in a defensive posture, only to react to challenges launched at times and places of India's choosing, would make no sense to any strategists.

Given how high India's valuation was, it would be exceedingly costly to coerce it into backing down. Even critics of China conceded that under the circumstances, it made more military sense for the PRC "to mount a general offensive on its own terms along the entire border" rather than defeat piecemeal attacks while Beijing hoped for New Delhi to change its mind.[49] The Chinese had to reassess their policy of denying India an advantage in emplacement throughout the disputed territories while studiously avoiding serious clashes and simultaneously offering to negotiate.

While the Chinese were stiffening their resistance, the Indian government confirmed the fatal order to evict them from the southern side of the Thagla (Thang La) ridge in the NEFA on September 22, 1962. When the determined fighting at Tseng-jong in early October revealed that the Chinese were prepared to resist—with overwhelming force—Indian intrusions, the government halted the advance on the 11th and decided to stay put at the Namka Chu river. However, when Nehru left for Ceylon on the following day, he gave an interview in which he left the impression that the army would eject the Chinese by force.[50]

This statement was interpreted much the same way around the world: in the United States, the *New York Herald Tribune* pronounced it tantamount to a declaration of war on China; in Britain, it was seen as an ultimatum, and in China itself it convinced observers that Nehru had authorized the army to attack, and that a strike was imminent.[51]. After thorough preparations they made no attempt to conceal from the Indians, the Chinese forded the river during the night and attacked at 5am on October 20th, crushing the Indian 7 Brigade and setting

[48] Maxwell (1970, 291).
[49] Maxwell (1970, 347), Gurtov and Hwang (1980, 139).
[50] Maxwell (1970, 342); Hoffmann (1990, 129).
[51] Maxwell (1970, 344–45).

off the war. Two days later, Kennedy announced that the US had detected Soviet missiles in Cuba and was blockading the island.

It is worth noting that Nehru's logic, while faulted by many to be an especially odious instance of wishful thinking, was not without merit. Given India's assessment of Chinese interest in the issue (Nehru had actually proposed to the Chinese that they could still use the Aksai Chin road for civilian traffic, the implication being that India would probably not look too closely at what actually was being transported), it made sense that there was a good chance that the PRC would not use force to resolve the dispute.[52] This assessment was strengthened by the stream of offers to negotiate emanating from Beijing as well as the repeated failure to follow through on threats to resist Indian incursions by force. The public debate in India had made its own high valuation crystal clear. Hence, the logic went, the Chinese would not be able to extract concessions, not unless they resort to force which they would not do on account of their relative disinterest.

The problem was that the status quo favored the Chinese position: if India were to change that, it had to challenge the PRC militarily. There was no other way to compel the Chinese to vacate the territories they held and India claimed. And vacate them they must, the government's own policy had obliged it to force them to if they did not. It was here that the forward policy entailed a serious escalation of risks. The Chinese would have been content to let the matter die from lack of interest but they could not sit idly by when the Indians were maneuvering their chess pieces in what amounted, from Chinese perspective, to an unwarranted land grab. Having decided that it would be too costly to coerce Nehru, the Chinese used their military superiority to impose the resolution they had hoped to achieve through negotiations. When they declared their unilateral cease-fire on November 21 and withdrew from the territories they had overrun in the course of the war, the Chinese enforced the solution that Zhou had offered India back in 1959. It was an exemplary military action in support of a clear political goal.[53] It also illustrates very nicely the MTM result that military coercion can be costlier than war, making fighting preferable under certain circumstances.

As a final note, it may be useful to reflect on the fact that the bargaining model of war would have trouble accounting for the war. The status quo distribution of benefits on the eve of the crisis reflected the distribution of power. It was more advantageous to the PRC, making India the dissatisfied state. The problem from a theoretical standpoint is that despite that dissatisfaction, the match between the distributions of power and benefits should have made war exceedingly unlikely. India would not be expected to fight to overturn the status quo—its military weakness did not make for a great expected benefit from doing so, and China, the satisfied actor, would then have no reason to fight to preserve the status quo.

Similarly, the preponderance of power school would be puzzled: given the obvious mili-

[52] Note from the Ministry of External Affairs to the Embassy of China, May 14, 1962, *White Paper*, 6: 43. The Chinese wondered why they would need "India's permission for using its own road on its own territory." Note from the Ministry of Foreign Affairs to the Embassy of India, May, 1962, *White Paper*, 6: 56.

[53] Maxwell (1970, 418). Gurtov and Hwang (1980, 141), however, argue that even though "the October offensive was politically controlled and carefully preplanned… despite their achievement of a military victory on all fronts, the Chinese failed to bring the Indian government to the conference table." That may be so, but the PRC secured its road in Aksai Chin and no longer had to contend with aggressive Indian troop deployments. Or, as Hoffmann (1990, 226) put it, "India's refusal to grant legitimacy to China's 'line of control' would not change the fact that India was now forced to tolerate it."

tary superiority of the Chinese, the Indians should have acquiesced to the status quo. There was no dangerous power transition: China was certainly not declining in relative strength compared to India. The only school that might have something to say here is the balance of power: the Chinese simply used their strength to their advantage. The problem here is that balancing is supposed to arrest expansionist tendencies of the powerful, and these are not in evidence. Instead, the Chinese settled their disputes amicably and fairly with anyone who cared to negotiate with them, often giving up a lot more than what they could have secured by force. The dramatic unilateral cease-fire in their war with India and the subsequent withdrawal from areas they could have occupied also reveals that they had no interest in doing so. It appears that the logic of costly coercion is perhaps most useful in illuminating this particular episode.

6

The Expansion of the Korean War, 1950

The Communists frequently adopt a threatening posture with the cold-blooded purpose of so frightening their enemies that the latter will surrender without a fight.
O. Edmund Clubb, November 7, 1950

Although the primary contribution of this study is theoretical, it is worth exploring the implications of the new insights offered by the militarized deterrence model. I do not wish to claim that this model somehow captures "enough" of the strategic interaction during crisis to apply it to any particular event and claim that it fully explains what transpired. My focus has been on clarifying the dynamics of military escalation, and as such I have entirely neglected other factors, such as domestic politics, the structure of the international environment, the potential behavior of third parties, and anticipated consequences of one's actions for future interactions. All of these are important for determining crisis behavior. They can, and have been, easily accommodated in the rationalist framework. Even traditionally competing approaches such as the organizational or bureaucratic explanations can be incorporated into this framework if we conceive of the domestic structure as a configuration of various interest groups with particular sets of preferences that make up the context of constraints and incentives in which leaders operate. I also happen to think that we have not paid sufficient attention to such factors as considerations of prestige, honor, and emotions such as anger and fear that must be accounted for if we are to gain a good grasp of the phenomenon we wish to study (O'Neill, 1999). This is a formidable set of omissions, and I have no doubt that the tendencies identified by the militarized deterrence model will sometimes be overwhelmed by some combination of other tendencies produced by these factors.

This is not to say that they all work at cross-purposes. In fact, some can be positively reinforcing. For example, if we assume that commitment of ground troops in a crisis itself creates audience costs for reasons of domestic politics and international prestige, then this can only reinforce the credibility of the escalatory threat. Although the model does allow for audience costs, I assumed that they do not vary with mobilization. That is, they are incurred by the failure to follow through on a threat and are independent on the precise nature or magnitude of that threat. This is not the case in Fearon (1994a), where costs increase in duration of escalation, or in Fearon (1997), where the magnitude of these costs can be chosen by the actor. These types of models are better suited to study the actor's willingness to engage audiences for coercive purposes. We could think then of militarization as a reinforcing tactic that turns backing down into an issue of honor. In other words, depending on perspective, we could think of audience costs enhancing the credibility of a military threat or as the military

threat creating the audience costs. It must remain a venue for future work to disentangle these nuances.

Therefore, on one hand I see the main purposes of this theoretical model to be explicative and generative (Clarke and Primo, 2007). The model explores the causal mechanism linking crisis escalation to the outbreak of war and produces previously non-obvious implications that must be investigated in further studies. On the other hand, if the tendencies it identifies are strong enough to warrant all that attention, we should see some evidence of that in the data patterns. Hence, I cannot in good faith ignore the predictive purpose even if I remain skeptical about the extent to which we should trust either supporting or disconfirming evidence. Unfortunately, data on military moves during crises are exceptionally hard to come by. In lieu of a multi-year data collection effort with very uncertain prospects of success, I offer a look at the Chinese intervention in the Korean War. I use the insights of the militarized deterrence model along with documentary evidence to challenge the explanation that seems to be prevalent in the political science literature.

It is not my purpose to investigate the origins of the Korean War. Rather, it is to offer an argument that explains why the US and China ended up fighting each other even though neither one initially wanted to. In doing so, I challenge some existing interpretations and show how the model can illuminate some of the complex dynamics during the crucial weeks in late September and early October.[1] The following discussion should not be treated as a "test" of the theory; in fact, I chose the historical case specifically to demonstrate how the model can be applied to clarify a hotly contested issue. On the other hand, the evidence seems to support the counter-factual claims that the model shows as necessary to sustain equilibrium behavior, and so in a sense, provides support for the causal mechanism identified by it.

6.1 An Outline of the Argument

When the Democratic People's Republic of Korea (DPRK) invaded the Republic of Korea (ROK) in June 1950, neither the United States nor the People's Republic of China (PRC) wanted to get involved in a war with each other over the disposition of the peninsula. Despite some calls urging the unification of Korea after the expulsion of the North Koreans from the South, President Truman and his advisors proceeded with great caution throughout the summer. Both the Soviet Union and China remained relatively quiescent, probably because they fully expected DPRK to win a quick victory, and seemed as unwilling to risk war with the United States as the United States was with them. And yet, by late November, China was at war with America. How did that happen? To make the question more precise, why did the United States attempt unification by force even though the original goal was to restore the *status quo ante bellum*? If the Chinese were serious about intervention, why did they fail to deter the Americans?

One widespread answer is that the Americans did not foresee that the Chinese would intervene to prevent the collapse of North Korea. The reasons for this failure could be rational—the Chinese did not signal their intent clearly—or non-rational—the Americans persisted

[1] For an excellent summary of Korean War studies, see Brune (1996). Chen (1996) offers a fascinating glimpse at the chequered history of the scholarship on the Chinese intervention itself.

in their delusion despite clear warnings to the contrary. Most explanations have tended to be of the non-rational variety, at least until the 1990s. As we shall see, there are numerous problems with that approach.

At the risk of oversimplifying my argument, let me attempt a summary. The Chinese signals were ambiguous, and that was *not* entirely unintentional. Mao Zedong did not want to make a clear commitment because he was not sure he wanted to fight the US over Korea and because of the uncertain support of the Soviet Union. There were many very good reasons for staying out of Korea. His moves were tentative and always allowed for the possibility of disavowing them. From the chosen method of communication with the US to the military preparations he undertook, all evidence suggests that even as late as October he was vacillating. When we add these factors to the missed opportunity for optimal military intervention in late September and the widely known technological inferiority of the Chinese army (plus the crippling lack of air support), we have what amounts to a persuasive case for dismissing the half-hearted Chinese threats as bluffs. The Americans had excellent reasons to doubt China's commitment for China had none.

On the US side, the worries about Chinese intervention subsided when the PRC failed to intervene after the stunning success of General MacArthur's September 15 amphibious assault at Inchon that shattered the North Korean army and allowed General Walker to break out of the Pusan perimeter. The military balance on the ground had become so favorable that the US administration shifted its war aims from liberation to unification, and it was quite prepared to risk Chinese intervention, something that would have been unthinkable just a few weeks earlier. There seems to have been a very narrow window of opportunity for the Chinese to deter the Americans: an overt entry across the Yalu River but north of the 38th Parallel would have prevented the US from attempting unification. This window extended roughly from September 15th (the landing at Inchon) and October 9th (when the Joint Chiefs told MacArthur to proceed even if there was evidence of entry of major Chinese units). Unfortunately, Mao did not prevail in the Politburo on intervention until October 5th and would not order mobilization until the 8th, when it was too late to deter the Americans with threats. At this point, war was inevitable.

The militarized deterrence model can provide a lot of insight into these dynamics. As we shall see, the model predicts that had the Chinese announced right around Inchon the entry of a massive force that would have given them roughly the 1.6 : 1 advantage they would have in October, then the probability of successful deterrence would have been more than 80% (and if they had done so at the 1.9 : 1 advantage they had before Inchon, it would have been 100%). In contrast, the model predicts that even with an open announcement of such an entry in early October, the probability of successful deterrence was zero! Add to this pessimistic assessment the military advantages of surprising the Americans, and there is little wonder that the Chinese concealed their preparations for entry.

To make the model's logic compelling as an explanation, it will be necessary to demonstrate that the United States would have been deterred by clear threats; that had the Chinese sent an unambiguous signal of their intention to intervene as massively as they ended up doing, then the US forces would have remained south of the 38th parallel. It is doubtless true that had China threatened with general war over Korea the United States would not have attempted forced unification. However, as we shall see, there are two crucial periods in which the extent of the threat mattered greatly: before the Inchon landing, a clear threat by either

China or the Soviet Union would have deterred the US. After Inchon, only a clear threat by the Soviet Union would have done so; a Chinese threat or even overt intervention would work only if it came with open Soviet support (which was not forthcoming), or if it was a start of a general war that would activate its treaty with the USSR (which would also bring the Soviets into the fray, something both Stalin and Truman were equally anxious to avoid), or if it involved full commitment to an intervention on a massive scale. In other words, by early October the United States had become undeterrable by the means the Chinese had at their disposal or were willing to use. Without open Soviet support, the sole remaining possibility was to reveal the scale of Chinese Communist forces (CCF) in Korea. Unfortunately, given the general opinion of their quality and the lack of air support, doing so would have exposed these armies to a devastating attack by the UN forces (UNF) thereby almost wholly negating their value. The only way to convince the Americans to abandon the reunification goal at this point was to bloody them sufficiently, and given the military inferiority of CCF, this meant surprising them, which required concealing the extent of mobilization, which meant losing its deterrent value.[2]

When the "new war" came on November 26, neither side was prepared to back down: the Chinese had missed the opportunity to deter the US because they were not sure they wanted to take the risk by a clear commitment at the time this window was open. When the window closed, only a massive threat would have stopped the march to the Yalu but the Chinese could not make this threat without risking losing its capability and at any rate the US administration was prepared to risk fighting China as long as the Soviet Union stayed out. Since the US could only be persuaded by a credible revelation of capability, this virtually ensured that the war would expand.

While this explanation absolves the US administration of much of the blame traditionally heaped upon it, it does not go to the other extreme by laying the fault with the Chinese. It really is difficult to see just what policy-makers on both sides could have done at the time given the information they had. Before Inchon, the Chinese had no reason to threaten intervention, and afterwards the brief opportunity to deter the Americans passed before they could do it credibly. As Maoz (1990, 119–23) concludes, the conflict between China and the United States over Korea was a tragic war which both sides had wished to avoid.

6.2 Militarized Deterrence in Korea

The explanation for the expansion of the Korean War offered here requires some counterfactual analysis. To make the argument that it was possible for the Chinese to deter the US on their own before Inchon but not in October, I have to show that US policy-makers had considered the circumstances under which MacArthur could be permitted to attempt unification and that these changed radically over the last few weeks of September. Before offering historical evidence for this, I will calibrate the military threat model (MTM) with values for the variables and then compare the October situation with the hypothetical scenario in which the Chinese make a clear threat right after the Inchon landing. In both cases, the US is the challenger and the PRC is the defender: if the UN Forces halted at the 38th parallel, then North Korea would survive as the status quo.

[2] I explore the incentives to feign weakness in such situations in Slantchev (2010).

The variables we need to consider are the two opponents' valuation of the disputed issue (the unification of Korea under South Korean rule), their beliefs about each other's valuations, the military balance, the costs of war, and the political costs of backing down. Table 6.1 shows the values used for the model in terms of the simulations used throughout the rest of the book.

	US (Challenger)		PRC (Defender)	
	Pre-Inchon	Inchon to Yalu	Pre-Inchon	Inchon to Yalu
Valuation	moderate	moderate	high	high
Beliefs	moderate	low	moderate	high
War Costs	moderate	moderate	moderate	moderate
Audience Costs	moderate	high	low	low

Table 6.1 *The Korean Conflict: Parameters for the US and the PRC.*

We shall assume that the opponents' respective valuations of the issue remained fixed (that is, their basic preferences did not change). The United States, although moderately interested in Korean unification on South Korean terms, did not consider it a vital interest. The US had to forestall unification of Korea by Pyongyang: the consolidation of Soviet rule in Europe had shown all too clearly what Communists intended to do once they were in control (Pollack, 1989, 214). The "loss" of China also contributed in at least two ways. First, the US now had to concentrate on helping Japan recover so that it could emerge as the American partner in the region. But if communists were allowed to have Korea, they would be able to threaten Japan which is barely 100 miles from Pusan. Second, the Truman administration was vulnerable to charges of "another Munich," and the domestic mood in America was not at all conducive to conciliatory diplomacy with any Red state (Kaufman, 1997, 22).

Although it was willing to work with the United Nations to prevent unification on North Korean terms, anything beyond that was outside of the famous "defense perimeter" articulated by Dean Acheson. The one argument for crossing the 38th parallel that found currency early on was articulated by the Joint Chiefs of Staff who insisted that restoring the *status quo ante bellum* would not guarantee the security of the South Korean regime. They "felt that if we were required to stop at the 38th Parallel, nothing would have been done to solve the real problem.... If you stopped at the 38th Parallel, then the North Koreans, supported by the Chinese and the Russians, could once again attack when they were ready to. The 38th Parallel had no defensive merit whatsoever."[3]

For China, on the other hand, the situation was more important. Korea was right at its doorstep, and a hostile presence there would certainly jeopardize the conquest of Taiwan and would tie down a significant number of forces along the Yalu River precisely when the leadership would either need them for Taiwan (and Tibet) or would want to demobilize them for economic development reasons. In other words, a unified hostile Korea was something that worried the Chinese quite a bit. Since the Chinese valuation is of crucial importance for

[3] Lieutenant General J. Lawton Collins, Truman Library Institute conference comment, May 1975. Cited in Heller (1997).

the argument I am making as well as the explanations I want to contest, I will expand on that later in the chapter.

In terms of beliefs about the opponent's valuation, the US started out thinking that Chinese intervention, although far from certain, was a distinct possibility because the PRC could not but care about what happened to Korea. However, because the US administration knew it had no designs on China itself and because it assumed it could make this fact clear beyond doubt to the Chinese, they judged that the Chinese interest was moderate. After the Chinese failed to intervene when they should have, the administration revised its estimate of PRC's valuation downward. The threats, vague and late as they were, were correspondingly interpreted as the Chinese making a play for concessions on Taiwan and their seat in the United Nations. The Chinese, on the other hand, started out thinking that the US would not be so foolish as to attempt unification—all American policy up to that point was in some sort of uneasy cooperation with the Soviet Union. However, when MacArthur's forces shattered the North Korean army and the UNF began its advance north, the Chinese revised their expectation upward: it seemed that the Americans were willing to take serious risks to unify Korea, which doubtless meant that they cared about it more than previously thought. Every disadvantage China would have from a hostile Korea was an advantage to the Americans. Not surprisingly, the very act of making a challenge causes the defender to become more pessimistic about the type of challenger it is facing.

In this book we have assumed that the war costs are independent of the levels of forces committed to the fighting (see the discussion in Chapter 7 on p. 177). This implies that these would not have changed in September for either actor. I assume the war costs are moderate for both actors: the United States did not have to mobilize for a full out war and China had no way of striking at the American homeland to inflict pain even if it had wished to, which it did not. The Chinese were also able to mount a successful invasion of Tibet right before intervening in Korea, and although they had postponed their planned invasion of Taiwan, they were probably confident that the US would not start a general war with them because doing so would have drawn in the Soviet Union. Hence, both sides expected significant but not overwhelming costs of war.

Audience costs, on the other hand, may be assumed to have changed for the United States. If the administration had announced its intention to unify Korea before Inchon (or shortly thereafter), then backing down after a clear Chinese threat would have incurred some political costs. However, since the alternative would have meant starting a new war with China over something that had not even been the original aim these domestic costs would have been, at most, moderate (not to mention that expanding the goal would not have been endorsed by the United Nations). Over the next few weeks, with the Chinese failing to go in, with the UN moving toward endorsement of unification, and with the military situation dramatically changing against China, the domestic clamor in support of the new goal predictably grew to a crescendo. The Republicans in Congress were especially vocal. Backing down in the face of a Chinese threat, although still possible, would have carried correspondingly higher audience costs. Mao, on the other hand, faced relatively low audience costs for backing down over Korea. Not only was he careful not to engage China's prestige by making clear public threats, but he even had to persuade a doubtful Politburo to support the

intervention.[4] In other words, whatever misgivings his supporters might have had about a unified Korea at China's doorstep, it would not have been too hard to convince them of the wisdom of not engaging the superpower in war, had Mao wished it. [5]

With these estimated values of the parameters in hand, we can see what the model would predict for a threat made in mid September and one made after the first week of October. We now need to consider the distribution of power for the two contingencies under consideration.[6] The military balance is not easy to ascertain for several reasons. First, as the fighting raged on, casualties on both sides mounted just as the United States was reinforcing the United Nations Forces (UNF), making for serious fluctuations in the totals. Second, it is very difficult to be certain about the numbers China began to ready for a possible intervention. On September 1, the United Nations ground forces, including the remaining ROK units, numbered close to 180,000 men. The North Korean army ranged for the assault on Pusan on that date had about 98,000 men (Appleman, 1961, 382,395). Far East Command had been worried about possible Chinese intervention almost from the outset of the war (its daily intelligence summary of June 28 explicitly raised the possibility). The PRC had been moving units to Manchuria for the past several months, and by the end of September the Americans estimated that perhaps up to 250,000 men of the Chinese Communist Forces (CCF) were in position to enter the war (Appleman, 1961, 758). The actual number seems to have been around 300,000 CCF troops. Using the US estimate, this gives us a rough idea of the military balance around the time of the Inchon Landing: 180,000 UNF versus a combined total of 348,000 CCF/DPRK should the Chinese intervene. In other words, 1.9 : 1 ratio in favor of the defender with distribution of power (DOP) of approximately 66%. In real terms, the CCF/DPRK total was closer to 400,000 which yields the ratio of 2.2 : 1 with DOP of about 69%.[7]

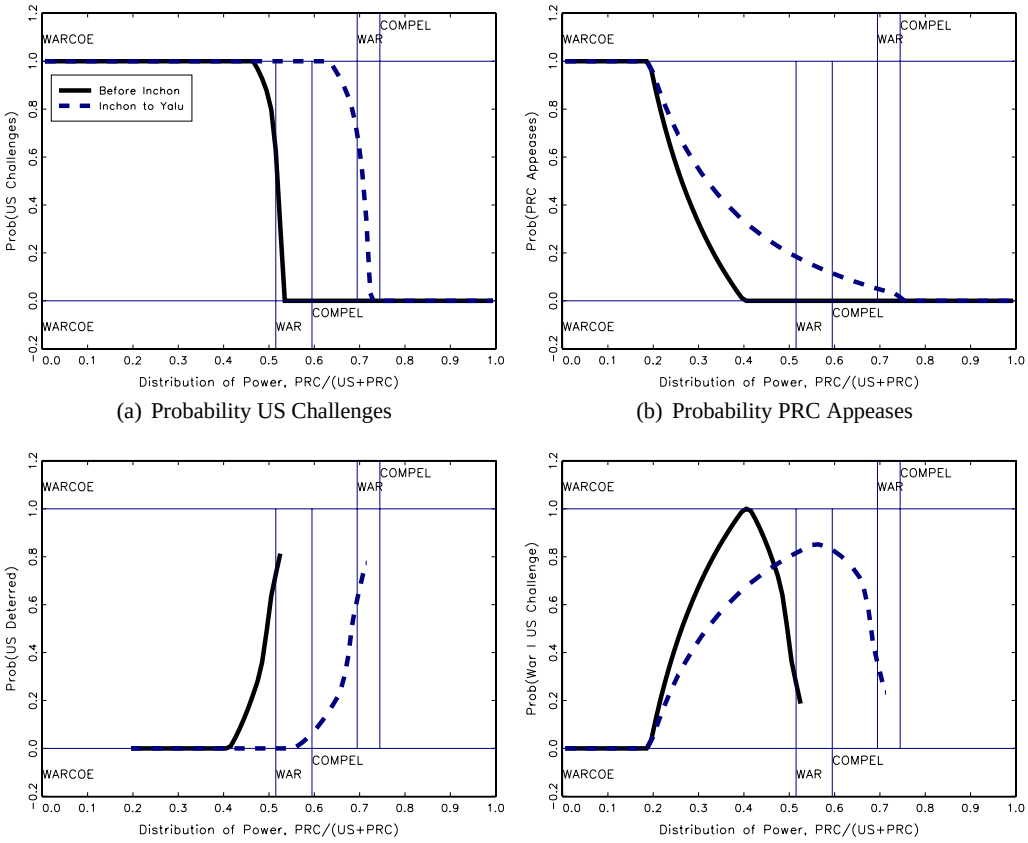
By the end of September, UNF had swollen to 229,772 men and by mid October there were approximately 400,000 CCF troops ready to cross the Yalu (Appleman, 1961, 606,751). Since the North Korean army was all but destroyed by that time, we can exclude its remnants from the estimates. Using the actual CCF strength (of which 120,000 were already across the river in North Korea), the ratio is 1.74 : 1 with DOP of about 64%. MacArthur himself told Truman on October 15th that he believed there were about 300,000 CCF in Manchuria but no more of 60,000 could actually cross the river. The October 20th US intelligence estimate put the CCF number at 400,000 at the river and on alert to cross—in reality, 180,000 had already made it into North Korea (Appleman, 1961, 760,767). This puts the US-estimated

[4] See Gurtov and Hwang (1980, 47-56) for a discussion of the internal divisions about intervention.

[5] The model parameters are as follows. The valuation ranges are $[5, 20]$ (low interest), and $[10, 25]$ (high interest). Moderate interest is the intersection of the two, $[10, 20]$, and the upper bound is used to compute the costs. War costs are set at 20%, while audience costs are set at 1% (low), 4.25% (moderate), or 7.5% (high). For instance, after Inchon the US (the challenger, S_2) believes that interest of the PRC (the defender, S_1) is low, whereas the PRC believes the US interest is high. The calculations are performed under the assumption that S_2 believes that v_1 is distributed uniformly on $[5, 20]$, whereas S_1 believes that v_2 is distributed uniformly on $[10, 25]$. War costs are $c_1 = c_2 = 4$, and audience costs are $c_1 = 0.2$, $c_2 = 1.5$.

[6] Since PRC is the defender, these represent the ratio of Chinese and DPRK forces to the total which includes those of the United Nations and ROK.

[7] This obviously ignores the quality of the troops and the serious technological edge UNF enjoyed. However, even if we degrade Chinese capability somewhat, the basic logic holds.



(a) Probability US Challenges

(b) Probability PRC Appeases

(c) Probability US Deterred Successfully

(d) Crisis Instability

Figure 6.1 The Military Threat Model and the Korean War.

ratio at somewhere between the low 1.3 : 1 (DOP of 57%) and the actual 1.74 : 1 (DOP of 64%).[8]

We are now ready for our counterfactual analysis. Figure 6.1 shows the predicted probabilities given our assumptions.[9] Had the Chinese openly threatened intervention before the end of September, the threat would have worked even at the American estimate of the DOP at 66%, and certainly at the real DOP of nearly 70%. The probability that the US would challenge the PRC by crossing the 38th parallel would be zero. Had the Chinese made their intentions clear, the US would have no reason to believe that they were lying—the prob-

[8] In fact, Far East Command underestimated the number of CCF in North Korea up until the UNF faced them in battle in late November. On the eve of the November 24th attack, the maximum number of CCF in North Korea was put at about 70,051. In reality, there were more than 300,000 Chinese troops there already (Appleman, 1961, 763,768). By that point, however, the UN force had increased to 440,000 troops "of vastly superior firepower" compared to the CCF (Whiting, 1960, 122). That is, deterrence was bound to fail.

[9] The equilibrium types are listed in the bottom information pane for the pre-Inchon scenario and in the top information pane for the post-Inchon scenario.

ability that the PRC would appease a challenge is zero. Since the US would have been successfully deterred, crisis instability is zero, so the war would not have occurred.

Contrast this with the situation that developed by early October. Neither the estimated DOP of 57% nor the real DOP of 64% would have deterred the US. That is, even if the PRC had revealed its strength at this point, the US would have still continued to press toward the Yalu. This is despite the high probability that the PRC would resist: the likelihood of appeasement is low, at around 10–15%. The probability that this escalation would deter the US is itself negligible, topping off at 15% at the real DOP (and not even reaching 1% at the estimated DOP). The crisis would be highly unstable, with probability of war around 80%. Observe in particular that under the new circumstances, the Chinese would have had to achieve a DOP of at least 70% before the Americans would become deterrable in an appreciable degree. Given the major UNF buildup, this they could not do without open support from the Soviet Union.

It is crucial to note that the US misperception about the extent of Chinese preparation in early October is *not* the reason why the Chinese threats failed. While it is true that the Americans had severely underestimated the size of CCF they were facing, even at the actual distribution of power the Chinese could not deter them. If the Chinese had somehow managed to reveal the real DOP, the US would still have challenged with a very high probability (around 90%), and the chances that militarization would have deterred the US successfully would be only about 20%. The logic of the MTM is merciless: given the altered military situation on the ground, the changing beliefs about the opponent, and the domestic mood in the US, deterrence was beyond China's reach by early October. This rationalizes Mao's decision to enter the war in utmost secrecy—there was not much to gain from trying to deter the Americans at this highly disadvantageous DOP.

I now turn to the historical evidence. In doing so, I will deal with an alternative explanation of the failure of the Chinese threats and I hope I will be able to make a persuasive case that they did not fail because the US government irrationally dismissed them as bluffs. Of course, it is impossible to prove that the administration was not irrational. My goal is more modest: I just want to prove that one need not assume that it was—its actions (and those of the Chinese) are readily understood within the rationalist framework of the military threat model.

6.3 The Evolution of US War Aims and Chinese Signals

6.3.1 Liberation Without Unification

The initial war aim was precisely what was authorized by the UNSC Resolutions of June 25 and 27: reverse the North Korean advance and expel the invading army back across the 38th parallel and reoccupy South Korea "quickly and cheaply."[10] There was a disagreement about the wisdom of conducting any military operations north of that line. Paul Nitze and George Kennan both argued against crossing even for military purposes because the risk of Soviet or Chinese intervention was too high. Dean Rusk and John Allison, and "nearly every other government quarter" disagreed. They supported operations across the parallel if these would secure the South. Allison even registered "emphatic dissent" with a memo from the Policy

[10] Acheson (1987, 450), Offner (2002, 387), Stueck (2002, 87).

Planning Staff (PPS) that argued against crossing except for strictly tactical purposes. This forced the PPS to moderate its line: in the second draft, it conceded that the decision what to do upon reaching the parallel should be "deferred until military and political developments provide the additional information necessary."[11]

The July 31 Pentagon memo argued that while a "return to the *status quo ante bellum* would not promise security," stopping at the parallel (which it disparaged as the "geographical artificiality violating the natural integrity of a singularly homogenous nation") would be least likely to provoke the Soviets. The military suggested that unification was possible provided the US was willing to mobilize sufficient resources to effect it and the Soviet Union did not intervene.[12] John Foster Dulles concurred: barring Soviet or Chinese intervention, there was no reason to stop at the parallel.[13] The CIA and Allison both emphasized that while unification was a nice goal, trying to accomplish it was just too risky.[14] The Department of State added that for any such policy, UN endorsement would be necessary.[15].

On August 24, Truman accepted NSC 73/4, which combined the desirability of unifying Korea with everybody's misgivings into an essentially wait-and-see policy. The UNF would be allowed to operate north of the parallel for military purposes provided neither the Soviet Union nor China had intervened or announced intention to intervene. The NSC agreed that under no circumstances should the Korean incident lead to a war with either USSR or PRC. This, however, should not prejudice MacArthur's efforts to destroy the DPRK army in the south.[16]

On September 11, Truman signed NSC-81/1, formally adopting this policy. It is worth quoting the relevant passages in full:

> It would be expected that the UN Commander would receive authorization to conduct military operations... in pursuance of a roll-back in Korea north of the 38th parallel, for the purpose of destroying the North Korean forces, provided that at the time of such operations there has been no entry into North Korea by major Soviet or Chinese Communist forces, no announcement of intended entry, nor a threat to counter our operations militarily in North Korea. Since such operations would involve a risk of major war with the Soviet Union... the UN Commander should, prior to putting any such plan into execution, obtain the approval of the President.... The United Nations Commander should undertake no ground operations north of the 38th parallel in the event of the occupation of North Korea by Soviet or Chinese Communist forces, but should reoccupy South Korea up to the 38th parallel.[17]

[11] Memorandum by Allison to Rusk, July 15, 1950. US Department of State, *Foreign Relations of the United States, 1950* (Washington, DC, 1976), 7: 393–95. Henceforth, all references to this volume are abbreviated as FRUS. Memorandum by Allison to Nitze, July 24, 1950, FRUS: 458–61; Offner (2002, 387), Stueck (2002, 88), Kaufman (1997, 40).

[12] Draft Memorandum Prepared in the Department of Defense, July 31, 1950, FRUS: 502–10.

[13] Memorandum by Dulles to Nitze, August 1, 1950, FRUS: 514–16

[14] Memorandum Prepared in the Central Intelligence Agency, August 18, 1950, FRUS: 600–03.

[15] Draft Memorandum Prepared in the Department of State for National Security Council Staff Consideration Only, August 23, 1950, FRUS: 635–39

[16] Memorandum of Conversation, August 25, 1950, FRUS: 646–48. Memorandum with Rough Notes on NSC Senior Staff Meeting on Korea, August 25, 1950, FRUS: 649–52. Memorandum of a Teletype Conference Prepared by the Department of the Army, August 30, 1950, FRUS: 659–60. Draft Memorandum Prepared in the Department of State for National Security Council Staff Consideration Only, August 30, 1950, FRUS: 660–66.

[17] Report by the National Security Council to the President, September 9, 1950, FRUS: 712–21. The passages quoted are on p. 716.

The authors of the document went out of their way to make it abundantly clear that this was *not* to be an authorization to attempt unification—the crossing would be strictly for military purposes only. Anything beyond that would require explicit approval from the President and the United Nations. Furthermore, if there was an open announcement of either Soviet or Chinese intent to occupy North Korea, MacArthur should not attack their forces and should instead immediately refer to the Security Council.[18] In other words, four days prior to the Inchon landing, the Americans would have been deterred from crossing the 38th parallel, *even for military purposes*, if the Chinese entered North Korea or even if they stated a clear intention of doing so in force. The policy was also adamant that while military operations north of the parallel were authorized by the June UNSC Resolutions, unification was not.

6.3.2 Unification Becomes a Tempting Possibility

As September wore on and the total collapse of the North Korean army became evident, American policy-makers directed their attention to the Soviets and the Chinese. The temptation to pursue unification was getting stronger and the Chinese were only making vague verbal threats. The US administration attempted to ascertain what they intended to do. Would the PRC send token volunteers or would it intervene in force? Acheson had attempted to warn off the Chinese through the Indian Ambassador Panikkar. His argument was that the rapidly changing military situation indicated that the UN might be able to "restore peace quickly in Korea" and because this was an action authorized by the UN, there was no threat to China.[19] It appears that Panikkar did not deliver that message when he met wet Zhou Enlai on the 20th because the latter stated that the PRC would not intervene in Korea short of a global war caused by the open Soviet entry after the UNF crossed the 38th.[20] Zhou also espoused that Mao-approved principle of PRC intervention: "war… should be conducted as a protracted war on the basis of self-reliance."[21] This was three days after the Chinese had sent military officers to Korea to survey the situation and "prepare for future battles."[22]

The quiescence of the Soviet Union was puzzling. On the 22nd, John Davies of the PPS wrote that by remaining uncommitted, the Soviets "abandoned the optimum opportunity for guaranteeing that UN forces would be prevented from pressing north of the 38th parallel." He speculated that this may be a ploy to lure the US into overextension and then strike with overwhelming force. His recommendation was for the US to proceed north with great caution, launching probes to test Soviet resolve, and expanding the areas under control if there was no response.[23] On the same day, there came the first public warning from the PRC

[18] NSC-81/1 which was circulated on the 9th was a slightly revised version of NSC-81 circulated as the Draft Report by the National Security Council on United States Courses of Action with Respect to Korea, September 1, 1950, FRUS: 685–93.

[19] Telegram from James Webb (Under-Secretary of State) to the Embassy in India, September 16, 1950, FRUS: 733.

[20] Telegrams from Henderson to Acheson, September 20, 1950, FRUS: 742–43.

[21] Chen (1992, 16), Zhang (1995, 77).

[22] Chen (1992, 16), Zhang (1995, 74). On the 16th, the Chinese also assisted Ho Chi Minh in a surprise attack on the French in Vietnam. By October 10th, this offensive secured the Sino-Vietnamese border (Zhang, 1995, 69–70).

[23] Draft Memorandum by Davies of the Policy Planning Staff, September 22, 1950, FRUS: 753–55.

that it would "resolutely oppose the criminal acts of American imperialist aggression."[24] The timing of the official hardening of the Chinese position may have had something to do with the failed initiative to get the PRC the UN seat whose outcome had become obvious the day before. Rusk's take on this was more belligerent than Davies': he interpreted NSC-81/1 as an authorization to fight a Chinese intervention; only an overt Soviet entry—which would indicate the start of a general war—would stop the UNF. In his opinion, neither was likely.[25]

On the 23rd, the US accidentally bombed targets in Manchuria. Mao was livid but Truman offered compensation. State received a report from Wilkinson which estimated that PRC's aid to North Korea would be token. There was evidence that about 250,000 troops might be sent to Korea but in North Korean uniforms.[26] On the next day, Zhou protested US incursions over Andong to the United Nations and warned that if the General Assembly did not pay more attention to these aggressive tactics, it would "share in the responsibility for lighting up the war-flames in the [Far] East" (Zhang, 1995, 77). This was contradicted by Wilkinson's report that Chu Teh (Commander-in-Chief of the PLA) had said the PRC would not intervene because it would need much more preparation. Wilkinson acknowledged that his source was of dubious reliability but that it nevertheless confirmed the substance of Zhou's statement on the 21st.[27]

On the 26th, Seoul fell to the UNF and Alan Kirk (US Ambassador to the USSR) reported that Soviet summaries of Mao and Chu Teh speeches on the 25th indicated that "these leaders now assert foremost task Chinese Communists is to build up strong army." In his opinion this was highly significant because it was tantamount to a shift in their priorities away from economic development and reconstruction toward the creation of a military strong enough to defend the frontiers. This, however, was to be a long-term goal, not a preparation to enter the Korean War.[28]

With all this information at hand, the US government transmitted the NSC-81/1 instructions to MacArthur. In a memo to the commander, the Joint Chiefs of Staff (JCS) also ordered him to "continue to make special efforts to determine whether there is a Chinese Communist or Soviet threat to the attainment of [his] objectives." Although he was authorized to fight even major PRC units if they intervened south of the parallel, he was not to proceed north if the Chinese had entered or threatened to enter North Korea.[29]

Panikkar now "reinterpreted" his talk with Zhou on the 21st and began to claim that the PRC was going to respond more aggressively and would intervene indirectly in North Korea.[30] Acheson and Ernest Bevin (the British Foreign Secretary) had essentially the same

[24] Zhang (1995, 77).

[25] Memorandum of Conversation, September 23, 1950, FRUS: 760.

[26] Offner (2002, 390). Telegram from Wilkinson to Acheson, September 22, 1950, FRUS: 765. Memorandum about conversations between Panikkar and Chinese Communist Officials, September 27, 1950, FRUS: 793–94.

[27] Telegram from Wilkinson to Acheson, September 25, 1950, FRUS: 768.

[28] Goncharov et al. (1993, 170). Telegram from Kirk to Acheson, September 26, 1950, FRUS: 779–80.

[29] Telegram from Webb to the US mission at the UN, September 26, 1950, FRUS: 781.

[30] Hubert Graves at the British Embassy in the US warned the US administration not to take Panikkar too serious because he was "volatile and an unreliable reporter." See the September 27 memo of the conversation with Graves, FRUS: 794. Panikkar did have credibility problems. Webb also urged that the US not use the "dubiously reliable intermediary Panikkar" who had "predispositions and free-wheeling proclivities." See Webb's telegram to the Embassy in India, October 4, 1950, FRUS: 875. Even Bajpai remarked that some of

reaction to Panikkar information: it was a predictable reaction to China's disappointment over the UN seat. Despite Indian fears that a UNF crossing of the 38th would provoke an intervention, the British estimate was that such an intervention "would be basically contrary to Chinese interests and not likely to occur."[31] The Dutch were sounding the tocsin as well but Kirk wrote back from the Soviet Union dismissing both warnings. From where he stood, it looked likely that the "Chinese Communists, thru press propaganda and by personal contacts with foreign diplomatic personnel Peiping, have taken strong line since Inchon landing bluff UN on 38th parallel issue."[32]

It was on this day that George Marshall cabled MacArthur the infamous order instructing him to "feel unhampered tactically and strategically to proceed north of the 38th parallel."[33] Much has been made of this particular memo: Acheson even claimed that while it was within JCS guidelines, it was nevertheless misinterpreted by MacArthur; it was supposed to be nothing more beyond the approval of MacArthur's plans (submitted on the 28th as required under his orders implementing NSC-81/1). Certainly, it was not *carte blanche* to invade North Korea.

This interpretation seems essentially correct. Rusk, for instance, was not sure whether MacArthur would have to make a last-minute check with Washington before UN forces crossed the parallel. This would only be correct if MacArthur was following the NSC-81/1 orders and was going north for military purposes.[34] MacArthur, in his usual manner, declared all Korea open for war on the 30th. Warren Austin chimed in by calling the 38th parallel an "artificial barrier." Zhou threatened that the PRC would not "supinely tolerate seeing their neighbors being savagely invaded by imperialists."[35]

6.3.3 The Closing of the Window of Opportunity

On October 1, as MacArthur was calling on the North Koreans to surrender, ROK units began crossing the 38th parallel. Although MacArthur would not announce this for two days, the Chinese learned about it immediately.[36] Kim sent desperate appeals for help to the Soviet Union and the PRC. Stalin urged Mao to move 5 or 6 divisions south but would not commit himself. Mao called a session of the Politburo to debate intervention.[37]

It was on the 2nd that Mao supposedly sent Stalin the famous telegram purporting to show a firm resolve to intervene in Korea. Many studies rely completely on this as *prima facie* evidence of Chinese intent to enter the war, which in turn substantiates the view that because they were serious about their intent, the Chinese had communicated it as best as

Panikkar's arguments "did not reflect much credit on Panikkar's reasoning ability." See Henderson's telegram to Acheson, October 5, 1950, FRUS: 876.

31 Telegram from Acheson to Web, September 28, 1950, FRUS: 797, and Telegram from Henderson to Acheson, September 28, 1950, FRUS: 809.

32 Telegram from Kirk to Acheson, September 29, FRUS: 822.

33 Telegram from Marshall to MacArthur, September 29, 1950, FRUS: 826.

34 Memorandum of Conversations, October 4, 1950, FRUS: 862. See also Stueck (1998, 92), and Kaufman (1997, 56,103).

35 Telegram from Wilkinson to Acheson, October 2, 1950, FRUS: 852. See also Offner (2002, 390–92), Goncharov et al. (1993, 175), and Whiting (1960, 111).

36 Appleman (1961, 615), Zhang (1995, 77).

37 Offner (2002, 389), Goncharov et al. (1993, 176–80), Stueck (1998, 89).

they could to the Americans, and the latter wilfully ignored it.[38] We shall have more to say about this interpretation later on in this chapter. Right now it will suffice to say that this telegram appears never to have been sent to the Soviets! The only copy is from the Chinese archives, and the version in the Russian archives (a report by the Soviet Ambassador in Beijing Roshchin that surfaced recently) is drastically different. According to Roshchin, Mao had verbally instructed him to tell Stalin that *no Chinese intervention was forthcoming* because the Politburo was too divided.[39] This version makes the subsequent events–from the failure to threaten openly to the last-ditch effort to secure Soviet support—much more intelligible.

On the 3rd, Zhou told Panikkar that China would intervene if American troops cross the 38th parallel but not if only ROK units do so.[40] This was indeed convenient: the Chinese already knew that ROK units had gone north, so the exception could clearly be seen as a way out of the commitment (anything else would have required immediate intervention) and an attempt to bluff the US into staying south. More conspiracy-minded interpretations have Mao conditioning Chinese entry on an event that had not yet occurred in order to conceal his decision to intervene anyway. This is an odd logic because by its own admission, if the US at this point opted to say south, the Chinese would have been deprived of the pretext under which they were supposedly planning their inevitable entry. It appears more likely that Zhou was telling Panikkar what the true state of affairs in the Politburo allowed him to. Namely, China could not yet definitely commit to intervention—which meant it had to acquiesce to the *fait accompli* of a ROK crossing—but that it may still resist the unification of North Korea by the Americans. The fact that he chose not to communicate this message to the United Nations—where the Chinese had gone before when they protested about the bombings in Manchuria—also undermined the immediacy of the warning, as both Kirk and Webb argued.[41]

Nevertheless, the US made a frantic effort to obtain some corroboration of Zhou's representations to Panikkar.[42] On the 4th, Clubb argued that Zhou's demarche could not be "safely regarded as bluff" and urged the administration to ascertain the Chinese moves. Indeed, Webb himself said that the question is not "whether Chi Commie intend to intervene in Kor[ean] conflict, but only of degree of their intervention."[43] However, in his talk with the British, Acheson articulated the infamous "poker game" analogy which many have seized upon as *the* evidence that the Americans were not taking the Chinese seriously. As reported by Allison, while Acheson

[38] Alternatively, it was a ruse by the Chinese who had already decided on war with the United States, as Zhang (1995, 80) and Chen (1992, 18) argue.

[39] The draft telegram can be found in Goncharov et al. (1993, 275–76). The Russian copy as Telegram No. 25199 from Roshchin to Stalin conveying Mao's October 2 message to Stalin, October 3, 1950, can be found as Document 12 in Mansourov (1995). Shen (1996) also discusses the discrepancy between the two versions.

[40] Telegram from Holmes to Acheson, October 3, 1950, FRUS: 839.

[41] Telegram from Kirk to Acheson, October 3, 1950, FRUS: 850. Telegram from Webb to the Embassy in India, October 4, 1950, FRUS: 874. See also Acheson (1987, 452), Offner (2002, 390), and Appleman (1961, 757–58).

[42] Circular telegram from Webb to Certain Diplomatic and Consular Offices, October 5, 1950, FRUS: 877.

[43] Memorandum by Clubb to Merchant, October 4, 1950, FRUS: 864–66. Telegram from Webb to the Embassy in India, October 4, 1950, FRUS: 874.

> agreed that there was a risk in going ahead in view of the Chinese Communists position as conveyed to the Indian Ambassador in Peiping,… the Chinese Communists were themselves taking no risk in as much as their private talks to the Indian Ambassador could be disavowed, that they had not made any statement directly to the United Nations or to the Unified Command and if they wanted to take part in the "poker game" they would have to put more on the table than they had up to the present.[44]

It is ironic that many scholars have blasted Acheson for this assessment for it was, in fact, correct. On the very same day, strong opposition to intervention surfaced in the Politburo. Lin Biao, the commander of the crack 4th Field Army, which was to bear the brunt of the fighting, declined command of the Chinese Forces in Korea, "fearful of this task," according to PLA's acting General Chief of Staff Nie Rongzhen. It would not be until the following day that Mao was to prevail at the Politburo on intervention. The Chinese signals had been mixed for the very good reason that they themselves had not decided what to do, even in principle.[45]

While the Chinese were making up their mind, time was slipping by. Henderson reported that Indian newspapers carried stories about 'neutral authority" in Peking saying that major conflict in Korea was now 'almost inevitable' and that when UNF crosses the parallel, it would certainly clash with Chinese troops. In his estimate, however, this was calculated to "contribute to war of nerves over Chinese intervention in Korea." The Belgians and the Swiss joined the British in their assessment that the PRC threats were not genuine, or at the very least did not indicate an intent to start a major war over Korea.[46]

On the 7th, UN Resolution 376 authorized the unification of Korea and US forces crossed the 38th. The Joint Chiefs informed Truman that neither NSC-81/1 nor the directive of September 27 actually tell MacArthur what to do should the CCF intervene without announcement. The President agreed and approved the new instructions amplifying the 9/27 directive. Sent to MacArthur on the 9th, the "amplifying" memo directed the General that should the PRC forces enter *anywhere* in Korea without prior announcement he was to "continue the action as long as, in [his] judgment, action by forces now under [his] control offers a reasonable chance of success."[47]

Although Acheson would later claim that MacArthur misinterpreted the UN Resolution as *carte blanche* for unification, the 10/9 instructions leave no doubt that he was to proceed with his invasion of North Korea even if he encountered substantial Chinese presence there.[48] The Joint Chiefs may have abdicated some responsibility for the campaign by letting MacArthur decide, on his own, whether it offered a "reasonable chance of success," but this was not inconsistent with letting the field commander exercise judgment that his superiors in Washington did not have enough up-to-date information to second-guess. It is doubtful that the JCS, which had expressed serious misgivings about the Inchon landing, would overrule the apparently invincible MacArthur.[49]

[44] Memorandum of Conversation, October 4, 1950, FRUS: 868.

[45] Zhang (1995, 80–81), Li et al. (2001, 42).

[46] Telegram from Henderson to Acheson, October 6, 1950, FRUS: 892. Telegram from Murphy to Acheson, October 6, 1950, FRUS: 901. Telegram from Vincent to Acheson, October 7, 1950, FRUS: 902.

[47] Draft directive to MacArthur submitted by the Department of Defense to the President for approval, October 7, 1950, FRUS: 911–12. Telegram from the Joint Chiefs of Staff to MacArthur, October 9, 1950, FRUS: 915.

[48] Acheson (1987, 455).

[49] Given MacArthur's well-known proclivities, authorizing him to proceed as long as there was some chance of

The bottom line, however, is clear: by early October the Americans had decided that they could risk war with the Chinese—provided they entered on their own without open Soviet support—in pursuit of Korean unification. The war aims had shifted to what would turn out to be their most extreme within a span of a few weeks. The United States had become undeterrable by China.

6.3.4 The Chinese Make Up Their Minds

The Chinese, in the meantime, had their own problems. Although Mao issued a mobilization directive on the day after the UN vote and told Kim Il Sung that China was entering the war, intervention was not immediate.[50] First, the CCF had invaded Tibet to reclaim central authority that had lapsed with the fall of the Qing dynasty. Second, and more importantly, Mao was bargaining with Stalin over the terms of PRC's entry in Korea. On the 7th, Roshchin cabled Stalin that Mao had said that China could not pay for the Soviet support (as originally agreed) and could not intervene without it. Mao was proposing to send Zhou to talk to Stalin.[51] In other words, on the day Mao was ordering mobilization, Zhou was on the plane to Moscow to plead with the Soviets for assistance.[52]

Although Chen (1992, 20) claims that Mao decided to send all CPV forces south of the Yalu on the 9th, Zhou told Stalin on the very next day that China would not enter without Soviet support. Startled by this revelation, Stalin initially agreed to help but on the 11th, he reneged. Molotov informed Zhou that the Soviet Union no longer supported intervention! Upon receiving this distressing news from Moscow, Mao informed Gao Gang that there would be Soviet air cover and withdrew for 72 hours to ponder his options.[53]

The Chinese leadership knew that it could not compel the US to leave North Korea without war but was loath to enter into the fray without Soviet assistance. On the 12th, the Politburo postponed the 10/9 order to the CPV to cross the Yalu. Stalin told Kim that the Chinese had reneged on their promise to help him and blandly advised him to evacuate.[54]

This is not the place to trace the devious and exceedingly intricate maneuvering among Mao, Stalin, and Kim with regard to the Korean War. Even then, Soviet intent may not be difficult to grasp in aggregate, even as it remains infuriatingly elusive in detail. Stalin seems to have feared that too large an involvement in Korea might embroil the USSR into a

success was tantamount to authorizing him to wage war on China even if the administration could not bring itself to admit that. Acheson (1987, 447) faults the JCS for being too timid. But they had every reason to: Inchon was not the first exceedingly risky operation that MacArthur had pulled off despite the strenuous objections his superiors had expressed about its execution. The 1944 Los Negros campaign was eerily similar in that respect. It is only failure that distinguishes an incredibly stupid and obviously doomed operation from a bold and brilliant one, and thus far MacArthur had performed boldly and brilliantly (Manchester, 1978, 340–44).

50 Goncharov et al. (1993, 184-90), Zhang (1995, 82), Chen (1992, 19–20).

51 Telegram No. 25348 from Roshchin to Stalin, October 7, 1950. The translated version can be found in Hershberg (2004, 377–78).

52 Zhang (1995, 83).

53 Zhang (1995, 83), Chen (1992, 20), Goncharov et al. (1993, 190,280), and Hao and Zhai (1990, 111). Stalin was probably worried that the US could bomb China, which would activate the Sino-Soviet mutual defense treaty and drag the Soviet Union into a war with the United States.

54 Chen (1992, 21), Stueck (1998, 89), Zhang (1995, 83), and Hao and Zhai (1990, 110).

general war with America (Pollack, 1989, 224). This was something he clearly did not want, so he directed Kim to consult with Mao when the Korean leader visited Moscow secretly between March 30 and April 25. Of course, Stalin had also promised Mao support for the invasion of Taiwan, so perhaps he expected that Mao would not want to share the limited resources (Goncharov et al., 1993, 146). At any rate, Stalin did not want to refuse Kim, so he pushed the Korean to Mao, hoping perhaps that this would do the trick. It did not, of course, because Kim was nobody's pawn. The Korean War was a civil war, waged for unification by Koreans on both sides of the 38th parallel, and not, as it is often portrayed, a superpower confrontation by proxy. While Kim relied on Soviet and Chinese assistance, he was no more controlled by Moscow or Beijing than Rhee was controlled from Washington.[55]

Once the war erupted, Stalin seems to have tried to ensure that the provisions of the Sino-Soviet treaty would not be activated. Since Article I obliged the Soviet Union to "immediately render military and other assistance with all means at its disposal" should China find itself at war with the US, this meant pursuing policies that would minimize the chances of that happening. Hence the otherwise puzzling failure to veto the UN resolutions branding North Korea as the aggressor: if US forces were under United Nations command, they would be less likely to declare war on China.[56] Unlike Mao, Stalin seems to have been prepared to accept the collapse on North Korea. As he remarked,

> Let the United States of America be our neighbors in the Far East. They will come there, but we shall not fight them now. We are not ready to fight.[57]

He must have understood that Mao could not let his neighbor disintegrate, so he moved toward a policy that would ensure China's entry in a manner that would be least likely to provoke a declaration of war by the United States. Since he did not want to precipitate a similar declaration by the Chinese, he could not have been too encouraging with his promises for aid.

A curious incident occurred during Zhou Enlai's visit with Stalin on October 10. Zhou, quite disingenuously, claimed that China was unable to enter the war without Soviet support. Stalin, who must have been informed about the preparations underway, called his bluff and explained the reasons why, while the Soviet Union could not intervene, China had to. He washed his hands of the decision, and when Zhou finally abandoned all pretense and asked for air support, Stalin agreed but placed limits on the operations of the Soviet pilots (Goncharov et al., 1993, 188-90). It was this supposedly final promise that Molotov repudiated on

[55] Bajanov (1995) claims that Stalin was firmly in control, setting even the date for the invasion, and that he pressured the reluctant Chinese into entering the war to save the Pyongyang regime. One evidence for this is Mao's October 2 telegram as recorded in the Russian archives, but, as we have already seen, things were not that simple. It does not appear that Mao was unduly influenced by Stalin. China entered the war mostly for national security reasons. There is no reason to suspect that Kim was an ideological pawn in Moscow's hands either, especially given Stalin's professed aversion to a war in the far east, a war that could potentially drag the USSR into conflict with the United States.

[56] The US was indeed somewhat restrained by the need to clear activities with the UN. That is why Marshall asked MacArthur to cross the 38th as a matter of military necessity, so as not to put the issue to a vote. Not much of a restraint, true, but not quite free either. The British were quite active trying to bring a negotiated end to the war, especially during the lull in fighting after the first encounter with the Chinese forces. Neither side was interested (Farrar, 1983).

[57] Cited in Goncharov et al. (1993, 191).

the following day that caused Mao's seventy-hour retreat into contemplation. When Soviet military assistance finally came, the Chinese even had to pay for it.

All this means that as late as October 12, three days after the Americans had resolved to risk Chinese intervention to unify Korea, the PRC leadership was still vacillating. No wonder all previous Chinese threats had a strong element of bluff to them! On that very day, the CIA estimated that overt PRC intervention was unlikely in 1950 despite Zhou's statements, troop movements in Manchuria, and newspaper propaganda.[58]

Mao emerged from his isolation on October 13 determined that the PRC would intervene even if Soviet support was not forthcoming. However, the scope of intervention was to be limited to holding the UNF away from the Yalu—fight should the UNF attempt to force the perimeter but prepare for a counteroffensive otherwise.[59] Zhou's new report from Moscow saying that Stalin had now firmly committed to support intervention provided the PRC entered first stiffened the backs of the Politburo members.[60] On the 15th, Mao cabled Gao and Peng that CPV advance units were to cross the Yalu no later than the 17th, and ordered the cessation of the large-scale Chinese offensive in Indochina.[61] This was the day on which MacArthur reassured Truman (who had flown to visit him on Wake Island) that the Chinese were unlikely to intervene and if they did try anyway, "there would be the greatest slaughter".[62] The Chinese began crossing—without Soviet air support—on the next day.[63]

Although the US persisted in the attempts to ascertain Chinese intentions and despite mounting evidence that the PRC had intervened on a massive scale, the fierce counter-attack with which they met MacArthur's November 25th offensive surprised the General. After his dramatic declaration that the United States now faced "an entirely new war," the administration reassessed its expectations about the conflict.[64] In the end, despite the unseemly hasty defection of its allies in the UN but also after ascertaining that the Soviet Union was not going to intervene, the US resolved to stay in Korea as long as possible. In mid December Truman declared a state of national emergency and put the economy on wartime footing.[65] It was in this way that the United States in China ended up fighting each other in Korea, a conflict that neither country had wanted but one they had failed to avoid.

[58] Memorandum by the Central Intelligence Agency, October 12, 1950, FRUS: 933–34.

[59] Chen (1992, 21–22), Christensen (1992, 135), Zhang (1995, 83–84).

[60] Stueck (1998, 89).

[61] Chen (1992, 22), Zhang (1995, 70).

[62] Substance of Statements Made at Wake Island Conference on October 15, 1950, FRUS: 953. It should be noted that when pressed by Rusk about the possibility of an open Chinese entry, MacArthur gave a guarded response that since the PRC was unlikely to do it without Soviet support, such an eventuality should be treated with utmost seriousness. Addendum to Notes on Wake Conference on October 14 by Rusk, undated, FRUS: 962).

[63] Chen (1992, 22-23). The Russians did transfer 231 planes to the PLA soon after Zhou left Moscow, and continued helping the Chinese from then on (Zhang, 1995, 84). The hasty crossing, which anticipated Mao's formal orders that were to come on the 18th, was probably meant to preempt the possible bombing of bridges by the UNF. See Chen (1992, 23) and Zhang (1995, 93).

[64] Telegram from MacArthur to the Joint Chiefs of Staff, November 28, 1950, FRUS: 1237.

[65] Kaufman (1997, 72).

6.4 Did the US Irrationally Dismiss Clear Chinese Threats?

Now that we have seen the changing war aims of the Americans and the tentative behavior of the Chinese, we need to ask whether a different logic could account of that pattern of behavior. The most common argument is that the American policy-makers made a mistake by underestimating Mao's resolve and his determination to intervene in the conflict to prevent the unification of Korea under the tutelage of the United States. That is, the Chinese were not bluffing but, for various reasons, the US administration completely misinterpreted their behavior. The debate is usually over the causes of this miscalculation.[66] The best summary is provided by Lebow (1981, 149-52):

> Despite differences of opinion about the causes of [the American] miscalculation [of Chinese intent], scholars are nearly unanimous in their view that it was not the result of a simple intelligence failure. In retrospect, it is apparent that there was ample evidence as to both Chinese capabilities and intentions. . . .
>
> MacArthur and his apologists aside, none of the analysts have sought to justify America's miscalculation of Chinese resolve as a reasonably drawn if incorrect assessment.
>
> . . . despite their methodological and interpretative differences these scholars start from the same fundamental premise: that American leaders were remarkably insensitive to Chinese warnings. They differ only in their explanations as to why this was so.

According to this view, the Chinese made clear threats and the Americans were adequately informed about their military capabilities, and yet the US chose to ignore the warnings. Among the reasons given for this remarkable imperviousness to apparently obvious facts are, in no particular order, (1) MacArthur's "closed mind," his manipulation of intelligence and his hubris; (2) the bureaucratic decision-making process itself; (3) the flaws in Chinese signalling (indirect communication and secrecy of their military moves); (4) the domestic political situation; (5) the desire to reestablish American prestige–shaken by the 1949 triumph of the Chinese communists—in the far east; (6) the failure to understand Marxist ideology; (7) the relative weakness of China and its repeated bluffs over Taiwan.[67]

While acknowledging the flaws in signaling inherent in using the Indian communication channel regarded as unreliable by the Americans or hiding the military preparations, George and Smoke (1974, 189–91) nevertheless conclude that "It would be a serious error. . . to fasten upon these flaws in signaling as decisive or critical factors for explaining either or both deterrence failures in this case.. . . For scholars to attribute the failure of Chinese warnings to achieve credibility to Peking's lack of skill in signaling is superficial and misleading."

It is indicative of the weakness of this argument that it requires us to believe that almost everyone in the US government was blind to what was apparently so obvious. As George and Smoke (1974, 208) themselves admit, "the miscalculation of Peking's intentions was by no means confined to top-level policy-makers. It also characterized the most careful and responsible estimates of professional intelligence specialists." Rather than giving the authors pause—is it reasonable to suppose that nobody could see the clear Chinese signals?—the conclusion is that the Americans simply did not understand how threatening to the Chinese their behavior was (George and Smoke, 1974, 213); that or else MacArthur had misled everyone (Lebow, 1981).[68] This strains credulity too: as Appleman (1961, 757)

[66] Lebow (1981), George and Smoke (1974, 184), and Sartori (2005), among others, advance this view.

[67] See Lebow (1981, 150–53) for a catalogue.

[68] How dare the Americans "ominously" reverse their policy on Taiwan? How dare they omit Taiwan from Acheson's defense perimeter speech and then send a fleet into the straight to "neutralize" it? How dare they

has documented, there were quite a few incorrect estimates of Chinese strength, all similar to MacArthur's. This does not even touch upon the fact the the American allies (and not just the British) essentially all reported the same thing. Is it all "psychological influences and misperceptions that distorted the views" of both intelligence specialists and policy-makers? Or perhaps there was no clear evidence? What is more reasonable and likely? But we need not ask this hypothetical, we have evidence that can answer the following question quite decisively.

Were the Chinese signals unambiguous? There is a chorus of scholars that says so:

> The Chinese gave ample evidence of both their intention and capability to intervene (Lebow, 1981, 157).
>
> [An] unusual consistency and lack of "noise" characterized Peking's efforts to signal its intentions (George and Smoke, 1974, 188).
>
> China's many and varied threats were clear in meaning. US leaders understood the warnings and knew that China was capable of intervening (Sartori, 2005, 20).

But were the Chinese signals really unambiguous? They were not. Indeed, it would have been rather strange if they had been given that the Chinese had not seriously faced the possibility of entering the war until September and Mao vacillated nearly until mid October. China would not be fully committed to the war until after MacArthur's "home by Christmas" offensive in late November. As Whiting (1960, 118) put it, "While China crossed the Yalu on October 15th, she did not cross the Rubicon until November 26th." In his recent study of the war, Stueck (1998, 106) similarly concludes that "[Mao's] movement toward a final decision did not even begin until early October, and the process was anything but linear. It is not surprising, therefore, that an explicit warning to the Americans came over two weeks after Inchon."

We can make our own assessment of these competing claims. There are two sets of arguments that support the case for ambiguity in Chinese behavior. First, we now know that Mao was not keen on entering the war, which explains the caution and tentativeness of his moves during the critical weeks after the Inchon landing. The Chinese hoped for a negotiated solution and could not risk an open commitment that would badly backfire if the United States itself turned out to be prepared for general war over Korea. In other words, the Chinese *would not have wanted to send clear threats* because they did not want to risk escalation without being sure they would be prepared to face the consequences. Second, evidence available at the time strongly suggested the Chinese were bluffing. From the concealed troop movements that led to grotesque underestimation of CCF in Korea, to using Panikkar (known to be suspected in the West of pro-Communist sympathies) or public radio broadcasts in Beijing (which could be dismissed as propaganda ploys), from failure to intervene when the military situation made success most likely to the widely known lack of air support, all of these factors added a rather strong support of Acheson's famous "poker game" dismissal of the Chinese threats as bluffs. In other words, *it was precisely because the Chinese did not want to send clear signals that they engaged in behavior that was ambiguous.*

> intervene in the "civil war"? Nowhere in this line of pretend-outrage is the simple fact that it was not the US that initiated the war in Korea but China's own friend North Korea. While one may quibble with actions that are ominous and threatening because they may be a sign of evil things to come, it is surely a stretch to weigh these more heavily than actions that implement said evil things.

It would be useful to begin by enumerating the reasons for intervention. The Chinese could not let the US reach the Yalu and destroy a friendly communist regime. This would mean a serious setback for communism's international prestige and would compromise the credibility of communist promises to assist others in the global revolution. It would mean the loss of a strategic buffer in North Korea and the emergence of a permanent hostile presence right at the border. This would make it necessary to maintain indefinitely a substantial force to deter further aggression against China. This would be ruinously expensive—it would strain the economy, jeopardize the major industrial centers in the Northeast, and give the enemy control of electric power in south Manchuria. American reassurances that there were no grand plan for further expansion into East Asia were not credible. Truman had promised to stay out of Chinese affairs over Taiwan but had reneged on his January pledge almost immediately once war in Korea began. The Korean problem was not isolated, and the defense of the Asian conversion to communism was the only way to defend China itself. Fighting the imperialists in Korea was more advantageous than fighting them in Vietnam or Taiwan.[69]

Furthermore, the Party feared losing popular support. Revealing fear by failure to intervene would encourage the counter-revolution. Disturbances were already breaking out in anticipation of an American invasion across the Yalu. Although Washington had disavowed MacArthur's heavily publicized visit to Chiang Kai-shek in Taiwan in July and had forced the General to issue a written repudiation, the Americans were not to be trusted. They were apparently siding with Chiang—what else would the Seventh Fleet be doing in the straits?[70] But if the Americans were coming across the Yalu anyway, then attacking them in Korea would save China itself from aggression. Seizing the initiative and exploiting the advantages of topography and short logistical lines were also in line with Mao's tactics.[71] Although inferior in weapons, the PRC was superior to the US in manpower, moral strength, and the support of the people. Even assuming that the Americans would bomb the mainland itself (the US was believed unlikely to use nuclear weapons), China had to fight because the US would not negotiate—on Taiwan, Korea, or even the UN seat—without suffering a serious setback first.[72]

In short, there were many reasons to intervene in Korea.[73] On the other hand, there were many good reasons to hesitate. Initially, Mao had not assigned much priority to Korea at all, even after the war had broken out. Instead, he had concentrated on the impending invasion of Taiwan.[74] The economy, ravaged by years of civil war, was in shambles, and the army had released nearly half its manpower back in April when the initial stages of the conversion to

[69] See Chen (1992, 18–40), Zhang (1995, 56), Hao and Zhai (1990, 104–106), Christensen (1992, 137–45), Goncharov et al. (1993, 184), Whiting (1960, 152), Roe (2000, 89).

[70] Wang (1989, 201), Gaddis (1989, 163), Goncharov et al. (1993, 159), Chen (1989, 189–91).

[71] Pollack (1989, 222), Chen (1989, 189–91).

[72] Sheng (1995), Hao and Zhai (1990, 108), Chen (1992, 20), Christensen (1992, 145).

[73] Some have stretched this argument too far. For instance, Chen (1994, 40) emphasizes Mao's revolutionary motivations and argues that these were more important than traditional security concerns. The notion that Mao was spoiling for a fight with the United States seems far-fetched to me.

[74] Mao sanctioned the "liberation" of Taiwan on March 11and on April 16 the PLA attacked Hainan Island and defeated the Nationalists in two weeks. It was units from this, Fourth Field Army, that were moved northeast. By early June Mao had decided to postpone the invasion until the summer of 1951 because of the slowness of the mobilization effort and the strains placed on the ravaged economy. Some skirmishing was to continue, however, and as late as June 12, the Chinese leadership authorized assaults on several offshore islands. The

civilian economy were getting under way. There was war weariness, internal banditry, and "unliberated" islands where resistance was still irksome. Most of the remaining troops were concentrated in the South, and were unprepared for winter operations. A ruinous war with America would certainly delay the reconstruction plans and cause great consternation to the people. Going to war with the United States in Korea would further divert the Party from its cherished goals of liberating Taiwan and assuming control of Tibet. Finally, given the precarious economic and military situation, it would likely make China even more dependent on Soviet aid, thereby exposing it to Moscow's influence and meddling. This war was not welcome, as the tumultuous Politburo session on October 1 amply demonstrates.[75]

The result? Ambiguous behavior until a decision to fight is made, then a surprise attack to maximize the chances of military victory. Making clear threats had serious disadvantages. If the PRC committed to fighting and the US ignored the threat, China had to attack openly, causing a war with the United States. Stalin—who was very apprehensive about being trapped into a war with the US by his defense treaty with China—could use this as an excuse to abrogate the treaty. If, on the other hand, the Chinese position remained ambivalent, the Soviets could be persuaded to help, at least covertly. Until late September, the Chinese probably also hoped to get Taiwan's seat in the UN, and making threats would have been counterproductive for that diplomatic effort.

It is worth noting that the Chinese had no problem going to the UN to voice their complaints about US incursions into Manchuria. It is, therefore, not surprising that their choice of diplomatic channel to convey the threats was interpreted, and rightly so, as allowing plausible deniability. Panikkar was entirely too sympathetic to the Chinese Communists.[76] So much was public knowledge. Using him to convey a supposedly unshakable resolve would be borderline idiotic unless one had no other options, which was clearly not the case (Zhang, 1995, 77). The Chinese even rejected a feeler for a direct contact the Americans put through the Indian government.[77]

In addition to their strange choice of Panikkar, the Chinese delivered oddly equivocal threats. For instance, why was it acceptable for ROK troops to cross the parallel but not for the Americans? Zhang (1995, 80) claims this was a ruse but his argument assumes that Mao had decided to intervene for sure at that point, which as we have already seen, was not the case. Since the threat was made *after* the ROK had crossed, the straightforward interpretation is that Mao wanted to keep his options open. A clear threat to intervene if the UNF moved north of the parallel when some of its units already had would have put the Chinese into the unenviable position to execute the threat immediately since the condition for its execution had already been met. But if one has not resolved to fight yet, then one better make a threat whose *if* had not yet occurred. There would be a political benefit if the threat worked and one could still back out if it failed. The need to disavow any intent to wage large-scale war would also explain why there was no talk beyond "volunteers."[78]

invasion plans were put on indefinite hold only on August 11, when Mao finally realized the Korean War would not be over soon (Goncharov et al., 1993, 157–58).

75 Harding and Yuan (1989, 189–214), Schaller (2002), Meisner (1999, 69), Goncharov et al. (1993, 176–80).

76 He had persuaded Nehru not to make trouble over PRC's invasion of Tibet on October 7 because an aggressive action against China would further weaken it just when developments in Korea threatened them gravely.

77 Telegram from Henderson to Acheson, October 10, 1950, FRUS: 921.

78 Telegram from Kirk to Acheson, November 14, 1950, FRUS: 1154.

The Chinese had not begun serious preparations for a long war with the United States despite their apparent conviction that any such war would be one of attrition.[79] Peking was not being secured against the air raids, there were no obvious military moves aside from troop movements in Manchuria. In fact, if the Chinese were serious about threatening the UNF, they would have either done so earlier or would have openly entered North Korea to demonstrate their resolve. For example, if they declared an intent to defend North Korea right after the Inchon landing, their chances of deterring the US would have been nearly certain. MacArthur himself noted that during the Wake Island conference with Truman. This was also the Department of State's assessment.[80]

Instead of making a show of force to buttress their supposedly credible threats, the Chinese took great care to conceal their preparations, even after they had decided to intervene. They maintained complete camouflage during the day, with standing orders for officers to shoot any stragglers that might reveal their presence to US reconnaissance overflights, and only moved under the cover of darkness.[81] Schelling (1966, 55) was puzzled by this secrecy and noted that while it doubtless gave the Chinese "stunning tactical advantages," it did so "at the expense of all deterrence and diplomacy." Achieving tactical surprise could only be a goal when one is resolved to fight and when one's opponent does not suspect an intent to do so. Making clear threats would obviously undermine the prospects for such a surprise and would ruin the Chinese strategy.[82] In other words, secrecy was necessary at first because the Chinese were not committed to fighting, and it was imperative for any threats to be capable of disavowal. It was also necessary after they had committed, but this time it was in order to achieve a military advantages. To claim that the Chinese had sent clear signals when both political and military factors strongly argued against such clarity, is absurd.

It should, therefore, come as not surprise that it was not just the Americans that thought the Chinese were either bluffing or, if they were not, their intervention would be on a very limited scale. So did the British, the Belgians, the Swiss, and even the Burmese.[83] The French, the Danes, and the Swedes had not information that could contradict that assessment, not information that did not originate with Panikkar anyway.[84]

It was not irrational for the US administration to believe the Chinese were bluffing for they were doing precisely this before mid October and were actively engaged in deception

[79] Telegram from Wilkinson to Acheson, October 7, 1950, FRUS: 912.

[80] Deputy Under Secretary of State (Matthews) to the Special Assistant to the Secretary of Defense for Foreign Military Affairs and Assistance (Burns, October 19, 1950, FRUS: 980.

[81] Appleman (1961, 753), Whiting (1960, 117–18), Zhang (1995, 93–94), Hao and Zhai (1990, 112–13).

[82] Revealing their preparations would expose the Chinese to US air strikes they could not counter without Soviet support which, as we have seen, was not exactly forthcoming. As Van Evera (1999, 61) lucidly explains, "China could not persuade Truman that he would pay these costs [of unifying Korea] without making itself unable to inflict them."

[83] Telegram from the Ambassador in Belgium (Murphy) to Acheson, October 6, 1950, FRUS: 901; Telegram from the minister in Switzerland (Vincent) to Acheson, October 7, 1950, FRUS: 902; Telegram from Wilkinson to Acheson, October 7, 1950, FRUS: 912; Telegram from the Ambassador in Burma (Key) to Acheson, October 14, 1950, FRUS: 944.

[84] Telegram from the Ambassador in France (Bruce) to Acheson, October 6, 1950, FRUS: 900; Telegram from the Ambassador in Denmark (Anderson) to Acheson, October 6, 1950, FRUS: 891; Telegram from the Ambassador in Sweden (Butterworth) to Acheson, October 7, 1950, FRUS: 906. The Norwegians confirmed the original report without the reference to allowing ROK troops to cross. Telegram from the Chargé in Norway (Snow) to Acheson, October 7, 1950, FRUS: 903.

after that. It is unnecessary to engage in mental gymnastics (e.g., "defensive avoidance") to explain why Truman, Acheson, Bradley, Marshall, MacArthur—just to name a few—thought so little of the prospect of Chinese intervention. It is difficult to argue that the US was not prepared to risk war with China over Korea. The problem was to avoid a larger war with China in China itself (which would involve the US too far east for Europe-firsters like Acheson), and to discern just what the Soviet role was in all of this. The dominant fear, as the documents clearly show, was that the Soviet Union would intervene, precipitating a general war in the process. Once it became clear that this was not going to happen, the US policymakers dug in their heels and prepared to wage a long war of attrition in Korea against China. As long as the war remained localized, they could hope to extract a negotiated settlement.

6.5 Conclusion

Sartori (2005, 41) argues that "China's threats to enter the Korean War in the event that US or UN troops crossed into North Korea were clear," and that "US leaders understood them, but nevertheless incorrectly dismissed them as bluffs." This interpretation goes directly against the intuition offered by the MTM. Signals that the opponent disbelieves never cause war in this model. When war does occur, it is always after the uncertainty is resolved, when there is no residual doubt about the credibility of the threat. In the MTM, war occurs when the two actors lock themselves in a situation where war is inevitable.

The arguments that the US misperceived genuine Chinese threats as bluffs must show, at the very least, two things. First, they must demonstrate that the Chinese were, in fact, resolved. Second, they must further demonstrate that the Chinese signals were credible and would have been believed by a common-sense person that was not hampered by psychological, bureaucratic, and situational, biases as the US administration allegedly was.

I have gone to great lengths to establish that the Chinese leadership was not committed to fighting the United States over Korea, and would almost certainly not have done so had the US forces remained south of the 38th parallel.[85] That is, the Chinese had not decided they would fight the US until *after* the American forces crossed the parallel, which implies that the Chinese would have had some difficulty deterring the US from crossing. After all, doing so would require them to establish a credible commitment to war.

In itself, this should be sufficient to undermine the traditional argument because it is difficult to see just how the PRC could have credibly signalled resolve that it did not possess and would not acquire until mid October at the earliest. However, I have also tried to prove that the US administration did indeed make vigorous efforts to ascertain just what the Chi-

[85] Almost every historian agrees that the PRC would not have intervened had not the US forces crossed into North Korea. See, for instance, Stueck (1998), Foot (1991, 419), Hao and Zhai (1990, 113). Farrar (1983) even argues that a buffer zone closer to the 39th parallel would have been acceptable. There are some who disagree. Chen (1996), for instance, cites broader domestic and international concerns in arguing that there was no way to avoid China's entry. Zhang (1995) also brings in Mao's military optimism and his desire to punish the arrogant Americans to argue that they trumped security concerns. However, both Christensen (1992) and Chen (1994) rely on the October 2 telegram that Mao supposedly sent to Stalin as their main evidence of Chinese intent to enter *before* the crossing. However, as we now know, this telegram was never sent and Mao did not himself decide, with great difficulty I might add, until after the crossing.

nese were up to. Given that the Chinese themselves were unsure at the time, it should come as no surprise that these efforts did not yield a definitive answer. Under the circumstances and given the military situation on the ground, chancing a limited war with the PRC for the sake of Korean unification appeared an acceptable risk. The Chinese intervention resulted from the military and political momentum created by the success at Inchon on the US side, and Soviet support for it on the Chinese side. Consistent with the logic of the model, by mid October the two antagonists had committed themselves to war, and whatever residual uncertainty over Chinese intentions there may have been, it was not the cause of the war's expansion.

My goal is not to be an apologist for Truman, Acheson, Marshall or MacArthur. However, it is hard to see just where it was that they made the awful mistake that subsequent analysts have been quick to blame them for. The situation on the ground and the information they had at the time reinforced the impression that unification was possible without the risk of a major war with the Soviets. A risk of a limited war with China, something the administration was anxious to avoid until two weeks after the Inchon landing, was well-worth taking once the North Korean army collapsed and the Chinese had missed the opportunity to intervene decisively—or even threaten to—when it was most advantageous to do so. Mao, Zhou, and the rest of the Politburo could not be faulted for bad policies except insofar as they did not make up their minds in time to deter the Americans.

The merciless logic of militarized deterrence points to the unpleasant possibility that war between the US and the PRC could have averted as late as early October had the Chinese threatened to enter in late September. It further points to the readily understandable reasons why the same threat would no longer deter the Americans after that. This in turn explains why the Chinese chose not to make it. That these possibilities appear substantiated by the available evidence only makes the situation all the more tragic. By mid October 1950, war between the United States and the People's Republic of China—a war that neither one of them had wanted just two weeks prior—had become inevitable.

7

The Price of Peace and Military Threat Effectiveness

I have yet to see any problem, however complicated, which, when looked at in just the right way, did not become still more complicated.
Paul Anderson

After the December 13, 2001 terrorist attack on the Indian Parliament, the strained relations between India and Pakistan reached the breaking point. India accused Pakistan of supporting the terrorists and demanded that Pakistan apprehend the leaders of their organizations and stop supporting them financially. When Pakistan's forces went on high alert in response to these bellicose statements, India mobilized over half a million troops and deployed them in Kashmir and Punjab toward the border with Pakistan. At the height of the conflict, "more than 1 million troops [stood] toe-to-toe along the 1,800-mile India-Pakistan border."[1]

The two countries, both officially in possession of nuclear weapons since 1998, also traded nuclear threats during the tense standoff. However, after some limited clashes in May and June 2002, the crisis deescalated and in October the two antagonists began demobilizations that led to a ceasefire in 2003. The massive Indian mobilization coupled with the bellicose statements of many of its politicians calling out for blood, worried many observes at the time. There was great concern that should war erupt, it may escalate into a nuclear exchange, a scenario for which the Defense Intelligence Agency projected over 8 million dead in the strikes alone.[2]

Instead of leading to nuclear war, the crisis petered out and the relations between the two countries have improved considerably since. India's highly aggressive behavior during the crisis and the subsequent normalization stand in interesting contrast that becomes even more puzzling when we recall that both countries have nuclear capability. In this chapter, I suggest a possible explanation that accounts both for situational stability—that is, crises becoming less likely—and for very costly mobilizations in crises that do erupt and which involve risk of war. I show why the burden of peace can be substantial. This does not mean that these resources were wasted even in situations that get resolved short of war with the affirmation of the status quo. As we shall see, sometimes it is necessary to pay the price to keep the peace as is.

[1] Laura Bradford. "Path to War," *Time,* June 2, 2002. http://www.time.com/time/magazine/article/0,9171,257113,00.html, accessed April 20, 2008.

[2] See Sean Howard's "India and Pakistan Camped on Brink of War over Kashmir" in *Disarmament Diplomacy,* 65, July-August 2002, http://www.acronym.org.uk/dd/dd65/65nr01.htm, accessed April 20, 2008.

7.1 The Paradoxical Burden of Peace

7.1.1 Stability and the Costs of War

It is probably fair to say that if there is anything approaching consensus in the scholarly literature on war, it is that when war becomes more expensive for everyone involved, it should also become less likely. This argument is at the heart of balance of power theory, with the idea being that when military capabilities are relatively evenly distributed, any fighting will be indecisive, protracted, and costly for all participants. This is supposed to induce them to exercise restraint. This stabilizing influence of the costs of fighting can also be seen in the bargaining model of war where an increase of these costs reduces the risk that negotiations will break down in war (Powell, 1999, 111).

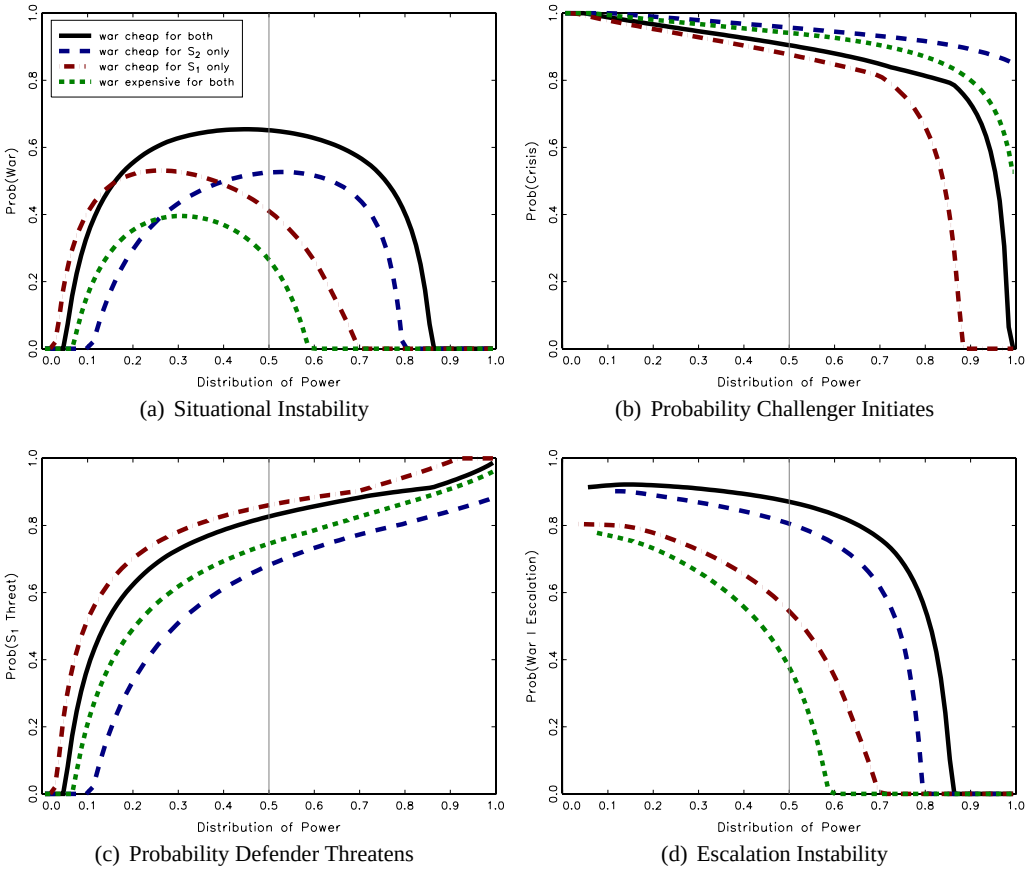
Many theorists treat war costs as a function of the military balance in a fashion similar to the balance of power theory. For instance, Huth et al. (1993, 612) write that "as the balance of military capabilities shifts toward the challenger, it becomes more likely that it will be able to prevail in an armed conflict. Additionally, under these circumstances the costs of armed conflict decline, increasing the net utility of a victory on the battlefield." In the MTM, the military balance does affect the expected payoff from war but it does so independently of the costs of war. Separating the two variables analytically allows us to investigate the effect of different military technologies, both in terms of their decisiveness (e.g., how they determine the probability of victory) and their capacity for destruction (e.g., how they affect the costs of fighting).

I certainly agree that costs should be a function of the duration of war and the tactics used (Leventoğlu and Slantchev, 2007). However, since military technology does influence costs as a separate component (e.g., aerial bombardment offers opportunities to inflict costs unavailable to actors lacking the capability), it will be useful to keep them distinct for analytical clarity. To investigate situational stability under varying assumptions about the costs of warfare, I assume general uncertainty about interests and consider the four scenarios we obtain by matching low and moderate cost players: (a) warfare expensive for both actors; (b) warfare expensive for the challenger but cheap for the defender; (c) warfare cheap for the challenger but expensive for the defender; (d) warfare cheap for both actors. Figure 7.1 shows four quantities of interest for each of these scenarios. The upshot of this exercise is the finding that whereas each of the constituent components behaves in an intuitive way, their combined effect has some surprising implications for situational stability that call into question the common wisdom.

Let us begin by observing that when it comes to comparing situations in which warfare is uniformly either expensive or cheap, then the intuitive relationship holds:

RESULT 7.1 *Situational stability is much worse when fighting is cheap for both actors than when it is expensive for both regardless of the distribution of power.*

This is essentially the same as the result in Morrow (1989b, 957). In Figure 7.1(a), the curve representing the unconditional probability of war under the assumption of universally expensive warfare is everywhere below the corresponding curve under the assumption of universally cheap warfare. The MTM therefore validates the notion that a military technology that affects all actors in the same manner should have a predictable effect on stability.



(a) Situational Instability (b) Probability Challenger Initiates

(c) Probability Defender Threatens (d) Escalation Instability

Figure 7.1 Stability and War Costs (general uncertainty).

This is where support for the intuitive relationship ends. Specifically, we can observe the following scenarios:

- warfare can become *more expensive* for one of the actors and, all else equal, situational stability can *decrease*;
- warfare can become *less expensive* for one of the actors and, all else equal, situational stability can *increase*.

To understand why these happen, let us parse the various components that together contribute to situation stability. Consider the first possibility, which can happen at DOP up to around 12%, and to make things concrete, take DOP of around 6%. The unconditional risk of war is 7% if warfare is cheap for both actors and more than three times that, 24%, if remains cheap for S_1 but becomes expensive for S_2.

At first glance, the figures do not provide a clue because all the dynamics there are in the intuitive direction. Specifically, the risk of war conditional on escalation is always higher when war is cheap for both than when it is only cheap for S_1. Since the defender is running a lower risk when the balance of costs favors him, he is more likely to escalate the

crisis, as seen in Figure 7.1(c). Finally, since this makes a costly war more likely from S_2's perspective, she is less likely to challenge the status quo.

Considered separately, each of these dynamics makes perfect sense, and yet their overall effect is to increase situational instability. The reason is that the improvement in stability resulting from higher escalation stability and lower likelihood of a challenge is dominated by the destabilizing effect of the rapidly increasing probability that the defender will escalate. In other words, the dramatic drop in the risk of war after escalation that arises from the balance of costs getting more favorable for the defender can make threats disproportionately attractive to him, and since the decrease in the probability that the challenger initiates is too small to offset that, the net effect is to increase the risk of war.

To see the second possibility, which can happen at DOP anywhere between 5% and 25%, take as example the DOP set at 13%. The unconditional risk of war is 24% if warfare is expensive for both actors and almost three times lower, 8%, if it remains expensive for S_1 but becomes cheap for S_2. Again, the component probabilities are all intuitive but their interaction produces the surprising result. Observe that the risk of war conditional on escalation is always lower when war is expensive for both actors than when it is only expensive for S_1, and that the difference is quite substantial. Making fighting cheaper for S_2 decreases escalation instability, as expected. This now leads to a higher probability of S_1 making threats when war is expensive for both: from his perspective war is just as costly as in the other scenario but escalation is a lot less risky. Hence, making fighting cheaper for S_2 actually makes escalation less likely precisely because it destabilizes it. Finally, this leads to S_2 challenging with a lower probability when war is expensive for both than when it is cheap to her. However, because fighting is not prohibitively costly, the decrease is pretty minimal. Overall, the stabilizing effect of a less likely escalation dominates the destabilizing effect of a (minimally) higher probability of a challenge and the greater escalation instability. The net effect is to lower the unconditional risk of war.

RESULT 7.2 *Although escalation stability always increases and the probability of challenge always decreases when warfare becomes more expensive for the challenger, situational stability may* decrease *because the defender is much more likely to escalate. Conversely, although escalation stability always decreases and the probability of challenge always increases when warfare becomes cheaper for the challenger, situational stability may* increase *because the defender is much less likely to escalate.*

This echoes Morrow's (1989b, 957) conclusion that "the effect of the costs of war on the probabilities of crises and war are complex." He also finds that although increased costs of fighting for the challenger make her less likely to initiate a crisis and more likely to back down when resisted, they also make the defender more likely to escalate. The key to all these results is the defender's propensity to escalate, and in particular whether the change in that propensity swamps the change in escalation stability. The interaction of these marginal effects produces the counter-intuitive result when DOP is relatively disadvantageous to the defender. Much of these dynamics rests on the expected costs of coercion for the defender, to which we now turn.

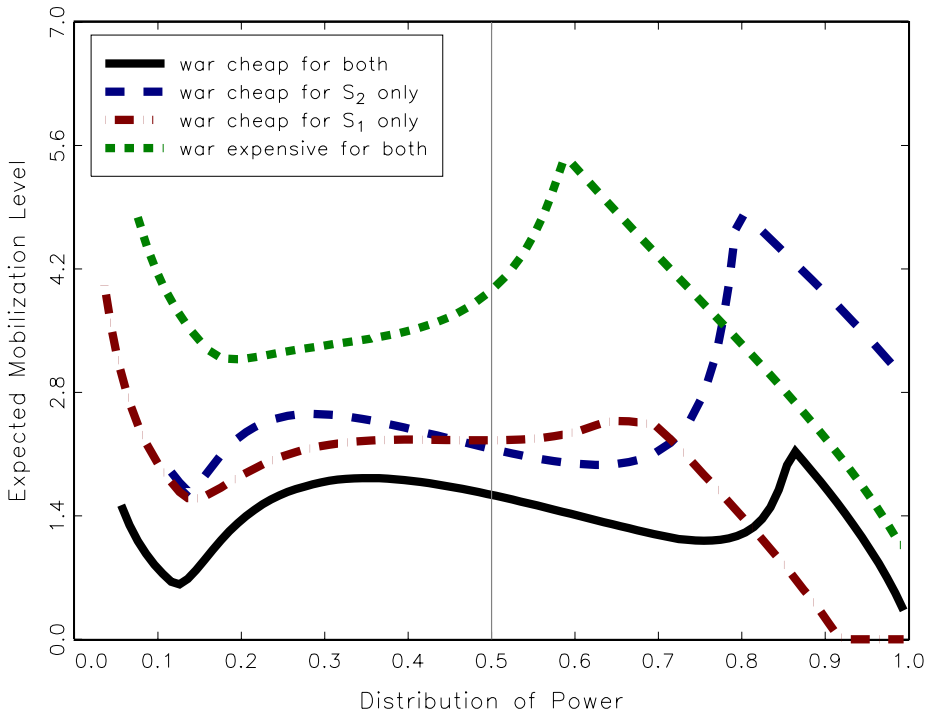


Figure 7.2 Defender's Mobilization and War Costs (general uncertainty).

7.1.2 High War Costs and Aggressive Mobilization

The costs of mobilization are crucial in determining the probability that the defender will escalate, and these costs depend on how easy to coerce his opponent is, and so, indirectly on the costs of war. Figure 7.2 shows the mobilization levels for the four scenarios we have been investigating. One thing immediately leaps to our attention: whenever the defender is willing to escalate, the more expensive the war, the higher his mobilization tends to be. In other words, S_1 is more aggressive in his threats precisely when fighting is expensive.

It is instructive to compare his behavior when war is very costly for both actors to the situation when it is costly only for him. Observe first that the expected mobilization level is always strictly greater in the universally costly war scenario than in the one where fighting is relatively cheap for the challenger for all distributions of power less than about 78% where S_1 escalates. In this range, for all DOP greater than about 25%, situational stability is also highest in the former case. In other words, for DOP between 25% and 78%, peace is simultaneously more likely and more costly when war is expensive for both actors.

Although this may sound counter-intuitive, the logic is simple: in both cases S_1 would like to avoid war because of its high costs; hence, his optimal allocations will tend to be coercive or compellent rather than war-fighting. When fighting is expensive for S_2 as well, she is relatively easier to compel, so S_1 is willing to pay to achieve compellence, which means that his mobilization is significantly higher but escalation instability is noticeably lower, as seen in Figure 7.1(d). When S_2's costs of fighting are lower, the defender is at

a disadvantage: it now will take a lot more effort to compel S_2 because war is relatively more attractive to her while it is just as unattractive to the defender. Since compellence is now much harder and more expensive, S_1 contends himself with a coercive strategy which saves a bit on mobilization costs but does expose him to the risk of a less stable escalation. This trade-off is especially noticeable in the upper regions of DOP where mobilization in the second scenario is higher: in both cases escalation stability is full, which means that S_1 achieves assured compellence but since S_2 is harder to compel when her war costs are low, the mobilization level is necessarily higher.

The analogous logic produces the other side of the coin: S_1's allocation will also tend to be higher when war is cheap for him only than when it is cheap for both actors. In other words: as fighting becomes more expensive for the challenger, the defender becomes more aggressive in his mobilization. We can already anticipate the reason: since S_2 is easier to coerce when her war costs are higher, doing so becomes a more attractive strategy for the defender, and consequently he is willing to pay more to make her capitulation more likely. Except in the region where the distribution of power is so skewed in S_2's favor that there is no deterrent effect (hence worse situational stability), this generally reduces the unconditional risk of war as well. Again, peace is both more likely and more expensive. As before, when DOP becomes so favorable for S_1 that he can achieve assured compellence, the mobilization level that he must pay for to do that will be higher when war is cheap for S_2.

Finally, note that this dynamic shows up without any qualifications in the comparison between the universally cheap war scenario and the one where fighting is cheap only for the challenger: both situational stability and mobilization levels are higher in the latter case. This suggests the following conclusion:

RESULT 7.3 *The* paradox of costly peace: *when war costs increase for one of the actors, situational stability will generally improve but the defender will use more aggressive mobilizations whenever escalation involves some risk of war; peace will be simultaneously more likely and more expensive.*

The war scare of 2001–2002 over Kashmir provides an interesting illustration of this result. With India alone moving half a million soldiers toward the border with Pakistan at a cost of 1.1 billion pounds, this standoff between India and Pakistan saw the largest massing of troops in the region since the 1965 war between the two states. Even the Kargil War that took place only a year after both countries had detonated nuclear devices in 1998 was much more circumscribed in scope with about 30,000 troops on the Indian side and 5,000 on the Pakistani.

Although Pakistan has consistently lost the military confrontations with India since the war over Bangladesh in 1971, its acquisition of nuclear capability has certainly made warfare potentially more expensive for India. The two countries maintain different attitudes toward the first use of nuclear weapons in a war. During the crisis, India's Foreign Minister Singh publicly committed his country to a no-first-use policy. When President Musharraf failed to reciprocate and insisted on Pakistan's right to resort to nuclear weapons first, India's Defense Minister Fernandes escalated the tensions by saying that whereas India could easily absorb a nuclear strike, Pakistan would "cease to exist."[3]

While the possession of nuclear weapons has certainly increased the costs of war for both

[3] *CNN*, January 8, 2003. http://www.cnn.com/2003/WORLD/asiapcf/south/01/08/pakistan. india/index.html. Accessed April 20, 2008. See also Sean Howard's "International Concern over Danger of Conflict in South

countries, given India's military superiority in conventional weapons, it may have affected them disproportionately in the sense that it reduced Pakistan's vulnerability to conventional attacks. In this situation, our results suggest that on one hand we should observe fewer crises but on the other these would tend to involve very aggressive mobilizations. The relations between the two countries have seen a marked improvement since the standoff, beginning with the 2003 ceasefire and the fencing of the Line of Control in Kashmir by the Indian Army. The model also suggests that despite all the bluster, the actual risk of war in this confrontation was not great.

If we assume that higher mobilization also makes war costlier to the opponent (in addition to its effect on probability of winning), then this also leads to a mutually reinforcing effect and should further improve situational stability because the coercive function of the military instrument will be even more pronounced, making higher allocations even more attractive. It is worth noting that whereas the conventional wisdom has generally captured the relationship between stability and the two polar scenarios of war being either expensive for both actors or cheap for both actors, the attendant corollary has been missed. Yes, the unconditional risk of war is lower when war is very costly for all than when it is cheap for all. But this peace will generally be dearly bought: the crisis mobilizations will be unusually aggressive precisely because they will tend to focus on avoiding war by achieving compellence. Given the enormous expenditures on defense during the last century, it is worth emphasizing that the strategic imperatives of military coercion may make peace a necessarily expensive state of affairs.

7.2 Do Audience Costs Improve the Prospects for Peace?

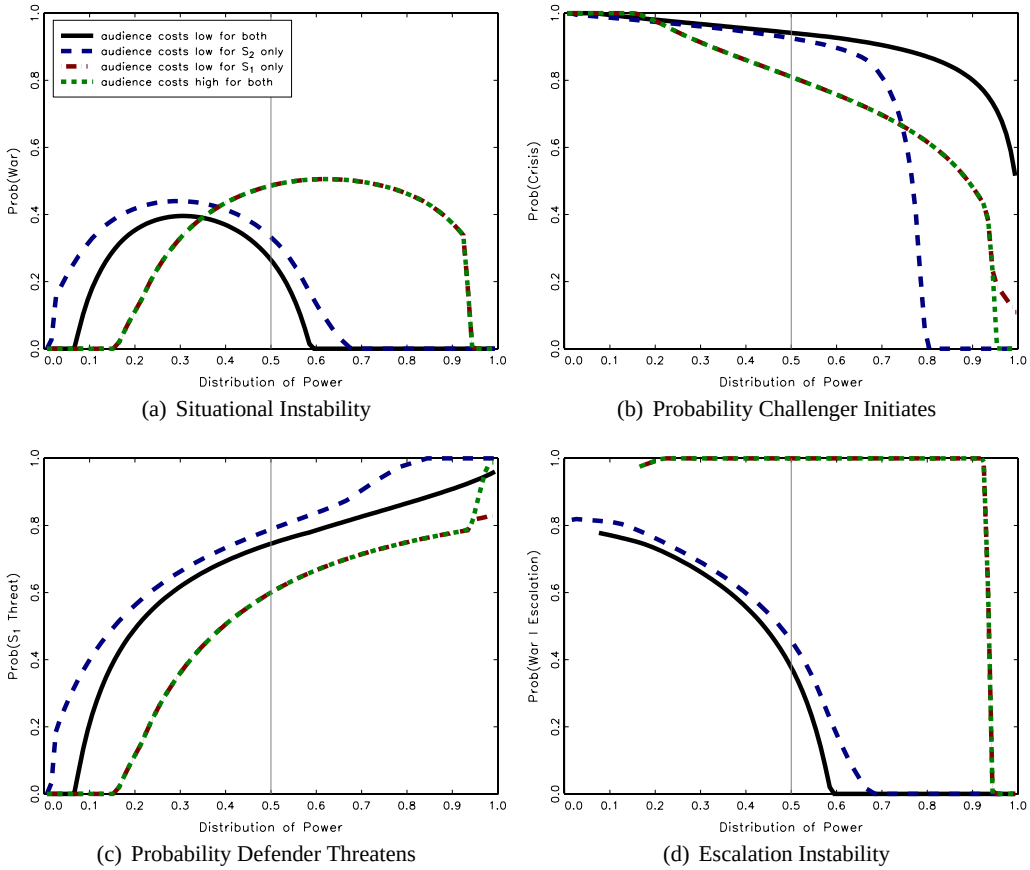
Recall from our discussion in Chapter 2 that audience costs are a major signaling mechanism that has gained prominence in explaining crisis dynamics. In that chapter, I also raised some concerns about its theoretical and substantive plausibility. In contrast, now I will take it as a given and will explore the deterrence model under varying assumptions about the magnitude of these costs. As we shall see, this will generate some surprising insights that run counter to the conventional wisdom that has developed from the original theoretical model.

For the simulations that follow, I assume general uncertainty and medium costs of war. As in the previous section, I examine four scenarios with audience costs high for both actors, low for both actors, and high for one actor but low for his opponent. Since we need to maintain Assumption 2.1, I set a_i at 10% of c_i to simulate low audience costs and at 90% of c_i to simulate high audience costs (that is, in latter case the costs of capitulating during a crisis are almost as high as the costs of war). In keeping with common practice in the literature, one can think of a high audience cost actor as a democracy, and a low audience cost actor as an autocracy.[4] Figure 7.3 shows the various measures of stability for these four scenarios.

The first thing to note is that the scenario in which audience costs are high for both actors is virtually indistinguishable from the one in which they are high for the challenger only.

Asia" in *Disarmament Diplomacy*, 62, January-February 2002, for a summary and some quotes about the risk of nuclear exchange. http://www.acronym.org.uk/dd/dd62/62nr01.htm, accessed April 20, 2008.

[4] The fact that this is common practice should not obscure the difficulties with this particular assumption, as I have noted in Slantchev (2006).



(a) Situational Instability (b) Probability Challenger Initiates

(c) Probability Defender Threatens (d) Escalation Instability

Figure 7.3 Stability and Audience Costs (general uncertainty, medium war costs).

That is, provided that S_2's costs of capitulation are high, varying the audience costs for S_1 has almost no impact on stability. Why is this so? When S_2's audience costs are high, she will be very difficult to compel and the MTM continuation game will involve war with a very high probability (that is, it will mostly involve the war-fighting equilibrium types). From our analysis in Chapter 3 we know that S_1's audience costs will only matter when the MTM equilibrium involves strategic pooling or bluffing because only in these two cases can there exist types who can potentially escalate without being resolved, which means that resistance must lead them to capitulation and incurring of these audience costs.[5]

[5] In the first case, we found the mobilization level that made $\underline{\tau}$ precisely indifferent capitulation and war, which ensured that his escalation would be credible. It is in finding the minimum mobilization level of that type that audience costs then play a role. In the second case, S_1 never expects to capitulate but his assured compellence mobilization must ensure S_2's capitulation despite her knowledge that she might be facing a bluffer. Here, she must evaluate the risk of war if she resists, which means she must estimate the likelihood that S_1 is unresolved at $\overline{m}$. It is in this calculation that S_1's audience costs come into play. In all other types of equilibrium play S_1's mobilization is credible to begin with, and his audience costs are irrelevant for crisis dynamics. Hence, when S_2's audience costs are so high that the MTM equilibria are not among these two, varying S_1's audience costs should have no impact on stability.

RESULT 7.4 *Because the defender's mobilization is credible—which means he never backs down once he mobilizes—stability (situational, crisis, and escalation) is determined by the audience costs of the challenger except when escalation involves strategic pooling or assured compellence.*

This in itself is a surprising result: the claim is that in just about every serious crisis, the defender's audience costs play a very minor or no role at all. The ability to use the military instrument has relegated the alternative commitment device to a secondary position. I should caution that this line of reasoning should not be pushed too far because the only reason S_1's audience costs do not feature in the calculus of stability as prominently as those of S_2 is that the model does not permit S_2's coercive mobilization. If it did, then S_2's attempt to coerce the defender by varying her mobilization level will involve an estimate of his propensity to capitulate, which necessarily will be based on the magnitude of his audience costs.

Turning now to the cases where audience costs clearly matter, Figure 7.3(d) illustrates some startling points. As our intuition would have it, escalation can be extremely dangerous if S_2 has high audience costs: since she is unlikely to back down and since coercion is too costly, militarized threats inevitably lead to a lock-in and war. This is in keeping with the original argument that incurring audience costs is a very risky proposition. Compare either of the high-cost scenarios to any one of the remaining two in which S_2's audience costs are low: since coercion is now worthwhile, escalation need not lead to war. Indeed, once DOP favors S_1 sufficiently (e.g., over about 65%), escalation is free of any risk because the defender achieves assured compellence and S_2 capitulates whenever he makes a threat. In contrast, war is certain at all DOP between about 15% and 95% whenever S_1 threatens a challenger who faces high audience costs for backing down.

The next step, illustrated in Figure 7.3(c), also follows our original logic: because threats against a high-audience cost challenger are so dangerous, S_1 is much less likely to make them. The probability of escalation is lower than the corresponding probability when S_2 will find it easier to capitulate because of low audience costs. Observe again that there is a noticeable difference in S_1's propensity to issue threats in these cases even though the resulting risk is not too different whether his own audience costs are high or low. This seems paradoxical: Escalation stability is better when audience costs are low for both actors than when it is low for S_2 only, and yet S_1's propensity to threaten is *lower* in the former case. In other words, S_1 is less likely to escalate when this escalation is less likely to lead to war. This seems rather odd, so we need to take a closer look at the mobilization levels shown in Figure 7.4.

Comparing expected mobilization levels between the two scenarios, we readily observe that mobilization when audience costs are low for both actors is generally higher than mobilization when audience costs are low for S_2 only. Coercion is more expensive for a defender with low costs of capitulating. This makes sense: a defender who does not face high audience costs is in a relatively weak position because capitulation is tempting. To overcome this, credible mobilization must be rather high because fighting must yield a relatively high payoff. Conversely, a defender with high audience costs need mobilize relatively little because the payoff from fighting need not be all that great to become preferable to capitulation. This now implies that whereas a low audience cost defender would benefit from improved escalation stability, he is still less likely to make threats because these are much more expensive.

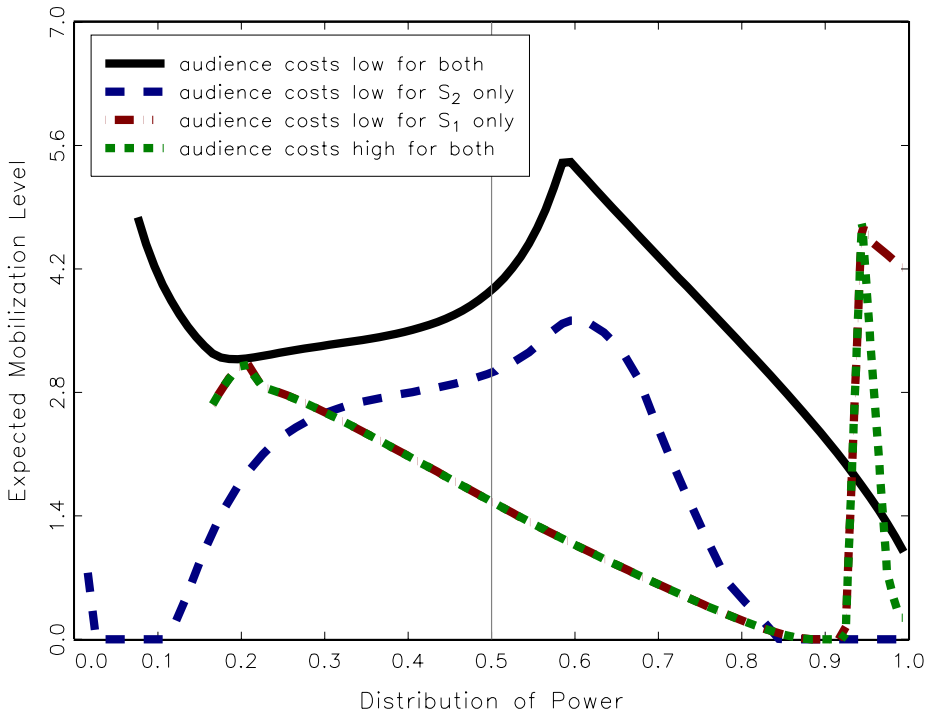


Figure 7.4 Defender's Mobilization and Audience Costs (general uncertainty, medium war costs).

RESULT 7.5 *When facing a challenger with low audience costs, the defender will be less likely to escalate if his audience costs are low than if they are high even though this escalation will be more stable.*

Consider now Figure 7.3(b) and these same two scenarios. When audience costs are low for S_2 only, the defender is both more likely to threaten and escalation is more likely to lead to war than when audience costs are low for both. Not surprisingly, the probability that S_2 will initiate the crisis is always lower in the first scenario. In fact, the more DOP favors the defender, the more dramatic the difference: at DOP of about 80%, the probability of a challenge to the status quo is zero in the former case and almost 90% in the latter.

Compare these now to the cases where audience costs are high for S_2 (as we have seen above, S_1's audience costs here are largely irrelevant). War is usually certain upon escalation, and this makes S_1 relatively reluctant to make threats. When DOP significantly favors the challenger, he will simply appease outright. In these cases, the probability that the status quo will get challenged (and peacefully overturned) is highest. But once DOP improves sufficiently to make S_1 willing to make some threats, the certainty of war quickly has a deterrent effect on S_2, and the probability of a challenge declines quickly. In fact, it is lowest for DOP up until about 75%.

To put it another way, for many reasonable distributions of power, the probability of a crisis is lowest when the challenger has high audience costs. This makes sense given how unstable escalation is going to be. But then at DOP of about 75%, situational stability is even

better when S_2's audience costs are low even though the risk of war is zero. This is also not surprising: here S_2 expects to be compelled for sure, so there is no benefit in challenging. This now reveals two different ways of achieving deterrence: threats that will certainly lead to war (one that S_2 herself will be willing to fight if she initiates the crisis and gets to that point) or threats that will certainly lead to peace because they will get S_2 to capitulate for sure. As we have seen, the literature has long recognized the first type. However, the second deterrent tactic is only possible through military coercion, and without a proper analysis of military threats, it has gone unrecognized. The upshot is that the ability to militarize a crisis can be a very effective way to stabilize the situation when the challenger has low audience costs.

RESULT 7.6 *General deterrence success can be maximized with threats that make either war or the challenger's capitulation certain upon escalation. The former tactic works against challengers with high audience costs, and the latter works against challengers with low audience costs.*

Thus far, most of our findings do not seem to contradict conventional wisdom about audience costs even if the specifics of the interactions between the various components of stability have produced some novel insights. Putting everything together, however, is another story altogether. If we follow the convention and argue that democracies have higher audience costs, then the democratic peace literature must lead us to predict that situational stability must be highest with audience costs are high for both actors.[6] Looking at Figure 7.3(a) reveals that for any DOP over 40%, this is precisely wrong: *situational stability is worst exactly when both actors have high audience costs.*[7]

There are at least two lines of attack here. First, one could argue that democratic governments do not necessarily face higher audience costs than other regimes.[8] Second, one could argue that it is not audience costs and the supposed attendant ability to signal resolve that is the crucial factor operating between two democracies.[9] Since our model is agnostic about regime types, the only proper thing to do is understand why high audience costs can be so destabilizing and leave the resolution of the new puzzle to future work.

As we have seen, when audience costs are low for both actors, the probability of the challenger initiating a crisis is high but because the likelihoods of the defender escalating and of the escalation leading to war are both low, overall situation instability will tend to be very low. If a high audience cost defender confronts a low audience cost challenger, the proba-

[6] And so it does. Fearon (1994a) and Guisinger and Smith (2002) argue from a theoretical perspective, and Eyerman and Hart, Jr. (1996), Palmer and Partell (1999), and Gelpi and Griesdorf (2001) provide some corroborating statistical evidence. Lipson (2003) gives a book-length treatment.

[7] Given that the defender's audience costs are irrelevant when a high-audience cost challenger is involved, the risk of war is the same as when his costs are low.

[8] I have shown, at least theoretically, that media freedom and unbiasedness is important in disciplining both the government and opposition, but mixed regimes are most vulnerable to audience costs (Slantchev, 2006). Bueno de Mesquita et al. (2003) argue that autocratic leaders face low-probability but extremely high-cost penalties for foreign policy failures, and Weeks (2008) argues that they are dependent on the support of a small subset of the population that is able to hold them accountable. Any of these approaches would undermine the usual association between audience costs and regime type.

[9] Even if high audience costs were detrimental as this model suggests they may be, the democratic peace can still persist because of other features of this regime type that makes democratic dyads more peaceful. These could be norms (Weart, 1988), institutional constraints (Bueno de Mesquita and Lalman, 1992; Maoz and Russett, 1993), military power (Lake, 1992), or some combination (Russett, 1993).

bility of a crisis is lower but escalation instability is worse and the likelihood of escalation higher, making the situation less stable overall. If, on the other hand, a high audience cost challenger confronts any defender, things get tricky. At DOP up to about 35%, situational stability will be better than under any of the other two scenarios. This is because for such asymmetric distributions of power, the defender is significantly less likely to make threats given that they must lead to war. Since the challenger herself is also less likely to initiate, overall situational instability is lower. (This is where "peace by threats of war" deterrence is most successful.) Once DOP begins favoring the defender, however, his propensity to make threats, while still lower than under the other two scenarios, is no longer drastically lower. Since escalation still leads to war, S_2's propensity to initiate also begins declining rapidly but not quickly enough to compensate, resulting in overall situational instability. (This is where threats of war simply lead to war, and "peace by threats of peaceful compellence" deterrence is most successful.)

RESULT 7.7 *When the distribution of power seriously favors the challenger, situational instability is worst when a high audience cost defender faces a low audience cost challenger. In all other cases, situational instability is worst when a defender confronts a high audience cost challenger.*

This illustrates some of the trouble with the traditional signaling approach to the democratic peace that is based on audience costs. The logic there usually goes something like this: democratic leaders can generate high audience costs; therefore, they can commit credibly not to back down during a crisis; therefore, a crisis against a democracy is very likely to lead to war. Foreseeing this chain of events, two democratic leaders know that if they get involved in a crisis, war will be inevitable, and as a result they do not even challenge each other. Our analysis shows that although each of the separate steps is correct (high audience cost players do face certain war upon escalation, the defender is less likely to threaten, and the challenger is most often less likely to initiate), the overall conclusion does not follow: situational instability can still be worse. The reason is that the increase in the probability of war upon escalation could be so great that it can swamp out the decreases in the probability of crisis and escalation.

RESULT 7.8 *Actors with high audience costs do face an extremely high risk of war upon escalation, so the defender is less likely to threaten and the challenger is less likely to initiate. However, overall crisis stability may still be worst, undermining the conventional link between audience costs and peace.*

What this implies for our strategic foundations of the democratic peace must be explored elsewhere. I now turn to another theoretical puzzle: what makes for an effective deterrent threat?

7.3 Deterrent Efficacy of Military Threats: Power or Beliefs?

How does the distribution of power relate to deterrence failure? If we take deterrence failure to refer to the probability that the challenger resists a threat and the crisis ends in war, then Figure 7.1(d) reveals that this probability is *decreasing* as DOP improves in favor of the defender. It we take it to refer the probability that status quo would be disrupted by war, then Figure 7.1(a) shows that for low to intermediate values of DOP the probability of deterrence failure *increases*, but for intermediate to high values of DOP, the probability of

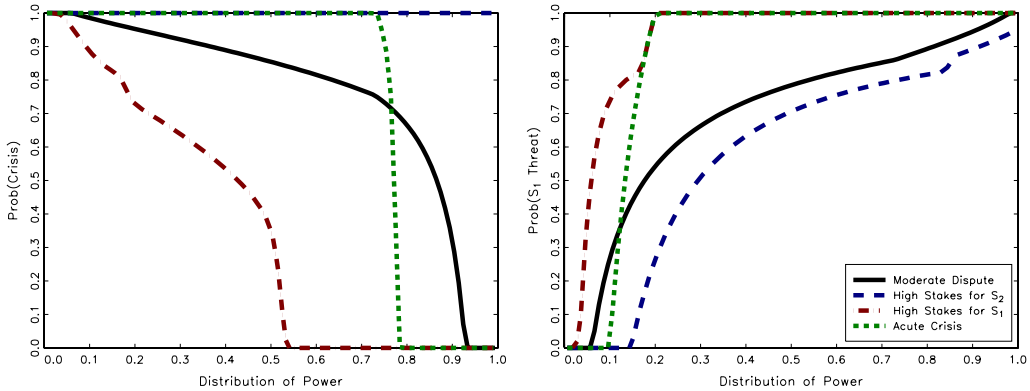
deterrence failure *decreases*. To nobody's surprise, the answer to the question depends on how we specify deterrence failure because doing so determines which constituent probabilities must be taken into account. As has been recognized now for some time, because S_2 can choose whether to turn a situation into a crisis, she effectively *self-selects* into crises. This implies that crises are not random events but are strategically produced by the actors, so the evaluation of the impact of some variable on deterrence failure must take that into account. Since this is the first study to incorporate militarized behavior, we must now consider how these selection effects are mediated through militarized escalation.

The best recent exposition of the selection problem that earlier work had ignored is in Fearon (2002).[10] Essentially, the argument concedes that credible threats would be more likely to deter a challenger from pressing her demands in a crisis. However, it does not follow from this that such threats would make deterrence more likely to succeed. The reason is that S_2's initial choice is strategic: she is very unlikely to escalate if her opponent's posture is credible unless she is truly serious about her demands. Hence, while credible threats would tend to discourage frivolous challenges (bluffs), the overall effect of observable escalation appears paradoxical: When S_2 *chooses* to press her demands even though she believes her adversary is very likely to resist, she is much more likely to do so even in the face of determined opposition. In other words, S_1's deterrence is much more likely to fail and because we have assumed a credible threat, the crisis ends in war. The overall impact of S_2's belief that her opponent is likely to stand firm is thus to enhance situational stability but undermine crisis stability.

This self-selection effect has profound implications for the empirical evaluation of factors that are conducive to deterrence success and for the methodology used to evaluate various theoretical claims. For example, basing inferences on a sample that includes only crises instead of a sample that also includes situations that could have become crises can seriously bias one's results (Achen and Snidal, 1989). The problem is further aggravated when our hypotheses make different predictions about the effect of explanatory variables at various stages of the crisis (Fearon, 1994b). Worse, limiting one's inference to the given sample may not help overcome the bias when the selection mechanism is correlated with the dependent variable (Signorino, 2002). In our case, S_1's decision—which determines the outcome of crisis escalation—is correlated with S_2's initiation—which determines the selection—and so the selection effect will necessarily lead to biased inferences unless it is properly incorporated in the statistical model. I will not delve into the empirical debate here. Rather, I want to evaluate the self-selection hypothesis in the context of the military deterrence model.

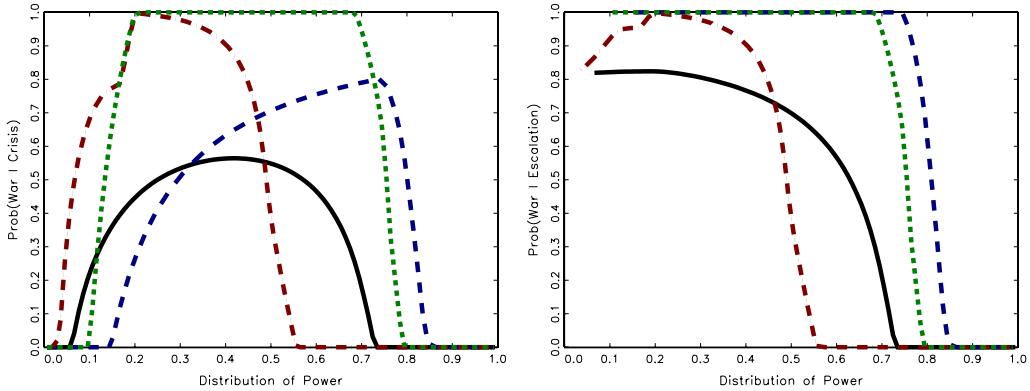
Keeping these subtle nuances in mind, we can now address a vexing theoretical question

[10] The literature on selection effects in classical rational deterrence theory and crisis bargaining is not vast but is quite sophisticated. In his review of work on deterrence at the time, Levy (1988) is quite aware of selection bias in studies that assume the existence of a prior threat by the challenger and the problematic inferences this implies for evaluating the impact of capabilities on immediate deterrence success. Achen and Snidal (1989) provide another early informal critique of selection bias in case study research on the topic. Morrow (1989c), Bueno de Mesquita (1990), Bueno de Mesquita and Lalman (1992), Smith (1996), and Smith (1998a) develop the ideas from a formal theoretical standpoint. Fearon (1994b), Bueno de Mesquita and Siverson (1995) and Reed (2000) analyze various propositions with explicit attention to selection effects. Harvey (1998) carefully outlines necessary and sufficient conditions for testing deterrence theories, while Danilovic (2001) and Signorino (2002) argue about the shortcomings in the extant statistical analyses of large-N data sets. Signorino (1999) and Smith (1999) develop appropriate techniques for empirical testing of such models.



(a) Probability Challenger Initiates [General Deterrence Failure]

(b) Probability Defender Threatens



(c) Crisis Instability

(d) Escalation Instability [Immediate Deterrence Failure]

Figure 7.5 Distribution of Power and Selection Effects (low war costs).

about efficacy of deterrent threats. Figure 7.5 shows the two types of deterrence failure along with the probabilities of escalation and war conditional only on a crisis for each of the usual four balance of interest scenarios with one important modification. Later in this section, I will want to investigate the effect from the perspective of a S_1 with a particular valuation. Under the original specification of peripheral and vital interests there is no type that is common to both sets. To accommodate this, I extend the range of types defining a peripheral interest to overlap somewhat with the range of types defining a vital interest: instead of the range $[0, 12.5]$, I use $[0, 17]$. To emphasize this, the plots refer to a "moderate dispute" rather than a minor one. The simulations assume low costs of war and baseline system militarization.

Since Fearon (2002, 245) is admirably clear about the selection effect, it is best to cite his theoretical hypothesis verbatim:

> The more the challenger initially expects the defender to prefer war to conceding the issue... the more likely is general deterrence to succeed, other things equal. But if general deterrence does

> fail, immediate deterrence will then be less likely to succeed, despite the defender's initial credibility. By the same token, if the challenger initially expected that the defender would probably prefer concessions to war, then general deterrence will be less likely to succeed, but subsequent efforts at immediate deterrence will be more likely to work.

To evaluate this with our simulation results, fix S_2's interest as peripheral and vary S_1's interest from moderate to vital: this allows a comparison of two scenarios. In the first (moderate dispute), the challenger believes to be facing a relatively low-valuation defender, whereas in the second (high-stakes for S_1), she believes to be facing a relatively high-valuation defender. Because high valuation defenders have higher expected payoffs from fighting, this means that in a high-stakes for S_1 crisis, the challenger initially expects the defender to be more likely to prefer war to conceding the issue.

Turning now to Figure 7.5(a), observe that S_2 is much less likely to challenge the status quo in the high-stakes for S_1 scenario; that is, general deterrence is much more likely to succeed. This is exactly in conformance with Fearon's hypothesis. Turning now to immediate deterrence shown in Figure 7.5(d), observe that for DOP up to about 50%, escalation is more likely to lead to war when S_1's stake is high than when it is peripheral. That is, immediate deterrence is more likely to fail even though general deterrence is more likely to succeed. This also supports the hypothesis. However, once DOP exceeds 50%, this no longer holds: at these distributions of power, immediate deterrence is more likely to fail in a minor dispute, precisely when general deterrence is more likely to fail as well. Stating this result parallel to Fearon's claim, we conclude that

RESULT 7.9 *If the distribution of power sufficiently favors the defender, then the more the challenger initially expects the defender to prefer war to conceding the issue, the more likely is general deterrence to succeed. If general deterrence does fail, immediate deterrence is still more likely to succeed as well. There will be no appreciable selection effect.*

This is not an artifact of assuming that S_2's interests are peripheral. In fact, the result is even stronger if we assume they are vital because in this case the selection effect disappears entirely. To see this, compare a crisis with high stakes for S_2 only to an acute crisis. As before, general deterrence is more likely to succeed when the challenger believes the defender is more likely to prefer war to concessions: the probability of initiation is never higher (and once DOP exceeds about 75%, significantly lower) in the acute crisis scenario. However, immediate deterrence is equally unlikely to work for all DOP up to about 70% under either scenario and much more likely to work in an acute crisis for values above that. In other words, once the distribution of power sufficiently favors the defender, both general and immediate deterrence is more likely to succeed when the challenger initially believes that the defender is more likely to prefer war to concessions.

In the equilibrium of our deterrence model, the more interested the challenger believes S_1 to be in the issue, the higher must her own interest be in order for her to be willing to initiate a crisis. This is entirely in accordance with Fearon's logic and is an equilibrium requirement for incentive compatibility. Hence, the claim about general deterrence will be correct, all else equal. This last qualification is crucial for without it, even this falls apart for it is no longer true that general deterrence is more likely to succeed if S_2 believes that S_1 is more likely to prefer war to concessions. To see this, compare an acute crisis to a moderate dispute. In the latter, S_1's interest is weaker and yet general deterrence is *more* likely to

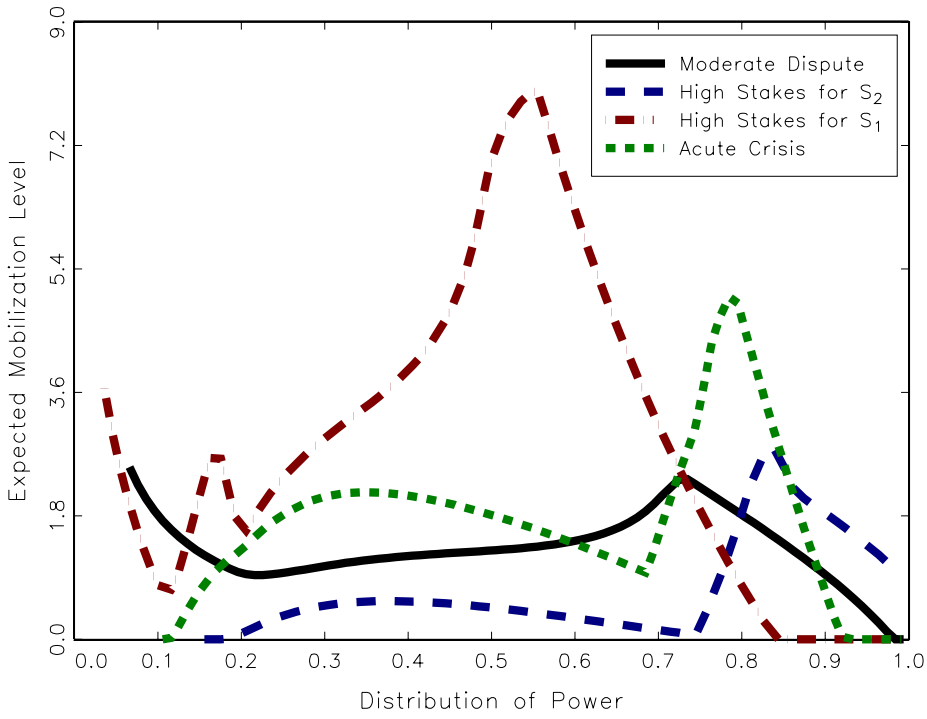


Figure 7.6 Mobilization Levels and Balance of Interests (low war costs).

succeed for distributions of power up to about 80%. In this region, general deterrence by a strong defender is less likely to work against a strong challenger than general deterrence by a weak defender against a weak challenger. Since we are holding constant all variables except the challenger's valuation in this comparison (relative to the previous two), it follows that the strength of the selection effect (indeed, its existence) depends on the balance of interests as well as the distribution of power, at the very least.

The selection effect logic appears unassailable, so it is important to understand why it sometimes fails. As one might expect at this point, it has to do with the characteristics of the military instrument. Figure 7.6 shows the expected mobilization levels for the four balance of interest scenarios we have been examining.

EXAMPLE 7.1 (Selection Effect). To see precisely how beliefs generated by high expectations of success of general deterrence lead to high probability of failure of immediate deterrence should crisis actually occur, compare a moderate dispute to a crisis in which stakes are high for the defender only. We shall look at the crisis from the perspective of $v_1 = 17$, which is the highest-valuation type in a moderate dispute and a medium-valuation type in the high-stakes crisis. Since we want to examine a case where the selection effect occurs, consider DOP of approximately 35%.

In a moderate dispute, the least-valuation S_2 to initiate a crisis is $t \approx 1.16$ in the pooling MTM continuation equilibrium. Since $\overline{\tau} \approx 10.53 < v_1 < v^{ca}$, it follows that S_1 with valuation v_1 will engage in coercion. His optimal credible coercive mobilization level is $\widehat{m}(v_1) \approx 2.66$, for which $v_2^*(\widehat{m}(v_1)) \approx 4.18$. Since conditional on crisis initiation S_1 believes that S_2's valuation is distributed on $[1.16, 12.50]$, this means that S_1 expects her to capitulate with probability $\Pr[v_2 < v_2^*(\widehat{m}(v_1))] \approx 0.27$.

When stakes are high for S_1, the least-valuation S_2 to initiate a crisis is $t \approx 5.07$, and the MTM continuation equilibrium is with war and coercion. Since $v^{wq} \approx 7.24 < v_1 < 21.39 \approx v^{wc}$, it follows

that S_1 with valuation v_1 will no longer opt for coercion but will instead mobilize for certain war. His optimal credible fighting mobilization level is $w^*(v_1) = 1.74$, and of course the probability that S_2 will capitulate after seeing this is zero. Since $v_2^*(5.07) = t$, it follows that the lowest coercive mobilization level is $\widehat{m}(v^{cq}) = 5.07$. In other words, to get S_2 to capitulate with positive probability assuming she believes the mobilization to be credible, S_1 would have to mobilize more than this very high level.

Observe now that since conditional on mobilization S_1 is certain to attack when resisted, the crucial factor in determining immediate deterrence success is the defender's belief about what it would take to get S_2 to capitulate. That is, the key variable is t, the least-valuation type to initiate the crisis. In a moderate dispute, this is very low, which means that S_1 learns almost nothing from crisis initiation. Since an overwhelming majority of types challenge the status quo (probability is 91%), S_1 remains fairly confident that he may be facing a very low valuation challenger. This makes coercion attractive and by mobilizing at the relatively low level of 2.66, the defender manages to get a decent 27% likelihood of compellence success. To wit, when general deterrence is expected to fail even for very weak challengers, coercion in an actual crisis remains an attractive possibility, and the defender attempts it, increasing the chances of immediate deterrence success.

In a crisis with a high-valuation defender, t is significantly higher. Really low valuation challengers are now successfully deterred and only relatively strong ones initiate (the probability is now only 60%). Because of this self-selection of S_2, once crisis begins S_1 must believe that he is facing a relatively high-valuation opponent. This makes coercion unattractive: he would have to go all the way over 5.07 to get a minuscule probability of compellence success when in a minor dispute he could get 27% probability with barely 2.66. Since coercion is prohibitively expensive, the defender switches to war preparation and mobilizes only at the fighting level of 1.74, knowing full well that this would lead to certain war. (War is not certain in expectation because the higher-valuation defenders will engage in coercion.) When general deterrence is expected to be fairly successful in screening out low-valuation challengers, coercion against self-selected high-valuation ones becomes too expensive and is not likely to be attempted, decreasing the chances of immediate deterrence success.

This illustrates Fearon's self-selection logic precisely: the entire result hinges on the defender's beliefs updated from how successful general deterrence is expected to be. The more successful general deterrence, the stronger the challenger who chooses to initiate, the less attractive coercion, and the dimmer the prospects for immediate deterrence success.

Why, then, does this logic no longer work once the distribution of power becomes sufficiently favorable for the defender? The clue is in the increasingly high mobilization levels of the defender in the high-stakes scenario. As we have seen, self-selection by the challenger leaves S_1 to cope with a fairly determined opponent, making coercion quite costly. However, the defender is actively coping with the problem as best as he can. As DOP improves, more and more types resort to coercion and they do so at increasing levels because they can afford to minimize the chance of war. At DOP of around 20%, the expected probability of war begins to decrease, as seen in Figure 7.5(d), and as DOP improves further, the rate of decrease accelerates because coercion becomes even more affordable. To make matters more concrete, let us look at an example.

EXAMPLE 7.2 (Coercion Trumps Selection Effect). Consider the same two scenarios as before from the

perspective of a defender with valuation $v_1 = 17$. Since we want to examine a case where the defender's ability to afford coercion overcomes the selection effect, consider DOP of about 51%.[11]

In a moderate dispute, the least-valuation S_2 to initiate a crisis is now slightly higher at $t \approx 1.84$, but the MTM continuation equilibrium remains the one with pooling. Since $\overline{\tau} \approx 8.40 < v_1 < v^{ca}$, it follows that S_1 with valuation v_1 will engage in coercion, as he did at DOP of 35%. His optimal credible coercive mobilization level is $\widehat{m}(v_1) \approx 2.82$, for which $v_2^*(\widehat{m}(v_1)) \approx 6.35$. Since conditional on crisis initiation S_1 believes that S_2's valuation is distributed on [1.84, 12.50], this means that S_1 expects her to capitulate with probability $\Pr[v_2 < v_2^*(\widehat{m}(v_1))] \approx 0.42$.

When stakes are high for S_1, the least-valuation S_2 to initiate a crisis is much higher at $t \approx 8.45$, and the MTM continuation equilibrium is now with war and compellence only. Since $v^{wq} \approx 4.95 < 16.76 \approx v^{wa} < v_1$, it follows that S_1 with valuation v_1 will mobilize for assured compellence (as opposed to certain war at DOP of 35%). The unique compellent mobilization level is $\overline{m} = 10.37$, and of course S_2 will certainly capitulate after seeing this.

As one would expect, when the distribution of power favors the defender considerably more than in the previous example, the challenger will be more reluctant to initiate under any scenario. In a moderate dispute, the lowest valuation type to self-select is only slightly higher at DOP of 65%, and as a result, the defender's behavior is strikingly similar: he mobilizes for coercion and since he can now do this more profitably, he uses a slightly higher level and his chances of obtaining S_2's capitulation are also better at 42% instead of 27%. In comparison to a moderate dispute under the less attractive DOP, the probability of general deterrence failure, although lower, remains high at 85% (so the defender learns little from initiation) but the prospects for immediate deterrence success are actually better.

In a crisis with a high-valuation defender, t is significantly higher compared to either the moderate crisis under the same DOP or the high-stakes crisis under the less attractive DOP. The probability of general deterrence failure is now only 32%, so initiation leads S_1 to believe that the challenger's valuation is very likely to be quite high. In other words, the self-selection works precisely as before. In fact, because S_1 is so pessimistic once crisis actually begins, he will not attempt coercion, it is not worth it given how hard the opponent would be to coerce. Instead, he opts either for optimal war or, if we can afford it, for assured compellence. The moderately high valuation type we have focused on happens to be among the latter types and he manages to obtain S_2's certain capitulation and enjoy no risk of war, albeit at great cost. That is, precisely because self-selection has screened out even moderate-valuation challengers, the defender faces a stark choice. However, the military instrument can achieve something signaling cannot, and S_1 takes full advantage of it, making sure that immediate deterrence will succeed.

RESULT 7.10 *Because of strategic self-selection by the challenger, the less likely is general deterrence to fail, the more pessimistic the defender will be should a crisis actually occur. However, if the distribution of power is favorable enough, the defender will mobilize more aggressively precisely because of his pessimism, improving the prospects of success for immediate deterrence. Hence, both general and immediate deterrence will be enhanced.*

Danilovic (2001, 104) cites a discussion with Paul Huth where he advances an argument according to which ex ante measures of defender's interests that are positively correlated with general deterrence success need not be negatively correlated with immediate deterrence success. It turns on which side is "more" motivated than the other, with motivation being a

[11] We have to be careful not to select a DOP so high, e.g., 60%, that we end up with a zero-probability event: a crisis like this will never occur in equilibrium because the challenger will stay with the status quo for sure.

key unobservable variable. Specifically, if the defender is more motivated than the challenger, then immediate deterrence should still work despite the self-selection. Levy (1988) argued along similar lines when summarizing the state of the art in deterrence studies in 1988. He concluded that the "finding that superior military capabilities along are not necessarily sufficient for deterrence can be explained in part by the tremendous importance of the interests and resolve of the initiator of the crisis" (510). He cautions that this does not handle the potential selection bias but at least could explain the empirical results. As we have seen the balance of interests does affect the strength of the selection effect but somewhat surprisingly the impact of the selection effect is *least* pronounced when the challenger has vital interests at stake. Furthermore, the MTM suggests that whereas self-selection is *always* present and it should shift the defender's priors toward a more pessimistic evaluation of his own position, this does not necessarily mean that immediate threats are less likely to work.

Relative motivation, although prominently figuring in many studies of deterrence, is quite a problematic concept. Maoz (1983) and Karsten et al. (1984), for example, criticize previous studies of deterrence for focusing on capabilities, commitment, and signaling to the exclusion of motivations. From our perspective, it is difficult to disentangle the concept of resolve (or motivation) from capabilities because in the MTM, "resolve" is the actor's willingness to fight rather than concede, and this willingness is a function of his expected benefits from war, which in turn depend on the distribution of capabilities as well as interests at stake. It may well be the case that an actor keenly interested in some issue (and therefore highly motivated to win it) may lack the resolve in the sense that the distribution of power makes using force too unattractive. Conversely, when his military advantage is pronounced, he may be quite resolved to fight over a minor issue. In the MTM resolve and commitment are interchangeable labels for the same phenomenon because each refers to whether an actor would rather fight than give up. Signaling, while clearly important in communicating this commitment, is not a separate concept in itself as I have argued throughout this book: the very process of committing oneself can also reveal one's resolve which itself can change during the crisis. Hence, this entire emphasis on motivations can be seriously misleading if it is posed in opposition to the traditional ideas.

Leaving aside the controversial comparison of relative motivation and its subsequent relation to observable indicators of interest, the argument developed here shows that what we have to focus on is the defender's reaction to his own updated beliefs. In particular, the defender's immediate threat can be enhanced precisely because the selection effect makes him pessimistic about the chances of milder forms of coercion, which in turn leads him to adopt a more aggressive strategy. Since both informal and formal discussions of deterrence usually come from underlying models that simply do not allow for graduated responses by the defender, they are likely to miss that particular effect.

Since coercion is a strategic choice, when the revision of beliefs that follows upon self-selection convinces the defender that coercion would not be worth trying, he may well opt for a compellent strategy that will increase the prospects for immediate deterrence success. In this case, for lower valuation defenders escalation will lead to certain war but for higher valuation ones it will lead to certain peace. In expectation, the probability of immediate deterrence failure in the high-stakes dispute is lower than the corresponding probability in a moderate dispute. Fearon is precisely correct about how self-selection interacts with beliefs. The military deterrence model demonstrates how this may sometimes lead to behavior by

the defender that negates the consequences of the selection effect for immediate deterrence. The crucial factor here is the subverting aspect of the military instrument for without it the defender would have been quite helpless to deal with the stronger challenger.

RESULT 7.11 *The selection effect must persist in all non-military models. The weakening (or elimination) of its consequences for immediate deterrence can only be achieved with the military instrument.*

This discussion of selection effects has only scratched the surface of the implications of the military deterrence model for empirical studies. One interesting question is how the immediate balance of forces affects immediate deterrence; that is, what escalation stability is as a function of the new distribution of power produced by the defender's crisis mobilization. Unfortunately, statistical analysis must await the collection of data on military deployments.

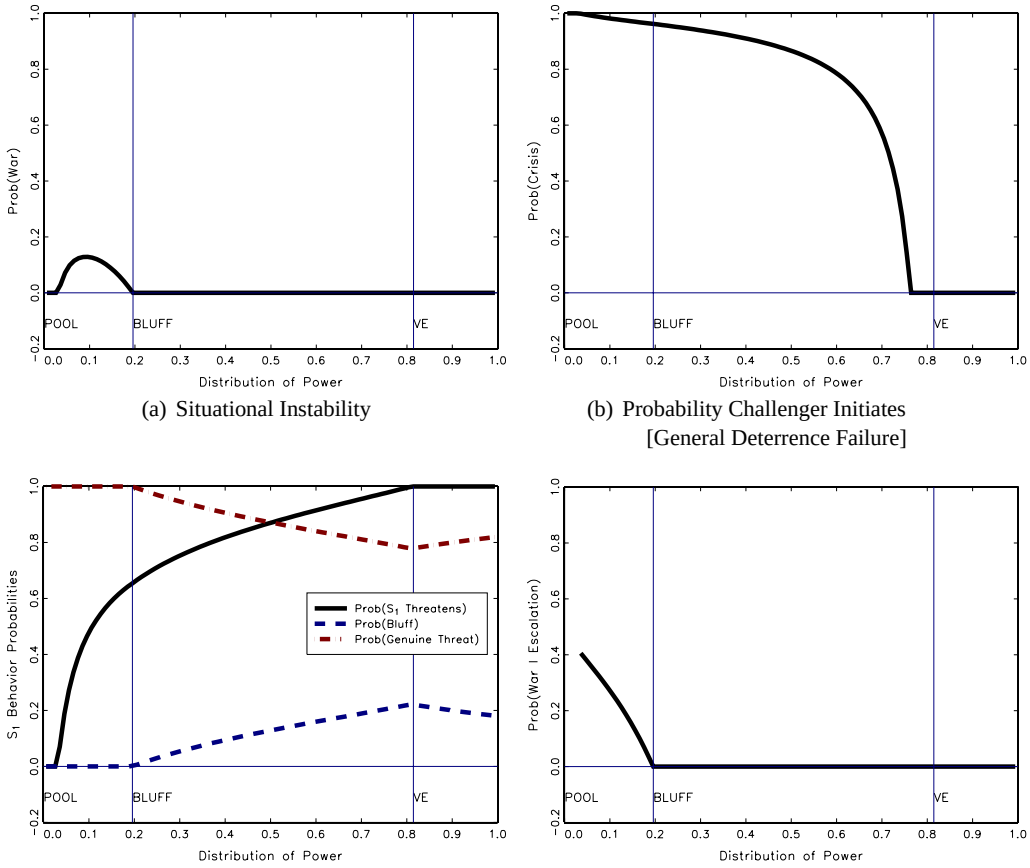
7.4 Bluffs, Sham Crises, and Deterrence

In Chapter 3 we found that the military instrument permits equilibrium bluffing when escalation is expected to lead to assured compellence. One may wonder whether the challenger would ever initiate a crisis in which the defender is expected to bluff. After all, since bluffing only happens when S_2 capitulates for sure, why would she start a crisis in which she will surely back down in the endgame? If she never initiates, then all this discussion about bluffing is devoid of empirical content: we should never expect to see any of it. I now show that it is quite possible for a bluffing defender to coerce successfully a bluffing challenger. While this appears to be a strange "sham" crisis with zero probability of war, the costs the defender incurs are quite real. This behavior can lead to a lot of sabre-rattling without any serious chance of fighting actually breaking out. This sort of thing can be entirely rational even though in retrospect it may appear empirically to have been a rather costly exercise in pointless bluster. This finding has serious implications about empirical studies of deterrence.

Since bluffing only happens when S_2 is comparatively easy to compel, assume medium war costs for the defender and high costs for the challenger. Figure 7.7 shows the quantities of interest for the general uncertainty case under baseline system militarization. In addition to the usual probability of S_1 threatening, Figure 7.7(c) also plots the two complementary probabilities of the threat being genuine or a bluff.

Despite the relatively high probability of general deterrence failure, overall situational instability is quite low. This means that with such a configuration of interests and costs we should see relatively few wars despite relatively frequent challenges: the status quo will either be preserved or be revised in favor of the challenger, but in both cases the crisis is likely to be resolved peacefully. The only exception is when the distribution of power favors the challenger so much that the defender cannot afford to practice assured compellence and must accept some chance that his threats will fail despite being credible. Once DOP exceeds 20%, however, this is no longer an issue and the risk of war is zero.

Consider now what happens as DOP improves further in defender's favor. Escalation remains fully stable even though the defender is not only more likely to threaten but also more likely to bluff. To understand why the probability of bluffing must increase without affecting escalation stability, recall that when practicing assured compellence, S_1 need only mobilize sufficient forces to compel S_2's capitulation despite her belief that he might back down when resisted. As the distribution of power improves, the required mobilization level



(a) Situational Instability

(b) Probability Challenger Initiates [General Deterrence Failure]

(c) Defender Behavior Probabilities

(d) Escalation Instability [Immediate Deterrence Failure]

Figure 7.7 Bluffing by Both Actors (general interests, baseline militarization, medium war costs for S_1, high costs for S_2).

goes down, which in turns makes it more attractive to even lower valuation types. But since this also means that a larger proportion of types now escalates, the overall probability of a threat goes up. The effect this has on the challenger is intuitive. She faces a higher probability of escalation when it still leads to her certain capitulation. Not surprisingly, she becomes less likely to initiate. Observe now that all this straightforward logic translates into a counter-intuitive finding:

RESULT 7.12 *As the distribution of power improves from the defender's perspective, the challenger becomes* less likely to initiate *a crisis even though the defender becomes* more likely to bluff *in his threats.*

The crucial point here is that challenger may be quite successfully compelled even if the defender is bluffing. In other words, immediate deterrence can succeed even in cases when a careful examination of the historical record reveals that the defender never seriously intended to fight in the first place if his threat was resisted. Furthermore, since S_2 will still initiate

such a crisis with positive probability, it follows that if that examination also revealed that she never really intended to fight, then we still cannot conclude that this was not a deterrence encounter. It is, in fact, quite possible for an initial bluff by the challenger to be successfully countered with a bluff by the defender. It is also quite possible for this bluff to work and the defender to appease rather than risk war by calling it.

For example, at parity in the distribution of power, the challenger is over 80% likely to initiate a crisis even though the defender is not only over 80% likely to make a threat, but there is a chance of close to 15% that this threat will turn out to be a bluff. For a 20% chance of appeasement by S_1, the challenger is willing to trade an 80% chance of capitulation despite knowing that even when the defender fails to appease, he may still be bluffing.

This seemingly innocuous example has profound implications for our research on deterrence. In their influential critique of empirical studies of rational deterrence theory (RDT), Lebow and Stein (1990) charge that many of the cases identified by Huth and Russett (1988) are improperly classified as deterrence encounters. As they put it, "in thirty-seven cases, we find no evidence that the alleged attacker intended to use force or that the putative defender practiced deterrence; *both are necessary to identify valid cases of deterrence*" (337, emphasis added). Even more to the point, they specifically exclude bluffing (343) and insist that to include a case in a dataset designed to evaluate deterrence theory, "there must be evidence that the challenger considered an attack, as well as evidence that a defender attempted to deter." The simple example above clearly shows that these claims are wrong. It is quite possible to have cases where *neither* of the actors has any intention to fight and yet the encounter is still without doubt one of deterrence.

To understand why, observe that uncertainty is a crucial component of the MTM, as it is for all RDT models, as I noted in Chapter 2. Huth and Russett (1990, 478) quite appropriately chide Lebow and Stein (1990) for not appreciating its role in the theory. In particular, the insistence that cases be considered deterrence encounters only when challengers seriously contemplate war is

> conceptually wrong because it excludes a wide range of cases in which deterrence is still operative.... Lebow and Stein's definition fails to allow for the logical possibility that the potential attacker may be probing the resolve of the defender through demands and threats, and that he will decide whether to use force only after such probing has helped to clarify his beliefs about the defender's resolve.

As the example above shows, Huth and Russett are not only right but their argument can be further strengthened by the insights from the MTM. One of the most important things we found is that the decision to fight depends not only on clarifying the defender's resolve but also on the defender's escalatory behavior as well. The clarification itself may involve alteration in the distribution of power, and this in turn will have an effect on the challenger's incentives quite independent from the signaling that reveals the defender's commitment to fight.

To put it bluntly, an actor's decision to attack can only be specified *a priori* in the rare cases where he would do so no matter what the other does. (The military deterrence model admits these as well: any initiation that results in escalation toward certain war is essentially an illustration of that possibility.) In most other instances, however, the challenger's resolve will depend on the defender's behavior and as such cannot be specified before the crisis begins because of the uncertainty involved. The challenger can only estimate the likelihood

of various stances the defender's response can take, and it is quite possible that she will only fight when he fails to mobilize sufficient forces to compel her. (Note that this does not require one to assume that she is uncertain about her own motivations.)

The irony is that Lebow and Stein (1990, 342) criticize the use of military moves as indicators of intention to attack because "military deployments. . . can be used for a wide range of purposes. . . [and] only a few of the deployments identified by Huth and Russett in their cases are associated with the intention to attack." In other words, it is not like Lebow and Stein do not appreciate that mobilization may be more than simple preparation for war. In this we agree: as the MTM shows, sometimes a military move may be a necessary step toward war, other times it may be an attempt to coerce involving some risk of war, and yet other times it may be compellent with the attendant expectation that it will not lead to war. The problem is that in their rush to criticize RDT, Lebow and Stein do not take seriously the implications of uncertainty for behavior.

By selecting on cases of clear intent to fight, Lebow and Stein's research design will essentially admit only those rare instances where military moves were preparatory for war. Recall that this means credibility is neither an issue (all such preparation is credible) nor a consideration (it is recognized as such by the other actor). In these encounters, war is certain so it should come as no surprise that Lebow and Stein find that most of these ended with deterrence failure. It is hard to see how it could have been otherwise. In fact, in response to a criticism by Orme (1987), Lebow (1987, 197) argues that in eight out of thirteen acute crises, the challenger forged ahead with her demands even though the defender's commitments were "clearly defined and repeatedly publicized" and despite the defender possessing the "military capability to defend them. . . [and giving] every indication of their resolve to do so." This is taken as evidence of the failure of RDT. But our analysis suggests precisely the opposite: by selecting on acute crises, Lebow is focusing on instances where the risk of war is high, and these in turn tend to be the cases where threats are often preparations for war. The MTM would predict that in these instances we should observe that credibility is not an issue and mobilization will either be followed by war or, less often, by the capitulation of the challenger. His finding that in all but three cases this is what happened *supports* the theory!

How we can analyze the deterrent efficacy of military threats by looking at cases where the challenger is ready to fight regardless of the defender's behavior, I do not understand.[12]

[12] Some examples of studies that select cases based on perceived high risk of war are Lebow (1981), Snyder and Diesing (1977), and George and Smoke (1974). I have already discussed the widely recognized selection problem resulting from looking at situations in which a challenge has been made. However, this recognition has not been enough when it comes to data collection. Although she explicitly acknowledges these selection-bias issues, Danilovic (2001, 107-8) includes only cases in which "at least one major power (challenger) upsets the status quo in general deterrence vis-á-vis another major power (defender) by getting into conflict with the third state (pawn)." Furthermore, what might be less obvious is that even studies of immediate deterrence restricted to crises may suffer from another bias introduced by the defender's propensity to appease. For example, borrowing from Morgan (1977), Huth and Russett's (1984, 498) third criterion for identifying relevant cases is "the officials of the defender state, either explicitly or by movement of military forces, threaten the use of retaliatory force in an effort to prevent an attack." Despite later refinements by Huth (1988) and Huth and Russett (1990), this data set is used for the evaluation of the selection-effect arguments in Fearon (1994b) and the strategic probit estimation in Signorino and Tarar (2006). This dataset essentially assumes that that the defender has escalated, and as such the outcome will also tend to be biased toward fighting. Such studies will be based on cases that are most likely to invalidate the analysis.

Achen and Snidal (1989) note that this approach will bias the empirical evaluation of rational deterrence theory because it over-samples on cases where deterrence failed. My argument goes further: these studies have led scholars to the conclusion that RDT has been falsified when in fact their findings may be fully in support of the version developed here. A failure of a credible threat does not mean that credibility was irrelevant (coercion can fail even when credible) and it does not mean that deterrence theory is wrong (it shows that military moves could be preparatory for war rather than coercive). Dogs that did not bark in the night can tell us much about that night because they could have barked. When behavior is conditional on expectations, things that *could* happen are relevant in explaining the things that *did* happen. When deterrence operates "of-the-path" of equilibrium play, it is empirically relegated to a counterfactual, but this does not mean that it was not at work.

Part IV

Conclusions

8

Implications

The maximum use of force is in no way incompatible with the simultaneous use of the intellect.
Carl von Clausewitz

The central conclusion from the analysis in this book is perhaps normatively troubling. Military threats can be useful in several ways. First, they can enable actors to establish commitments necessary to obtain better deals. Second, they can communicate these commitments credibly so that the opponent can believe them. Third, they can undermine the opponent's commitment and make it more likely that he will concede, which in turn can reduce the risk of war.

> Military threats can serve a useful purpose in crises. They may reduce the risk of war relative to purely diplomatic actions.

One may be tempted to think that military coercion is a cheap alternative to war preparation. As such one may wonder whether an opponent could be coerced by send a battalion of Marines or perhaps one task group. The analysis in this book suggests that this is unlikely to work if the force being sent is significantly smaller than what one would need for actual operations. In fact, the model suggests that it may have to be quite a bit *larger*.[1]

The reason for this has nothing to do with credibility—as we have seen, preparation for war is just as believable as an optimal military threat. The problem is that the opponent's expected payoff from war is only partially related to one's own. That is, whereas one could affect it by military preparations, if the opponent's valuation is too high, the overall payoff may remain high as well. In other words, if the opponent really cares about the disputed issue, then lowering his probability of success sufficiently may require very expensive moves by the threatener. Whereas preparation for war is "simply" preparation to fight the best possible battle, mobilization for coercion is more than that for it must also make the battle sufficiently bad for the opponent.

> One cannot succeed with military coercion on the cheap.

The unhappy implication is that sometimes this implies that a strong resolved actor would prefer to fight rather than spend the resources to convince his opponent to give up. When the British pressed their demands against the Chinese Emperor before the Opium Wars in

[1] Mathematically, it is easy to demonstrate that $m^*(v_1) < \widehat{m}(v_1)$. That is, the optimal credible war preparation is smaller than the optimal credible mobilization for coercion. However, as I stated before, this particular result depends on the assumption that each player's war payoff is independent of his opponent's valuation.

1840, they arrived with a small fleet to Chou-shan Island. On July 4, the commander of the British force invited a Chinese official to board the 74-gun ship-of-the-line HMS *Wellesley*, formidable battleship whose "overpowering broadside" would reduce anything the Chinese could hope to field against it. The Chinese official was allowed full access to inspect that firepower for himself and he was duly impressed. However, when the British commander demanded the surrender of the capital city of the island, the Chinese refused. Even while acknowledging the "unaccountable force that confronted him," he remarked, "Still, I *must* fight."[2]

The problem was not disbelief in British resolve or in their technological edge. The problem was that the Chinese felt compelled to resist mostly because they underestimated just how decisive this technological edge would be. They thought they still had a chance because the expeditionary force was not all that large (there was a lot of discussion back in Britain how to minimize the costs). The only way for the British to disabuse them of this idea without the commitment of expensive resources that had to be summoned all the way from India was to fight. This they promptly did and within hours they had captured the city, killed about 2,000 Chinese (losing nineteen men only), and introduced the word *loot* to the English language.[3]

This result contradicts Lauren (1994, 31–32) claim that military threats can be "token in character, in the hope that the appearance of a fraction of the power available would create the perception that more force would follow if necessary… the coercive effect of what little was actually done could be magnified substantially by linking it to a credible threat of additional action." According to the logic of the MTM, gunboat diplomacy is not very likely to work, at least not as well as land-based threats that are costlier and involve much firmer commitments (Blechman and Kaplan, 1978, 529–30).

When the opponent is not so weak as to make fighting preferable to coercion—as in the example above and in the 1962 Sino-Indian war—the threat may have to exceed the capability requirements in order to impart credibility. This problem is particularly acute when a state is very rich or very strong. If a state is wealthy, mobilization is not too burdensome, which makes military threats more attractive even for governments that may not care much about the contested issue. Analogously, if a state is powerful, then a relatively small mobilization can have a dramatic impact on the expected payoff from war for itself and its opponent. Again, this has the effect of making military threats too attractive. If the temptation to use them is too great, then bluffing becomes an issue, and then a resolved actor will have to mobilize overkill capability to reveal its commitment.

> Wealthier and militarily more powerful nations may have to engage in much more aggressive behavior.

This is another unhappy conclusion. First, it means that aggressive behavior during crises may have very little to do with aggressive intent; it may simply be an attempt to persuade the opponent of one's resolve to stand firm. Second, the problem will affect great powers asymmetrically, with the end result of making them appear to be bullies. While this may

[2] Beeching (1975, 115–16).

[3] The Hindi word *lūt* had been only rarely used before the sack of that city. As *India Gazette* reported, "a more complete pillage could not be conceived… the plunder ceased only when there was nothing to take or destroy" (Beeching, 1975, 116).

be true, we should keep in mind the explanation MTM offers for such behavior. Third, it would be incorrect to chide such an actor for wasting resources in compellence. Even when the threat works, it may have been necessary to make it seem all out of proportion given the strength of the opponent to make it effective.

Regardless of how aggressive these threats appear, they will not generally be associated with the gravest risks of war. After a certain point in military preparations, signaling becomes useless because even persuading the opponent that one is going to fight will not be enough to compel the opponent to concede. The greatest danger of war is when a nation has enough wealth to use risky military threats but not quite enough to adopt a strategy that ensures compellence.[4]

We saw that military threats can be less destabilizing than other signaling mechanisms. However, we also saw that all nations arming themselves is not good: when overall militarization increases, situational stability decreases. There are, then, two contrary impulses for actors that want peace: on one hand they need to keep the level of arming down to a tolerable level but on the other they do not want to expose themselves to military coercion. The resolution of this conundrum requires one to find a "sweet spot" in balancing the two sides, and it is unlikely that such a thing can be found. In general, as long as it is possible for actors to impose their will by force, military threats will remain among the instruments of statecraft. As long as the use of force is costly and the commitment to employ it doubtful, military threats will necessarily involve an irreducible risk of violence. War is with us to stay.

[4] Brito and Intriligator (1985) find an analogous result.

Appendix A

Formalities for Chapter 2

Proposition 2.2 summarizes the results of the two general cases. For ease of exposition, I will deal with them separately here. Observe that if S_1's threat is known to be genuine (because only resolved types escalate), S_2 would resist only if she is resolved as well. The probability that she capitulates then is $F_2(v_2^*)$. Since S_1 would then escalate when the condition in (2.3) is satisfied, we can simplify it:

$$v_1^* \leq \frac{(1 - F_2(v_2^*))a_1}{F_2(v_2^*)}. \tag{GT}$$

When (GT) is satisfied with strict inequality, some resolved types of S_1 will be unwilling to escalate even if doing so convinced S_2 that the threat is genuine.

PROPOSITION A.1 (Genuine Threat). *If, and only if,* (GT) *is satisfied, the following assessment constitutes the unique equilibrium. S_1 threatens if $v_1 \geq \hat{v}_1$, and fights if $v_1 \geq v_1^*$, where*

$$\hat{v}_1 = \frac{(1 - F_2(v_2^*))c_1}{p + (1 - p)F_2(v_2^*)} < \overline{v}_1, \tag{A.1}$$

and v_1^* is in (2.1)*. S_2 resists if $v_2 \geq v_2^*$, with beliefs updated by Bayes rule.*

Proof of Proposition A.1 (Sketch) By subgame perfection, S_1 attacks when resisted if, and only if, $v_1 \geq v_1^*$. In the postulated equilibrium only resolved types escalate, and so S_2 believes that resistance means certain war. This now implies that only resolved types would resist; that is, $v_2 \geq v_2^*$. From S_1's perspective, the probability that S_2 will capitulate is then $F_2(v_2^*)$. Let $\hat{v}_1$ solve $F_2(v_2^*)\hat{v}_1 + (1 - F_2(v_2^*))w_1(\hat{v}_1) = 0$. To see that this equation has a unique solution, observe that by (GT), $F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))(-a_1) \leq 0$. Since $w_1(v_1^*) = -a_1$, it follows that $\hat{v}_1 \geq v_1^* \Rightarrow w_1(\hat{v}_1) \geq -a_1$, which in turn implies that there exists $\hat{v}_1 \geq v_1^*$ that solves the equation, with the solution specified in (A.1). This now means that all $v_1 \geq \hat{v}_1$ will prefer to escalate rather than appease. □

If (GT) is satisfied but $\hat{v}_1 \geq \overline{v}_1$, then no type would strictly prefer threatening to appeasement. This implies that S_1 will never escalate the crisis, which puts S_2's decision off the equilibrium path, and so Bayes rule cannot help: one cannot condition on an event that is not supposed to occur. However, it turns out that S_2's beliefs in this case are immaterial. The most advantageous (from S_1's perspective) conclusion she could draw following escalation is that S_1's type is among the resolved ones, in which case only $v_2 > v_2^*$ would resist. This means that the probability that S_2 would resist is *at least* $F_2(v_2^*)$ even with these generous beliefs. But this is exactly the probability we used in (A.1), and since it has no solution

that is less than $\overline{v}_1$, all types of S_1 are deterred. In other words, regardless of S_2's beliefs, appeasement is assured, and the status quo is peacefully revised with certainty.

PROPOSITION A.2 (Assured Appeasement). *If* (GT) *holds and $\hat{v}_1 \geq \overline{v}_1$, then the following strategies and beliefs form the essentially unique equilibrium. S_1 never escalates and fights if $v_1 \geq v_1^*$, where $\hat{v}_1$ is defined in* (A.1)*. S_2 resists if $v_2 \geq v_2^*$, and believes that if escalation occurs, $v_1 \geq v_1^*$.*

This equilibrium is essentially unique because of the latitude we have in specifying S_2's beliefs: if we assign more optimistic beliefs, then even lower-valuation types would act. However, this has no bearing on S_1's equilibrium behavior, and all such equilibria are outcome-equivalent. Assured appeasement is possible because Assumption 2.1 rules out the possibility that S_1 would ever escalate if he is certain of resistance. If we relax it to allow that his valuation may be high enough for this to happen, then this equilibrium will not exist.

Finally, suppose that (GT) is not satisfied, so $\hat{v}_1 < v_1^* < \overline{v}_1$. The following result states the solution in this case.

PROPOSITION A.3 (Bluffing). *If, and only if,* (GT) *is not satisfied, the following assessment constitutes the unique equilibrium. S_1 threatens if $v_1 \geq \hat{v}_1$, and fights when resisted if $v_1 \geq v_1^*$, where*

$$\hat{v}_1 = \frac{(1 - F_2(\hat{v}_2))a_1}{F_2(\hat{v}_2)} \tag{A.2}$$

and v_1^* is in (2.1)*. S_2 resists if $v_2 \geq \hat{v}_2$, where*

$$\hat{v}_2 = \frac{(1 - F_1(v_1^*))c_2 - (1 - F_1(\hat{v}_1))a_2}{1 - F_1(\hat{v}_1) - p(1 - F_1(v_1^*))},$$

with beliefs updated by Bayes rule.

Proof of Proposition A.3 (Sketch) As before, S_1 attacks if, and only if, $v_1 \geq v_1^*$. Since all $v_1 \geq \hat{v}_1$ escalate but $\hat{v}_1 \leq v_1 \leq v_1^*$ are bluffers, S_2's posterior belief by Bayes rule is $G_1(v_1^*) = \int_{\hat{v}_1}^{v_1^*} f_1(x)\,\mathrm{d}x / \int_{\hat{v}_1}^{\overline{v}_1} f_1(x)\,\mathrm{d}x = (F_1(v_1^*) - F_1(\hat{v}_1))/(1 - F_1(\hat{v}_1))$. Using this in (2.2) and simplifying yields $\hat{v}_2$ specified in the proposition. This now implies that from S_1's perspective, the probability that S_2 will capitulate is $F_2(\hat{v}_2) > 0$. By definition, $\hat{v}_1$ is the type that is indifferent between escalation and appeasement, and so it solves $F_2(\hat{v}_2)\hat{v}_1 + (1 - F_2(\hat{v}_2))(-a_1) = 0$. To see that this equation has a unique solution, observe that the left-hand side is increasing in $\hat{v}_1$ and $\hat{v}_2$ is a function of $\hat{v}_1$ with $\hat{v}_1 = v_1^* \Rightarrow \hat{v}_2 = v_2^*$. Since (GT) is not satisfied, it follows that $F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))(-a_1) > 0$, which implies that for some $\hat{v}_1 < v_1^*$ the equation will have the solution specified in (A.2). All $v_1 > \hat{v}_1$ strictly prefer to escalate rather than appease. □

PROPOSITION A.4 (Sunk Costs / Fearon (1997)). *Assume that* (GT) *is not satisfied. Choose any $\hat{v}_1 \in (0, v_1^*)$, where v_1^* is defined in* (2.1)*. Let*

$$\hat{v}_2 = \frac{(1 - F_1(v_1^*))c_2 - (1 - F_1(\hat{v}_1))a_2}{(1 - F_1(\hat{v}_1)) - p(1 - F_1(v_1^*))},$$

and let $m^* = F_2(\hat{v}_2)\hat{v}_1 + (1 - F_2(\hat{v}_2))(-a_1)$. The following assessment constitutes an equilibrium of the sunk-costs game. Strategy for S_1: if $v_1 < \hat{v}_1$, appease with $m = 0$ and

capitulate if resisted; if $v_1 \geq \hat{v}_1$, threaten with m^* and fight when resisted if, and only if, $v_1 \geq v_1^*$. Strategy for S_2: act if $m \neq m^*$ or $v_2 \geq \hat{v}_2$; capitulate if $m = m^*$ and $v_2 < \hat{v}_2$. Beliefs for S_2: if $m = m^*$, update to believe that v_1 is distributed by F_1 on $[\hat{v}_1, \overline{v}_1]$, and if $m \neq m^*$, update to believe that v_1 is distributed by F_1 on $[0, v_1^*]$.

Proof See Fearon's paper. The derivation of m^* in that proof has a mistake: since $\hat{v}_1$ is not resolved when he escalates, if S_2 resists, he will capitulate. This means that his expected payoff from escalation is $F_2(\hat{v}_2)\hat{v}_1 + (1 - F_2(\hat{v}_2))(-a_1) - m^*$ rather than $F_2(\hat{v}_2)\hat{v}_1 + (1 - F_2(\hat{v}_2))(p\hat{v}_1 - c_1) - m^*$ as given in the original proof. This latter specification assumes that $\hat{v}_1$ will fight when resisted, which is only true in the intuitive equilibrium, but not in any of the others, all of which admit bluffing. □

Proposition 2.3 in the text summarizes the following corollary:

COROLLARY A.1 (Intuitive Equilibrium). *If $\overline{v}_1$ is sufficiently large, then only $\hat{v}_1 = v_1^*$ is intuitive. In the unique intuitive equilibrium, $\hat{v}_2 = v_2^*$ and $m^* = F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))(-a_1)$.*

If (GT) is satisfied, then v_1^* cannot profit by escalating even if doing so would convince S_2 that the threat is genuine. This now implies that for some $\hat{v}_1 > v_1^*$, the condition must be violated with equality. Recall that if (GT) is satisfied, then $F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))(-a_1) < 0$, and since for $m^* > 0$ we require this to be positive, no $m^* > 0$ exists that would make v_1^* willing to escalate. All $v_1 \geq \hat{v}_1$ would then be willing to escalate without sinking any costs because even by itself the failure to appease will be sufficient to persuade S_2 that they are genuine defenders.[1] In other words, sinking costs does not do anything for S_1 in this case, and the result is equivalent to Proposition A.1.

Suppose now that only genuine defenders threaten in equilibrium. This implies that v_1^*'s expected payoff from escalating is less than zero, and that $G_1(v_1^*) = 0$ as well. Hence, $\hat{v}_2 = v_2^*$, and $F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))[m(pv_1^* - c_1) + (1 - m)(-a_1)] = F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))(-a_1) < 0$, which, of course, is condition (GT). Since all types who escalate will fight for sure, the smallest type for which escalation is profitable is $\hat{v}_1$ as defined in (A.2). As in the sunk-cost model, this means that *any* risk that S_1 generates will reveal that he is resolved, and he will therefore pick the smallest such risk. Again, the game form is limiting because it does not allow threats with $m = 0$ but if it did, this is what $v_1 \geq \hat{v}_1$ would pick.[2] If (GT) is satisfied, then generating risk, like sinking costs, does not do anything for S_1, and the result is equivalent to Proposition A.1.

Proposition 2.4 in the text summarizes the following formal proposition.

[1] Technically, since the game form does not allow escalation without sinking some costs, there will be no equilibrium here. This happens because in equilibrium S_1 must pick the smallest $m > 0$, which does not exist because we can get arbitrarily close to zero. There are two ways we can deal with this non-existence problem. First, we can change the game form slightly to allow S_1 to appease (as before), followed by a choice of $m \geq 0$; that is, S_1 could escalate without sinking any costs. The results do not change in this case and in equilibrium $\hat{v}_1$ is the valuation that solves (GT) with equality. All $v_1 \geq \hat{v}_1$ escalate without incurring any sunk costs. The other method is to assume that the least costly escalation involves a strictly positive cost, $m \geq \epsilon > 0$. This now implies that for the smallest-valuation type willing to escalate:
$F_2(v_2^*)\hat{v}_1 + (1 - F_2(v_2^*))[p\hat{v}_1 - c_1] - \epsilon = 0$, which means $\hat{m} = \epsilon$ and
$\hat{v}_1 = [(1 - F_2(v_2^*))c_1 + \epsilon]/[p + (1 - p)F_2(v_2^*)]$.

[2] See fn. [1] for a discussion of this technical issue.

PROPOSITION A.5 (Randomized Threat). *Let $\hat{v}_2$ be defined in* (2.4)*. If, and only if,* (GT) *is not satisfied in the risky threat game, the following strategies constitute a unique equilibrium. S_1 escalates with $m = 1$ if $v_1 \geq \hat{v}_1$, and fights if $v_1 \geq v_1^*$. S_2 resists if $v_2 \geq \hat{v}_2$ and capitulates otherwise. If S_2 observes any $m < 1$, she updates to believe that S_1 is unresolved, otherwise she updates to believe that v_1 is distributed by F_1 on the interval $[\hat{v}_1, \overline{v}_1]$.*

Proof For any $v_1 > v_1^*$, the choice of m is irrelevant if S_2 resists because they all strictly prefer war to capitulation so they will maximize the probability of S_2's capitulation by using $m = 1$, so $\hat{v}_2 = v_2^*$, and so only resolved challengers would resist. Because (GT) is not satisfied, there exists $\hat{v}_1 < v_1^*$ for which escalation with $m = 1$ is equivalent to appeasement:

$$\hat{v}_1 = \frac{(1 - F_2(v_2^*))c_1}{p + (1-p)F_2(v_2^*)}. \tag{A.3}$$

What is the smallest valuation that would cause S_2 resist if she is certain the defender is not resolved? Using $G_1(v_1^*) = 1$ in (2.4) yields:

$$\hat{v}_2 = \frac{mc_2 - a_2}{1 - mp}. \tag{A.4}$$

Because choosing any $m < 1$ leads to $G_1(v_1^*) = 1$, any unresolved S_1 will escalate with m that solves the following program:

$$\max_m \Big\{ F_2(\hat{v}_2)v_1 + (1 - F_2(\hat{v}_2))[m(pv_1 - c_1) + (1-m)(-a_1)] \Big\},$$

with $\hat{v}_2$ defined in (A.4). It is not difficult to show that this expression is strictly increasing in m, and therefore any unresolved S_1 will pick the highest possible escalation.This now implies that $v_1 \in [\hat{v}_1, v_1^*]$ will mimic the resolved types and choose the riskiest escalation, and all $v_1 < \hat{v}_1$ appease. □

It is worth noting that S_1 may be bluffing in this equilibrium: all $v_1 \in [\hat{v}_1, v_1^*]$ would not actually fight if they had to choose between war and capitulation. S_2's beliefs are especially interesting. With $m = 1$, it is irrelevant what she thinks about her opponent's resolve: if she resists, war is certain no matter how much S_1 values the good. Although Proposition 2.4 specifies the beliefs after equilibrium escalation for the sake of completeness, any set of beliefs would work.

If (GT) is satisfied, then there is no point in incurring any audience costs: threats are already credible. However, since the defender does not pay the audience costs in equilibrium, any level of these costs can be supported in equilibrium. To see this, recall that if this condition is satisfied, then $F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))(-a_1) < 0$. Since $v_1^*(m^*) < v_1^*$ and $-a_1 - m^* < -a_1$, it follows that $F_2(v_2^*)v_1^*(m^*) + (1 - F_2(v_2^*))(-a_1 - m^*) < 0$ as well, and so no $m^* > 0$ exists such that all $v_1 \geq v_1^*(m^*)$ would be willing to escalate. That is, if this condition is satisfied, then the least-resolved type v_1^* would be unwilling to escalate even without incurring any audience costs and even if S_2 believes his threat to be genuine. Since matters are only more difficult for $v_1^*(m^*) < v_1^*$, there would exist no positive signal that would make him willing to escalate either. In this situation, the equilibrium would involve credible escalation for all $v_1 \geq \hat{v}_1$, where $\hat{v}_1 > v_1^*$ is the smallest-valuation type for whom credible escalation is profitable (if we wish this to be a costless escalation, the technical

considerations from fn. [1] apply). The ability to incur audience costs, just like sinking costs, is irrelevant in that case.

Consider now the situation where (GT) is not satisfied. If S_1 escalates with m^* that convinces S_2 that the threat is genuine, the probability that she will capitulate is $F_2(v_2^*)$, and since any type willing to send this signal must be willing to fight, it follows that $v_1 \geq v_1^*(m^*)$. The smallest-valuation type that would be willing to escalate must be indifferent between choosing m^* and appeasement. Noting that $w_1(v_1^*(m^*)) = -a_1 - m^*$, we obtain m^* by solving $F_2(v_2^*)v_1^*(m^*) + (1 - F_2(v_2^*))(-a_1 - m^*) = 0$, or:

$$m^* = \frac{F_2(v_2^*)c_1}{p + (1-p)F_2(v_2^*)} - a_1. \tag{A.5}$$

Proposition 2.5 in the text can be stated formally as follows (for the proof, refer to Fearon's paper):

PROPOSITION A.6 (Tying Hands / Fearon (1994a)). *Assume that* (GT) *is not satisfied. Let* m^* *be defined in* (A.5)*,* v_1^* *in* (2.5)*, and* v_2^* *in* (2.1)*. The following strategies form the essentially unique equilibrium of the tying-hands game. Strategy for* S_1*: if* $v_1 < v_1^*$ *appease and capitulate if resisted; if* $v_1 \geq v_1^*$*, escalate with* m^* *and fight if resisted. Strategy for* S_2*: if* $m < m^*$ *or* $v_2 > v_2^*$*, resist; if* $m \geq m^*$ *or* $v_2 \leq v_2^*$*, capitulate. Beliefs for* S_2*: if* $m < m^*$*, update to believe that* v_1 *is distributed by* F_1 *on* $[0, v_1^*]$*; if* $m \geq m^*$*, update to believe that* v_1 *is distributed by* F_1 *on* $[v_1^*, \overline{v}_1]$.

To see that the escalation threshold remains the same whether S_1 escalates with a randomized threat or by tying hands, observe that $v_1^*(m^*) = \hat{v}_1$ from (A.3). To see that the escalation threshold is lower than in the sinking-costs model, observe that: $v_1^*(m^*) = v_1^* - m^*/p < v_1^*$.

Appendix B

Formalities for Chapter 3

B.1 Modeling Military Threats

Proof of Claim 3.1 The claim is that if (M_1, M_2) and (M_1', M_2') are such that $p = p'$, then $M < M' \Rightarrow p(m) > p'(m)$ for any $m > 0$. Simplifying $p(m) > p'(m)$ yields $m(M' - M) > MM_1' - M_1M'$. The right-hand side is zero because $p = p' \Rightarrow M_1/M = M_1'/M' \Rightarrow MM_1' = M_1M'$. This means that the inequality is satisfied for any $m > 0$ if, and only if, $M < M'$. □

B.2 Threats With Complete Information

Proof of Proposition 3.1 The solution is trivial when $c_1 - a_1 \geq v_1$ because S_1 can never establish a credible threat, so in the unique SPE, S_2 would resist regardless of his mobilization level. Making a threat would then lead to certain capitulation in the endgame, which is strictly worse than appeasement. Hence, S_1 will appease immediately.

Assume now that $c_1 - a_1 < v_1$. Recall from the argument in the text that if S_1 threatens in equilibrium, he must be resolved at the chosen mobilization level m, so:

$$m \geq \underline{m} = \frac{(c_1 - a_1)M_2}{v_1 - (c_1 - a_1)} - M_1,$$

where $\underline{m}$ satisfies (CR_1) with equality. Consider any SPE in which S_1 escalates and S_2 capitulates for sure. Let $\hat{m}$ satisfy (CR_2) with equality:

$$\hat{m} = \frac{v_2M_2}{c_2 - a_2} - M,$$

and define $\overline{m} = \max\{\underline{m}, \hat{m}\}$. Because $p(m)$ increases in m, it follows that any $m \geq \overline{m}$ would achieve compellence too and since mobilization is costly, S_1 must mobilize uniquely at $m = \overline{m}$ in any SPE in which he escalates to obtain assured compellence.[1]

Consider now any SPE in which S_1 escalates and S_2 resists for sure. The optimal mobilization level is:

$$m^* = \text{argmax}_m \; p(m)v_1 - c_1 - m \Rightarrow m^* = \max\left\{0, \sqrt{v_1M_2} - M\right\},$$

[1] Strictly speaking, S_2 will be indifferent between capitulation and resistance with certain war when $\overline{m} = \hat{m}$, which means that she could resist with certainty or even randomize. However, there can be no SPE in which she resists or randomizes when indifferent: even the slightest deviation by S_1 to $\hat{m} > m$ would cause her to capitulate with certainty, which means that S_1's strategy that allows her to mix cannot be optimal. This is not an issue when $\overline{m} = \underline{m}$ because in that case S_2's payoff from war is strictly worse than the payoff from capitulation.

and the optimal war payoff is $W_1(m^*) = p(m^*)v_1 - c_1 - m^*$. Since S_1 escalates in this SPE, it follows that $W_1(m^*) > 0$, which implies $m^* \geq \underline{m}$, as required. That is, S_1 is resolved at the optimal war-fighting mobilization level.

Simplifying $\hat{m} \leq \underline{m}$ yields:

$$\frac{c_1 - a_1}{v_1} + \frac{c_2 - a_2}{v_2} \geq 1, \tag{PEACE}$$

which is the condition we shall use to distinguish between the two general cases. Suppose first that (PEACE) is satisfied, so $\overline{m} = \underline{m}$. War cannot happen in equilibrium because S_2 would capitulate at the smallest allocation that makes S_1 resolved. S_1 would not mobilize $m < \overline{m}$ because S_2 would resist any such allocation knowing that S_1 would not fight, making appeasement preferable for S_1. Thus, if $v_1 > \overline{m}$, then S_1 mobilizes $\overline{m}$ and forces S_2 to capitulate for sure; otherwise, he appeases her immediately.

Suppose now that (PEACE) is not satisfied, so $\overline{m} = \hat{m} > \underline{m}$. In this case, S_1 can ensure S_2's capitulation by mobilizing $\overline{m}$, or he can appease, or he can allocate some $m \in [\underline{m}, \overline{m})$ and fight a certain war. S_1 would not mobilize $m < \underline{m}$ because doing so would lead to capitulation. If S_1 were to escalate for war in equilibrium, it must be the case that fighting is better than appeasement, $W_1(m^*) > 0$, and also better than compellence, $W_1(m^*) > v_1 - \overline{m}$; these yield $\underline{m} \leq m^* < \overline{m}$. Hence, if $W_1(m^*) > \max\{0, v_1 - \overline{m}\}$, then S_1 mobilizes credibly at the unique level $m^* \in [\underline{m}, \overline{m})$, S_2 resists, and S_1 attacks. The outcome is certain war. If $0 > \max\{W_1(m^*), v_1 - \overline{m}\}$, then appeasement is strictly better than any escalation. Finally, if $v_1 - \overline{m} \geq \max\{0, W_1(m^*)\}$, then ensuring S_2's capitulation the best option for S_1, and consequently he mobilizes to obtain it. □

B.3 Threats Under Uncertainty

B.3.1 Optimal Credible Mobilization

We begin by defining several crucial military allocations by S_1 which differ depending on the strategic effect they have on S_2's behavior.

DEFINITION B.1. A mobilization level, m, is *credible* if, and only if, S_2 believes that $v_1 \geq v_1^*(m)$ upon observing it.

If a credible mobilization causes S_2 to capitulate for sure, the result is *assured compellence*, if it causes her to capitulate with positive probability, the result is *coercion*, and if it causes her to resist for sure, the result is *war*. A credible mobilization is *optimal* if it is the best allocation for type v_1 provided that S_2 will believe that he is resolved. When mobilization is credible, $G_1(v_1^*(m)) = 0$, and so S_2 will capitulate if $v_2 < v_2^*(m) = (c_2 - a_2)/p_2(m)$. Since we know that her valuation is in the interval $[t, u]$, we now have to distinguish three possibilities:

(i) $v_2^*(m) \leq t$, which implies she will resist for sure—in this case S_1 would *mobilize for war* by maximizing $p_1(m)v_1 - c_1 - m$;

(ii) $v_2^*(m) \in (t, u)$, which implies that she will capitulate with probability $G_2(v_2^*(m)$—in

this case, S_1 would *mobilize for coercion* by maximizing

$$G_2\left(\frac{c_2-a_2}{p_2(m)}\right)v_1+\left[1-G_2\left(\frac{c_2-a_2}{p_2(m)}\right)\right]\left[p_1(m)v_1-c_1\right]-m;$$

(iii) $v_2^*(m)\geq u$, which implies she will capitulate for sure—in which case S_1 would *mobilize for assured compellence* at the unique level that solves $G_2(v_2^*(m))=1$.

Solving these optimizations and equation yields the following:

DEFINITION B.2. A credible mobilization level is

- *optimal for war* for type v_1 if, and only if, it is defined by:

$$m^*(v_1)=\sqrt{M_2v_1}-M; \tag{B.1}$$

- *optimal for coercion* for type v_1 if, and only if, it is defined by:

$$\widehat{m}(v_1)=M_2\sqrt{\frac{uv_1}{\zeta}}-M, \tag{B.2}$$

  where $\zeta=(u-t)M_2-c_1(c_2-a_2)$; and
- *optimal for assured compellence* if, and only if, it is defined by:

$$\overline{m}=\frac{uM_2}{c_2-a_2}-M. \tag{B.3}$$

Deriving the necessary and sufficient conditions for these quantities to be valid is an extremely long and tedious process so is omitted here. (The replication package on the website has the full 60-page appendix.) The definition assumes that the mobilization levels are credible. We now investigate what it takes to make them such.

B.3.2 Minimum Credible Mobilization

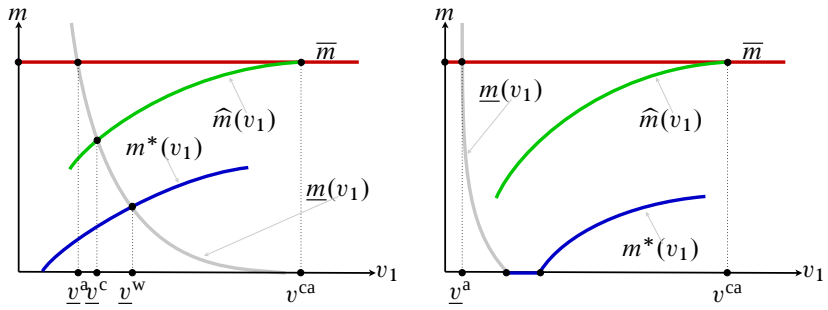
Any type v_1 is indifferent between attacking and capitulating at the final node when $v_1=v_1^*(\underline{m}(v_1))$. Since higher mobilizations can never undermine his commitment, it follows that v_1 would fight for any $m>\underline{m}$.

DEFINITION B.3. The *minimum credible* mobilization level for type $v_1>c_1-a_1$ is defined as:

$$\underline{m}(v_1)=\frac{(c_1-a_1)M_2}{v_1-(c_1-a_1)}-M_1 \quad \text{for all} \quad v_1>c_1-a_1, \tag{B.4}$$

and is the smallest allocation at which v_1 will attack.

The function $\underline{m}$ is discontinuous at c_1-a_1 and undefined for any $v_1\leq c_1-a_1$. It is possible to be resolved at the pre-crisis allocations, and the smallest-valuation type for which this is so solves $\underline{m}(v_1)=0$. I shall call this the *least inherently resolved* type. All higher valuation types have an inherently credible commitment. If the least inherently resolved type is smaller than $\phi^*=M^2/M_2$, the least-valuation type that could mobilize for war, i.e., the one that solves $m^*(\phi^*)=0$, then the types in that range would escalate without militarizing



(a) Militarized Escalation (b) Verbal Escalation

Figure B.1 Optimal and Minimum Credible Militarized Allocations.

the crisis even if S_2 is expected to resist for sure. The necessary and sufficient condition for this to be the case is:

$$c_1 - a_1 < \frac{MM_1}{M_2}. \tag{VF}$$

I shall refer to this as a *verbal escalation to war* to emphasize that it involves no additional mobilization. Analogously, if the least inherently resolved type is smaller than the least-valuation type that could mobilize for coercion, i.e., the one that solves $\widehat{m}(v_1) = 0$, then the types in that range would escalate without militarizing the crisis when S_2 is expected to capitulate with positive probability. The necessary and sufficient condition for this to be the case is:

$$c_1 - a_1 < \frac{\zeta MM_1}{uM_2^2}. \tag{VC}$$

I shall refer to this as a *verbal escalation to coercion*. If (VC) is satisfied, then (VF) must also be satisfied. The converse, of course, is not necessarily the case.

B.3.3 Credibility Cut-Point Types

The optimal credible mobilization levels are continuous, non-decreasing, and strictly positive whenever they exist, which implies that if $\underline{m}(v_1)$ intersects any one of them, it will do so at most once. The type at which the minimum credible mobilization level intersects the optimal mobilization level is a *credibility cut-point type* because it bisects the type space: the optimal mobilization is not credible for all types to the left of the cut-point but is credible for all types to the right. I will use the superscript mnemonic to denote the mobilization type of interest: 'w' for war, 'c' for coercion, and 'a' for assured compellence. For instance, $\underline{v}^a$, is the type for whom $\underline{m}(\underline{v}^a) = \overline{m}$. Thus, mobilizing $\overline{m}$ is not credible for all $v_1 < \underline{v}^a$ but is credible for all $v_1 > \underline{v}^a$. This particular type is easy to derive: $\underline{v}^a = \frac{u(c_1-a_1)}{u-(c_2-a_2)}$. The other types are derived analogously except care must be taken when verbal escalation is possible.

Figure B.1(a) shows one possibility where the minimum credible mobilization level intersects all three optimal mobilization levels when (VF) is not satisfied (and therefore neither is (VC)), and so any escalation is militarized. Figure B.1(b) shows an instance where verbal escalation is possible.

B.3.4 Expected Payoffs from Credible Mobilization

Using the definitions of optimal credible mobilization levels given by (B.1), (B.2), and (B.3), we can specify the expected payoffs from war, coercion, and assured compellence as follows:

$$\mathcal{W}_1(v_1) = v_1 - 2\sqrt{M_2 v_1} - c_1 + M; \tag{B.5}$$

$$\mathcal{C}_1(v_1) = \left(\frac{1}{u-t}\right)\left[(u - t + c_2 - a_2)v_1 - 2\sqrt{u\zeta v_1} - uc_1\right] + M; \tag{B.6}$$

$$\mathcal{A}_1(v_1) = v_1 - \frac{uM_2}{c_2 - a_2} + M. \tag{B.7}$$

The following lemma is crucial in establishing the relationship among the payoffs from different strategic contingencies.

LEMMA B.1. $\mathcal{A}_1' = 1 > \mathcal{C}_1' > \mathcal{W}_1' > 0$, $\mathcal{W}_1'' > 0$, *and* $\mathcal{C}_1'' > 0$.

The proof requires straightforward algebraic manipulations. Lemma B.1 shows three things. First, it establishes that the payoffs from optimal war, optimal coercion, and optimal compellence are strictly increasing in type. Second, it shows that they are increasing at faster rates as well. Third, it demonstrates the crucial result that the payoff from compellence increases the fastest, followed by the payoff from coercion, followed by the payoff from war. Because these functions are continuous, each pair may have *at most* one intersection, if any.

B.3.5 Escalation Cut-Point Types

There are four general strategies that S_1 can adopt, depending on what S_2 is expected to do: he can appease, mobilize for war, mobilize for coercion, or mobilize for compellence. As we have seen, Lemma B.1 implies that any two payoff functions can have at most one intersection—the type at which this intersection occurs is indifferent among the two strategies. I shall call this an *escalation cut-point type* because bisects the type space: all types to the left strictly prefer one of the strategies and all types on the right strictly prefer the other. With four strategies, there are six pairs to consider, so there are six of these escalation cut-point types.

The superscripts on these escalation cut-point types consist of two-letter mnemonics designed to indicate which two actions the type is indifferent between. The codes are as follows: 'q' (quit) for appeasement, 'w' for certain war, 'c' for coercion, and 'a' for assured compellence. For example, v^{ca} is the type that is indifferent between optimal credible coercion and assured credible compellence. It can be found by noting that $v_2^*(\widehat{m}(v^{\text{ca}})) = u$, or $\widehat{m}(v^{\text{ca}}) = \overline{m}$. Solving this yields $v^{\text{ca}} = \frac{u\zeta}{(c_2-a_2)^2}$. All $v_1 > v^{\text{ca}}$ prefer assured compellence to coercion, whereas all $v_1 < v^{\text{ca}}$ prefer the opposite. It is also easy to define v^{aq}, the type that is indifferent between assured compellence and appeasement. Solving $\mathcal{A}_1(v^{\text{aq}}) = 0$ gives us $v^{\text{aq}} = \frac{uM_2}{c_2-a_2} - M$. All $v_1 > v^{\text{aq}}$ prefer assured compellence to appeasement, whereas all $v_1 < v^{\text{aq}}$ prefer the opposite. Although conceptually straightforward, deriving the other cut-point types is quite tedious and a bit involved. The complication arises from the possibility that both war and coercion can involve verbal escalation, which necessitates checks on the exogenous parameters to establish conditions under which the various possibilities might arise. Details can be found in the replication package.

B.3.6 Equilibrium Analysis

The analysis now essentially reduces to examining various configurations of the escalation cut-point types and their relationship to the credibility cut-points. Before continuing with the analysis, however, it is worth discussing briefly two technical issues with respect to updating beliefs in equilibrium and establish some fundamental results we will use throughout.

Consistent Beliefs

As we have seen in (B.1) and (B.2), the optimal nonzero war and coercion mobilization levels are unique for each type. Intuitively, any type that uses its unique mobilization level in equilibrium would be revealing its valuation to S_2. Unfortunately, Bayes rule cannot be used to obtain that inference because there are no atoms in the distribution of types (recall that they are uniform on $[0, \overline{v}_1]$), and so the probability of any one particular type being selected from this continuum is zero. Bayes rule does not allow conditioning on zero-probability events, and consequently is undefined. Technically, we are dealing with a non-empty set that has measure zero. That is, a set of types who mobilize at the same level that is not empty but the occurrence of the set itself has an equilibrium probability of zero.

To deal with this situation, I will assume that the support of S_2's beliefs conditional on such mobilization is restricted to the set of types that mobilized at this level. This requires S_2 to infer the type for whom the given allocation level would have been optimal even though only one type would use such an allocation in equilibrium.

Plausible Beliefs

Perfect Bayesian equilibrium does not restrict much beliefs players can have after observing behavior that is not supposed to occur in equilibrium. As we have seen in the sunk-cost game in Chapter 2, this lack of specificity can permit the existence of equilibria supported by very unlikely beliefs. For instance, if S_1 were to mobilize unexpectedly a very large force relative to the equilibrium level m, then S_2 concludes that he is unresolved. To eliminate this problem, I require that upon observing the unexpectedly high mobilization, S_2 infers that S_1 is at least as committed as the least-committed type at m. Formally, if a set of types $[\underline{v}, \overline{v}]$ with $v_1^*(m) < \overline{v}$ pool on a common threat m in equilibrium, then observing any $\hat{m} > m$ should cause S_2 to believe that $v_1 \geq v_1^*(m)$.

This affects only beliefs after zero-probability events and serves to rule out some rather bizarre equilibria in which S_1 cannot signal his commitment with a larger mobilization because S_2 "threatens" to infer that he is less resolved than what he would have been at the smaller mobilization. Because larger mobilizations are always more committing than smaller ones for any given valuation, such an inference is clearly implausible. Without this restriction, S_2 can threaten S_1 to have incredible beliefs, and it makes sense to require that if actions are credible in equilibrium, then so must beliefs be as well.[2] The equilibrium I study is supported by such plausible beliefs.

[2] This is analogous, in spirit, to the restriction of beliefs Kreps and Wilson (1982) impose.

Ordering of Minimally Resolved Types

The first important result is that any type whose payoff from escalating to war is at least as good as appeasement will be resolved at his, possibly verbal, optimal escalation level. (The proof is just algebraic and is omitted.)

LEMMA B.2. *All types who can profit from war are resolved:* $\underline{v}^{w} < v^{wq}$.

When war is preferable to coercion for all types who can choose between them, potential equilibrium choices are limited to appeasement, preparation for war, and mobilization for assured compellence. What can happen in equilibrium critically depends on the credibility of these mobilizations. The following lemma establishes a fundamental relationship between these types.

LEMMA B.3. *All types resolved for war are also resolved for coercion, and all types resolved for coercion are resolved for compellence:* $\underline{v}^{a} \leq \underline{v}^{c} \leq \underline{v}^{w}$.

Proof Since $\underline{m}(\cdot)$ is continuous where it is defined and is strictly decreasing, it follows that $\underline{m}^{-1}(\cdot)$ is strictly decreasing as well. This now means that $\overline{m} \geq \widehat{m}(v_1) \geq m^*(v_1)$ implies that $\underline{m}^{-1}(\overline{m}) \leq \underline{m}^{-1}(\widehat{m}(v_1)) \leq \underline{m}^{-1}(m^*(v_1))$ for any v_1, which establishes the claim. □

This lemma implies that any type who can mobilize credibly for war could mobilize credibly for coercion (provided coercion is possible) and for compellence. Furthermore, any type who can mobilize credibly for coercion can also opt for credible assured compellence as well.

The General Cases

Any non-appeasement equilibrium action requires that its expected payoff is non-negative. This implies that the relationship between the escalation types is the fundamental differentiating aspect of the model. This suggests three basic cases to examine:

$$\text{war preparation} : v^{wq} \leq \min(v^{cq}, v^{aq}) \qquad \text{(WAR)}$$

$$\text{coercion} : v^{cq} \leq \min(v^{wq}, v^{aq}) \qquad \text{(COERCION)}$$

$$\text{assured compellence} : v^{aq} \leq \min(v^{wq}, v^{cq}) \qquad \text{(COMPELLENCE)}$$

Since these cases are mutually exclusive and exhaustive, we will examine each separately to find solutions for all possible configurations of the exogenous variables.

War Preparation

Assume (WAR) is satisfied, which means that the lowest-valuation type who will escalate in equilibrium is v^{wq}. By Lemma B.2, all types who can profitably escalate to war, $v_1 \geq v^{wq}$, are resolved, and by Lemma B.3, all of them will also be resolved for coercion and compellence. To see whether coercion will be attempted in equilibrium, observe that coercion is never optimal if $\mathcal{W}_1(v_1) > \mathcal{C}_1(v_1)$ for all $v_1 < v^{ca}$, which yields the following condition:

$$v^{wa} \geq v^{ca}. \qquad \text{(NC)}$$

If (NC) is satisfied, then credible coercion is always worse than war, and hence it will not be attempted in equilibrium.

PROPOSITION B.1. *Assume* (WAR) *and* (NC) *are satisfied. The following assessment constitutes the essentially unique equilibrium.*

Strategy for S_1: (i) if $v_1 < v^{wq}$, appease; if $v_1 \in [v^{wq}, v^{wa})$, mobilize at $m^*(v_1)$; if $v_1 \geq v^{wa}$, mobilize at $\overline{m}$; (ii) if resisted, fight if, and only if, $m \geq \underline{m}(v_1)$.

Strategy and beliefs for S_2: (i) if $m < m^*(v^{wq})$, update to believe that v_1 is uniformly distributed on $[0, v^{wq})$ and resist; (ii) if $m \in [m^*(v^{wq}), m^*(v^{wa}))$, then infer that v_1 is uniformly distributed on $[v^{wq}, \psi^*]$ if $m = 0$ or that $v_1 = (m^*)^{-1}(m)$ if $m > 0$, and resist; (iii) if $m \in [m^*(v^{wa}), \overline{m})$, infer that S_1 is resolved and resist if, and only if, $v_2 \geq v_2^*(m)$; (iv) if $m \geq \overline{m}$, update to believe that v_1 is uniformly distributed on $[v^{wa}, \overline{v}_1]$ and capitulate.

Proof It is not difficult to verify that this is an equilibrium given the configuration of the cut-points and the conditions in the proposition. The equilibrium is essentially unique because it is possible to specify alternative beliefs for allocations off the equilibrium path (e.g., $m < m^*(v^{wq})$, $m \in [m^*(v^{wa}), \overline{m})$, and $m > \overline{m}$) that rationalize the strategies. Equilibrium behavior will not be affected. □

COROLLARY B.1. *If $c_1 \leq MM_1/M_2$ is satisfied in addition to the conditions for Proposition B.1, then for some valuations S_1 escalates verbally to war.*

To see how this follows, observe that $c_1 \leq MM_1/M_2$ implies that (VF) is satisfied and that $m^*(v^{wq}) = 0$. This means that all $v_1 \geq v^{wq}$ can profit from verbal escalation. Of these, $v_1 \in [v^{wq}, \psi^*]$ will escalate verbally and $v_1 > \psi^*$ will militarize the crisis in equilibrium.

Suppose now that (NC) is not satisfied. Now all $v_1 \in [v^{wc}, v^{wa}]$ prefer coercion to fighting.

PROPOSITION B.2. *Assume* (WAR) *is satisfied but* (NC) *is not. The following assessment constitutes the essentially unique equilibrium.*

Strategy for S_1: (i) if $v_1 < v^{wq}$, appease; if $v_1 \in [v^{wq}, v^{wc})$, mobilize at $m^*(v_1)$; if $v_1 \in [v^{wc}, v^{ca})$, mobilize at $\widehat{m}(v_1)$, and if $v_1 \geq v^{ca}$, mobilize at $\overline{m}$. (ii) If resisted, fight if, and only if, $m \geq \underline{m}(v_1)$.

Strategy and beliefs for S_2: (i) if $m < m^*(v^{wq})$, resist (any beliefs would work); (ii) if $m \in [m^*(v^{wq}), m^*(v^{wc}))$, infer $v_1 = (m^*)^{-1}(m)$ and resist; (iii) if $m \in [m^*(v^{wc}), \widehat{m}(v^{wc}))$, update to believe that S_1 is resolved and resist if, and only if, $v_2 \geq v_2^*(m)$; (iv) if $m \in [\widehat{m}(v^{wc}), \overline{m})$, infer $v_1 = \widehat{m}^{-1}(m)$ and resist if, and only if, $v_2 \geq v_2^*(m)$; (v) if $m \geq \overline{m}$, update to believe that v_1 is distributed uniformly on $[v^{ca}, \overline{v}_1]$ and capitulate.

Proof As before, verification if straightforward. S_2's beliefs after out-of-equilibrium allocations, e.g., $m < m^*(v^{wq})$, $m \in [m^*(v^{wc}), \widehat{m}(v^{wc}))$, and $m > \overline{m}$, in the proposition provide one possible reasonable specification. □

Coercion

Assume (COERCION) is satisfied, which means that the lowest-valuation type who will escalate in equilibrium is v^{cq}. By Lemma B.1, this implies that no type will mobilize for certain war. Furthermore, $v^{cq} < v^{aq} \Rightarrow v^{cq} < v^{ca}$, which implies that all $v_1 \in [v^{cq}, v^{ca}]$ prefer optimal coercion to compellence in addition to appeasement. The question then reduces

to whether these types can engage in this coercion credibly. If $\underline{v}^{c} \leq v^{cq}$, then they can do so, and Lemma B.3 further implies that all $v_1 > v^{ca}$ who prefer assured compellence can mobilize at $\overline{m}$ credibly as well. Hence, the condition:

$$\underline{v}^{c} \leq v^{cq} \tag{CC}$$

can be used to distinguish between cases where coercion is credible for all escalation types at their optimal coercive allocations. Let α be the smallest valuation type whose optimal coercive mobilization does not cause certain war: $\widehat{m}(\alpha) = m*(\beta)$, where β solves $v_2^*(m^*(\beta)) = t$.

PROPOSITION B.3. *Assume* (COERCION) *and* (CC) *are satisfied. The following assessment constitutes the essentially unique equilibrium.*

Strategy for S_1: (i) if $v_1 < v^{cq}$, appease; if $v_1 \in [v^{cq}, v^{ca})$, mobilize at $\widehat{m}(v_1)$; if $v_1 \geq v^{ca}$, mobilize at $\overline{m}$. (ii) If S_2 resists, fight if, and only if, $m \geq \underline{m}(v_1)$.

Strategy and beliefs for S_2: (i) if $m < \widehat{m}(\alpha)$, resist (any beliefs would work); (ii) if $m \in [\widehat{m}(\alpha), \widehat{m}(v^{cq})$, update to believe that S_1 is resolved and resist if, and only if, $v_2 \geq v_2^*(m)$; (iii) if $m \in [\widehat{m}(v^{cq}), \overline{m})$, infer $v_1 = \widehat{m}^{-1}(m)$ and resist if, and only if, $v_2 \geq v_2^*(m)$; (iv) if $m \geq \overline{m}$, update to believe that v_1 is uniformly distributed on $[v^{ca}, \overline{v}_1]$ and capitulate.

Proof Given the configuration of the cut-points and the conditions, it is not difficult to verify the proposition. As before, care must be taken with S_2's beliefs after out-of-equilibrium mobilizations like $m < \widehat{m}(v^{cq})$ and $m > \overline{m}$. It is possible to find other beliefs that would support the strategies. □

COROLLARY B.2. *Assume the conditions for Proposition B.3 are satisfied and let $y = [u - (c_2 - a_2)]M - tM_1$. If*

$$c_1 < \frac{y(u-t)MM_2}{(c_2 - a_2)(y + uM_2)M - utM_2^2}$$

holds, then for some valuations S_1 escalates verbally for coercion.

Suppose now that (CC) is not satisfied. This is a very interesting case because it raises the specter of bluffing: $v_1 \in [v^{cq}, v^{ca}]$ want to mobilize for optimal coercion provided S_2 finds it credible, but $v_1 \in [v^{cq}, \underline{v}^{c})$ are, in fact, bluffing because they will not be resolved at their optimal coercive levels. Since $\widehat{m}(\cdot)$ is uniquely optimal, an attempt to use such allocation by these types would reveal their lack of resolve, causing S_2 to resist for sure, which in turn would destroy their incentive to mobilize in the first place.

Suppose that, upon observing any $m < \widehat{m}(\underline{v}^{c})$, S_2 inferred that S_1 was not resolved. This would presumably deter any attempts by these unresolved types to mobilize at their optimal coercive levels. Unfortunately, this is not enough to sustain an equilibrium. To see that, observe that $\underline{v}^{c} > v^{cq}$ means that the payoff from optimal coercion is strictly positive for $\underline{v}^{c}$, which implies that if some type $v_1 < \underline{v}^{c}$ but close enough to $\underline{v}^{c}$ mimicked $\underline{v}^{c}$ and allocated $\widehat{m}(\underline{v}^{c}) > \widehat{m}(v_1)$ instead of his own optimal coercive level, he would obtain a strictly positive payoff. In other words, such a type cannot be made to capitulate if S_2 were to believe $\underline{v}^{c}$'s mobilization level to be credible. This, in turn, destroys the credibility of $\underline{v}^{c}$'s own mobilization, and the supposed equilibrium unravels.

Suppose now that given these bluffers who are mimicking his allocation, $\underline{v}^{c}$ deviated slightly to some $m > \widehat{m}(\underline{v}^{c})$ and doing so convinced S_2 that he was resolved. Although he would incur the cost of this higher allocation, he will also reap the benefit of increased probability of obtaining the good peacefully. The latter always outweighs the former because the jump in the probability is discontinuous and occurs for an arbitrarily small increase in m. In other words, if $\underline{v}^{c}$ can persuade S_2 that he is resolved with an arbitrarily small increase in his allocation, it will be beneficial to do so. Hence, in equilibrium bluffing cannot occur as long as there is a discrepancy in S_2's capitulation probability depending on her belief about S_1's resolve.

Figure 3.4 depicts this situation. Although all $v_1 > v^{cq}$ prefer optimal coercion to appeasement, none of the types with valuations $v_1 \in [v^{cq}, \underline{v}^{c})$ will be resolved at $\widehat{m}(v_1)$ provided that S_2 infers their valuations (which she must in equilibrium). Take some $\overline{\tau}$ such that $\underline{\tau} \in (v^{cq}, \underline{v}^{c})$ is resolved at $\widehat{m}(\overline{\tau})$ and is indifferent between escalating using that level and appeasing. In other words, $\underline{\tau}$ is the type whose expected payoff from mobilizing at $\overline{\tau}$'s optimal coercive level is zero and he is resolved at that level. If S_2 observes $\widehat{m}(\overline{\tau})$, S_2 will infer that v_1 is uniformly distributed on $[\underline{\tau}, \overline{\tau}]$, and is resolved.

We now derive the bounds on the range of pooling types. Using the facts that $\underline{\tau}$ is minimally resolved at $\widehat{m}(\overline{\tau})$ and that he is indifferent between escalation with that mobilization and appeasement, we obtain:

$$\underline{\tau} = \frac{\sqrt{b^2 + 4(c_2 - a_2 - t)(c_1 - a_1)\left[(u - t)M - ua_1\right]} - b}{2(c_2 - a_2 - t)}, \tag{B.8}$$

where $b = [(u - t)M_1 - ua_1 + tc_1 + a_1(c_2 - a_2 - t)]$, and:

$$\overline{\tau} = \frac{\zeta \underline{\tau}^2}{u\left[\underline{\tau} - (c_1 - a_1)\right]^2}. \tag{B.9}$$

Observe now that as $\overline{\tau}$ approaches v^{ca}, $\underline{\tau}$ approaches v^{aq}, and in fact it must be the case that $\overline{\tau} = v^{ca} \Leftrightarrow \underline{\tau} = v^{aq}$. In this case the pooling types all choose $\overline{m}$, and their payoffs are fully described by $\mathcal{A}_1(v_1)$. Since $\underline{\tau}$ is the least-resolved type willing to escalate at $\overline{m}$, it follows that he must be equivalent to v^{aq}. (This is easy to see in Figure 3.4.)

This now implies that pooling can be credible only if v^{aq} himself can escalate credibly with $\overline{m}$. In other words, $\underline{v}^{a} \leq v^{aq}$ is a necessary condition for resolved types to induce credibility in the manner described above. This makes the condition:

$$\underline{v}^{a} \leq v^{aq} \tag{NB}$$

required for the existence of this equilibrium. As we shall see in the next section, (NB) is a general condition that is sufficient to rule out any equilibrium bluffing. We can now state the result succinctly.

PROPOSITION B.4. *Assume* (COERCION) *is satisfied but* (CC) *is not. If* (NB) *is also satisfied, the following assessment constitutes the essentially unique equilibrium.*

Strategy for S_1: (i) if $v_1 < \underline{\tau}$, appease; if $v_1 \in [\underline{\tau}, \overline{\tau}]$ mobilize at $\widehat{m}(\overline{\tau})$; if $v_1 \in (\overline{\tau}, v^{ca})$, mobilize $\widehat{m}(v_1)$; if $v_1 \geq v^{ca}$ mobilize at $\overline{m}$. (ii) If S_2 resists, fight if, and only if, $m \geq \underline{m}(v_1)$.

Strategy and beliefs for S_2: (i) if $m < \widehat{m}(\overline{\tau})$, update to believe that v_1 is uniformly distributed

on $[0, \underline{\tau})$ *and resist regardless of valuation; (ii) if* $m = \widehat{m}(\overline{\tau})$*, update to believe that* v_1 *is uniformly distributed on* $[\underline{\tau}, \overline{\tau}]$*, and resist if, and only if,* $v_2 \geq v_2^*(m)$*; (iii) if* $m \in (\widehat{m}(\overline{\tau}), \overline{m})$*, infer* $v_1 = \widehat{m}^{-1}(m)$ *and resist if, and only if,* $v_2 \geq v_2^*(m)$*; (iv) if* $m \geq \overline{m}$*, update to believe that* v_1 *is uniformly distributed on* $[v^{\text{ca}}, \overline{v}_1]$*, and capitulate regardless of valuation.*

Proof Most of the legwork is done in the derivation of the pooling range in the text. It is not hard to verify the proposition given the beliefs. Care must be taken with out-of-equilibrium mobilizations like $m < \widehat{m}(\overline{\tau})$ and $m > \overline{m}$. Figure 3.4 helps with visualizing the behavior of the payoff functions. □

Propositions B.3 and B.4 cover all possible configurations when (COERCION) obtains except when (NB) is not satisfied. As it turns out, this exception is a special case of the more general situation where (NB) fails irrespective of whether (COERCION) obtains or not. It is the subject of the following section.

Assured Compellence

Assume (COMPELLENCE) is satisfied, which means that the lowest-valuation type who will escalate in equilibrium is v^{aq}. If (NB) is satisfied, then $v_1 \geq v^{\text{aq}}$ can escalate credibly for assured compellence. Lemma B.1 implies that this yields a payoff strictly better than either war or coercion under the assumptions of this configuration. Therefore, all such types mobilize credibly for assured compellence in equilibrium, which yields a very simple solution.

PROPOSITION B.5. *Assume* (COMPELLENCE) *and* (NB) *are satisfied. The following assessment constitutes the essentially unique equilibrium.*

Strategy for S_1*: (i) if* $v_1 < v^{\text{aq}}$*, appease; if* $v_1 \geq v^{\text{aq}}$*, mobilize at* $\overline{m}$*; (ii) if resisted, fight if, and only if,* $m \geq \underline{m}(v_1)$*.*

Strategy and beliefs for S_2*: (i) if* $m < \overline{m}$*, update to believe that* v_1 *is resolved and resist if, and only if,* $v_2 \geq v_2^*(m)$*; (ii) if* $m \geq \overline{m}$*, update to believe that* v_1 *is resolved and capitulate.*

Proof All allocations except $\overline{m}$ are off the equilibrium path and S_2's beliefs for these allocations are the most pessimistic possible. It is easy to verify that S_1 would not deviate even under these favorable conditions, which implies that no alternative beliefs can sustain a profitable deviation. It is possible to come up with an infinite variety of beliefs that would rationalize the specified strategies, but equilibrium behavior will be equivalent. □

Suppose now that (NB) is not satisfied without assuming anything about the configuration of the escalatory types. This implies that the least-valuation type who can profit from assured compellence relative to appeasement will not be resolved at $\overline{m}$. The credibility problem arises if all $v_1 \geq v^{\text{aq}}$ attempt to use $\overline{m}$ because all $v_1 \in [v^{\text{aq}}, \underline{v}^{\text{a}})$ are not resolved at that allocation. These types have incentives to bluff if doing so would ensure S_2's capitulation, but because of this $\overline{m}$ can no longer compel S_2 with certainty. If these types were to mobilize in equilibrium, S_2's beliefs would have to reflect the possibility that she is facing a bluffer. This would increase the expected payoff from resistance, and some types may become willing to resist even at $\overline{m}$, which contradicts the assumption that $\overline{m}$ results in assured compellence. Hence, when $v^{\text{aq}} < \underline{v}^{\text{a}}$, the level defined by $\overline{m}$ will no longer produce assured

compellence because it is no longer credible (which is what we assumed when we derived it).

The first question is whether verbal escalation is possible. What would an equilibrium of this type look like? Since the escalation is verbal, S_1 will threaten regardless of his valuation if that would compel S_2 to capitulate for sure. Hence, all $v_1 \in [0, \overline{v}_1]$ escalate and so $G_1(v_1^*(0)) = F_1(v_1^*(0))$. Since we require that $v_2^*(0) \geq u$, using this in (CR_2') yields:

$$\frac{(1 - F_1(v_1^*(0)))c_2 - a_2}{1 - p_1(0) + p_1(0)F_1(v_1^*(0))} \geq u \tag{VB}$$

as the condition for verbal escalation. The following lemma summarizes the verbal compellence equilibrium.

LEMMA B.4. *Assume* (NB) *is not satisfied. If* (VB) *is satisfied, then S_1 escalates verbally regardless of his valuation (only $v_1 \geq v_1^*(0)$ attack if resisted), S_2's beliefs remain the same as her priors, and she capitulates regardless of her valuation.*

We now turn to militarized assured compellence when bluffing is possible. This compellent allocation, which I shall label $\widetilde{m} > 0$, will be such that even though S_2 capitulates for sure after seeing it, not all types who mobilize at that level are resolved. Since this allocation is strictly positive, (VB) cannot be satisfied. Define $\widetilde{v}^{\mathrm{aq}}$ as the type that is just indifferent between assured compellence at $\widetilde{m}$ and appeasement. We must look for an equilibrium where $v_1 \geq \widetilde{v}^{\mathrm{aq}}$ mobilize at $\widetilde{m}$ but only $v_1 \geq v_1^*(\widetilde{m}) > \widetilde{v}^{\mathrm{aq}}$ are resolved at that allocation. Given that $v_1 \in [\widetilde{v}^{\mathrm{aq}}, \overline{v}_1]$ mobilize at $\widetilde{m}$ but $v_1 < v_1^*(\widetilde{m})$ are unresolved, the probability that S_1 will capitulate when resisted is $\Pr[v_1 \leq v_1^*(\widetilde{m})]$, or $G_1(v_1^*(\widetilde{m})) = (v_1^*(\widetilde{m}) - \widetilde{v}^{\mathrm{aq}})/(\overline{v}_1 - \widetilde{v}^{\mathrm{aq}})$. Since S_2 capitulates for sure, we have $v_2^*(\widetilde{m}) = u$, or:

$$\frac{(\overline{v}_1 - v_1^*(\widetilde{m}))c_2 - (\overline{v}_1 - \widetilde{v}^{\mathrm{aq}})a_2}{(\overline{v}_1 - \widetilde{v}^{\mathrm{aq}})p_2(\widetilde{m}) + (v_1^*(\widetilde{m}) - \widetilde{v}^{\mathrm{aq}})(1 - p_2(\widetilde{m}))} = u. \tag{B.10}$$

We also know that $\widetilde{v}^{\mathrm{aq}}$'s expected payoff is zero, $\widetilde{v}^{\mathrm{aq}} - \widetilde{m} = 0$, and so we can substitute $\widetilde{v}^{\mathrm{aq}} = \widetilde{m}$ in (B.10) to obtain the cubic:

$$A\widetilde{m}^3 + (AM_1 + B)\widetilde{m}^2 + (BM_1 + C)\widetilde{m} + CM_1 + D = 0, \tag{B.11}$$

where $A = u + a_2$, $B = X + AM$, $C = Y + XM$, $D = -(M^2 - M_1M_2)(c_1 - a_1)c_2$, with $X = \overline{v}_1(c_2 - a_2) - (u + c_2)(c_1 - a_1)$, and $Y = M_1(c_1 - a_1)c_2 - M_2u\overline{v}_1$. The smallest real positive root is the solution for $\widetilde{m}$, and can then use it to obtain the value of $v_1^*(\widetilde{m})$. This leads to our next result, an equilibrium with a militarized escalation that involves bluffing.

PROPOSITION B.6. *Assume that neither* (NB) *nor* (VB) *is satisfied. The following assessment constitutes the essentially unique equilibrium.*

Strategy for S_1: (i) if $v_1 < \widetilde{v}^{\mathrm{aq}}$, appease; if $v_1 \geq \widetilde{v}^{\mathrm{aq}}$, mobilize at $\widetilde{m}$. (ii) If resisted, fight if, and only if, $m \geq \underline{m}(v_1)$.

Strategy and beliefs for S_2: (i) if $m < \widetilde{m}$, update to believe that v_1 is uniformly distributed on $[0, \widetilde{v}^{\mathrm{aq}}]$ and resist; (ii) if $m \geq \widetilde{m}$, update to believe that v_1 is uniformly distributed on $[\widetilde{v}^{\mathrm{aq}}, \overline{v}_1]$ and capitulate.

Proof All mobilizations except $\widetilde{m}$ are off the equilibrium path of play. The proposition specifies one set of beliefs for S_2 that rationalizes her behavior. There are infinite variations on these beliefs that can rationalize the strategy but they all produce the same equilibrium behavior. Verifying that S_1's strategy is optimal is straightforward. □

Observe now that it must be the case that $\widetilde{m} > \overline{m}$; that is the assured compellence level when bluffers are present must be higher than the level that can achieve compellence when mobilization is credible. To see why this should be so, suppose that it is not, so $\widetilde{m} \leq \overline{m}$. By definition, $\overline{v}_2$ capitulates after $\overline{m}$ because her payoff from capitulating equals the expected payoff from resisting, which equals her payoff from war because a credible escalation means that S_1 would fight for sure. If escalation involves bluffing, then resistance yields a payoff that is strictly better than war because S_1 capitulates with positive probability. Since S_2's payoff from war decreases in m, it follows that $\overline{v}_2$ will not capitulate for any $\widetilde{m} \leq \overline{m}$, contradicting the supposition that $\widetilde{m}$ is an assured compellence level.

This result only requires that the no-bluffing condition in (NB) and the no-verbal bluff condition in (VB) fail; it is not tied in any way to the distribution of escalatory types. Proposition B.5 handles the case where both (COMPELLENCE) and (NB) are satisfied. Since Lemma B.4 and Proposition B.6 collectively cover all cases where (NB) is not satisfied, these results exhaust all the possibilities, including the special case where (COERCION) holds.

Assured Appeasement

The analysis thus far has implicitly assumed that $\overline{v}_1$ is large enough to accommodate the highest types required by the various results. For example, Proposition B.3 assumes that $v^{\text{ca}} \leq \overline{v}_1$, which implies that there exist types whose valuations are high enough to make them willing to allocate $\overline{m}$. What would happen to our results if this were not the case?

Consider first the issue of assured compellence in any non-bluffing equilibrium (that is, all results except the one specified in Proposition B.6). In propositions B.2, B.3, and B.4, assured compellence is possible only if $v^{\text{ca}} < \overline{v}_1$. If this condition is not satisfied, then no equilibrium outcome will involve an allocation of $\overline{m}$. Rather, $\overline{v}_1$'s strategy would be determined by the region in which this valuation falls. Taking as an example Proposition B.3, if $\overline{v}_1 \in (v^{\text{cq}}, v^{\text{ca}})$, then the equilibrium strategy would be for all $v_1 < v^{\text{cq}}$ to appease (as before), and all $v_1 \in [v^{\text{cq}}, \overline{v}_1]$ to mobilize at $\widehat{m}(v_1)$. In other words, no type will mobilize at $\overline{m}$. The equilibrium outcome could still involve capitulation by S_2, but only as consequence of coercion, not assured compellence. If $\overline{v}_1 < v^{\text{cq}}$, then the equilibrium strategy would be for S_1 to appease regardless of valuation. I shall refer to this as an *assured appeasement* equilibrium.

It is not difficult to see that an analogous argument establishes the possibility for such a result in all remaining configurations. In general, the equilibrium outcomes will "shrink" as $\overline{v}_1$ decreases. Under Proposition B.1, $\overline{v}_1 \in (v^{\text{wq}}, \underline{v}^{\text{c}})$ means that only appeasement and war can happen in equilibrium, and $\overline{v}_1 < v^{\text{wq}}$ means that assured appeasement is the only possibility, which is what $\overline{v}_1 < v^{\text{aq}}$ will also lead to under Proposition B.5. Lowering the upper bound on the range of S_1 valuations depresses the possible outcomes until appeasement remains the only optimal strategy.

PROPOSITION B.7. *If $\overline{v}_1$ is smaller than the least-resolved type that would have escalated*

if it existed, the equilibrium outcome is assured appeasement: S_1 appeases regardless of valuation.

It is not hard to see this in the non-bluffing equilibria. For the equilibrium with bluffing, $\overline{v}_1 > \underline{m}(\overline{v}_1)$ is sufficient to guarantee that S_1 will appease regardless of valuation if the conditions for Proposition B.6 obtain.

Appendix C

Formalities for Chapter 4

We begin with the measures of stability, peace, and expected mobilization. Calculating the probabilities for equilibria that do not involve coercive mobilizations is straightforward, and so I will derive the expressions for the more complicated ones. Consider first the credible coercion equilibrium from Proposition B.3. Since threats are genuine in this equilibrium, we need to be able to calculate the probability that S_2 stands firm, which depends on the particular mobilization level, not simply on the event of escalation. Because only resolute types escalate, S_2 resists if $v_2 \geq v_2^*(m)$. All $v_1 \in [v^{\text{cq}}, v^{\text{ca}})$ mobilize at their type-dependent optimal coercive levels, $\widehat{m}(v_1)$, which in turn determines the cut-point for S_2. If we knew S_1's type, we would derive $v_2^*(\widehat{m}(v_1))$ and then compute $\Pr[v_2 \geq v_2^*(\widehat{m}(v_1))] = 1 - F_2(v_2^*(\widehat{m}(v_1)))$ using the assumption that v_2 is uniformly distributed by F_2 on $[t, u]$. Since we do not, we must integrate:

$$W_C\,(\text{Prop. B.3}) = \int_{v^{\text{cq}}}^{v^{\text{ca}}} \Big[1 - F_2\big(v_2^*(\widehat{m}(x))\big)\Big] f_1(x)\,\mathrm{d}x.$$

Escalation stability must take into account the fact that S_1 has mobilized, and his type is now uniformly distributed by H_1 on $[v^{\text{cq}}, \overline{v}_1]$, which implies that $h_1(x) = 1/(\overline{v}_1 - v^{\text{cq}})$ for any v_1 in this interval, and 0 elsewhere:

$$W_E\,(\text{Prop. B.3}) = \int_{v^{\text{cq}}}^{v^{\text{ca}}} \Big[1 - F_2\big(v_2^*(\widehat{m}(x))\big)\Big] h_1(x)\,\mathrm{d}x.$$

Analogous computations yield:

$$\begin{aligned} W_C\,(\text{Prop. B.4}) = {} & [F_1(\overline{\tau}) - F_1(\underline{\tau})]\big[1 - F_2\big(v_2^*(\widehat{m}(\overline{\tau}))\big)\big] \\ & + \int_{\overline{\tau}}^{v^{\text{ca}}} \Big[1 - F_2\big(v_2^*(\widehat{m}(x))\big)\Big] f_1(x)\,\mathrm{d}x \end{aligned}$$

and

$$\begin{aligned} W_C\,(\text{Prop. B.2}) = {} & [F_1(v^{\text{wc}}) - F_1(v^{\text{wq}})] \\ & + \int_{v^{\text{wc}}}^{v^{\text{ca}}} \Big[1 - F_2\big(v_2^*(\widehat{m}(x))\big)\Big] f_1(x)\,\mathrm{d}x. \end{aligned}$$

Escalation stability can be obtained from these expressions by substituting the prior with posterior beliefs.

Turning now to peaceful resolutions, the probability of appeasement is the probability that S_1's type is not among the ones that escalate. Since equilibrium escalation only involves two

possible outcomes—war or S_2's capitulation—the probability of that it ends peacefully is the complement of the probability that it ends in war. For instance, in the coercion equilibrium from Proposition B.3, S_1 only escalates if $v_1 \geq v^{\text{cq}}$, so:

$$P_C(\text{Prop. B.3}) = F_1(v^{\text{cq}}) \quad \text{and} \quad P_E(\cdot) = 1 - W_E(\cdot).$$

Finally, consider the calculation of the expected mobilization level. Suppose that the continuation game's equilibrium is from Proposition B.2. We now have to account for the three ranges of escalating types. Letting $h_1(x)$ denote the posterior probability density function and $H_1(x)$ the corresponding cumulative distribution function, we obtain:

$$M_E = \int_{v^{\text{wq}}}^{v^{\text{wc}}} m^*(x)h_1(x)\,\mathrm{d}x + \int_{v^{\text{wc}}}^{v^{\text{ca}}} \widehat{m}(x)h_1(x)\,\mathrm{d}x + (1 - H_1(v^{\text{ca}}))\overline{m}.$$

The expected mobilization levels for the other cases are computed analogously.

Turning now to S_1's expected equilibrium payoff, note that in the MTM we have to account for the equilibrium in the continuation game. The payoff from assured compellence is straightforward, so I will specify the other three general equilibrium types. Recall that $\mathcal{W}_1(v_1)$ is type v_1's optimal war payoff, $\mathcal{C}_1(v_1)$ is the optimal coercion payoff, and $\mathcal{A}(v_1)$ is the assured compellence payoff. We now have:

$$\begin{aligned}
\mathcal{U}_1(\text{Prop. B.1}) &= \int_{v^{\text{wq}}}^{v^{\text{wa}}} \mathcal{W}_1(x)\,f_1(x)\,\mathrm{d}x + \int_{v^{\text{wa}}}^{\overline{v}_1} \mathcal{A}_1(x)\,f_1(x)\,\mathrm{d}x \\
\mathcal{U}_1(\text{Prop. B.2 }) &= \int_{v^{\text{wq}}}^{v^{\text{wc}}} \mathcal{W}_1(x)\,f_1(x)\,\mathrm{d}x + \int_{v^{\text{wc}}}^{v^{\text{ca}}} \mathcal{C}_1(x)\,f_1(x)\,\mathrm{d}x \\
&\qquad + \int_{v^{\text{ca}}}^{\overline{v}_1} \mathcal{A}_1(x)\,f_1(x)\,\mathrm{d}x \\
\mathcal{U}_1(\text{Prop. B.4}) &= \int_{\underline{\tau}}^{\overline{\tau}} \Big[(1 - G_2(v_2^*(\widehat{m}(\overline{\tau}))))(p(\widehat{m}(\overline{\tau}))x - c_1 - \widehat{m}(\overline{\tau})) \\
&\qquad\qquad + G_2(v_2^*(\widehat{m}(\overline{\tau})))(x - \widehat{m}(\overline{\tau}))\Big] f_1(x)\,\mathrm{d}x \\
&\qquad + \int_{\overline{\tau}}^{v^{\text{ca}}} \mathcal{C}_1(x)\,f_1(x)\,\mathrm{d}x + \int_{v^{\text{ca}}}^{\overline{v}_1} \mathcal{A}_1(x)\,f_1(x)\,\mathrm{d}x.
\end{aligned}$$

Finally, consider S_1's expected equilibrium payoff in models with non-military threats. In the bluffing equilibrium of the basic escalation game, the probability that S_2 capitulates is $F_2(\hat{v}_2)$ as given in Proposition A.3. Since S_1 escalates if $v_1 \geq \hat{v}_1$ but attacks if, and only if, $v_1 \geq v_1^*$, the *ex ante* expected payoff is:

$$\begin{aligned}
\mathcal{U}_1(\text{BE}) = F_2(\hat{v}_2) &\int_{\hat{v}_1}^{\overline{v}_1} x f_1(x)\,\mathrm{d}x \\
&+ (1 - F_2(\hat{v}_2))\left[\int_{v_1^*}^{\overline{v}_1} (px - c_1)\,f_1(x)\,\mathrm{d}x - a_1 \int_{\hat{v}_1}^{v_1^*} f_1(x)\,\mathrm{d}x\right].
\end{aligned}$$

In the genuine-threat equilibrium of Proposition A.1, the expected payoff is even simpler:

all escalating types get either their valuation with probability $F_2(v_2^*)$ or war with complementary probability:

$$\int_{\hat{v}_1}^{\overline{v}_1} \left[F_2(v_2^*)x + (1 - F_2(v_2^*))(px - c_1)\right] f_1(x)\,\mathrm{d}x.$$

In the intuitive equilibrium of the sinking costs model, only genuine challengers are willing to pay m^* from Corollary A.1, and this signal is the same for everyone. The expected payoff is:

$$\mathcal{U}_1(\mathrm{SC}) = \int_{v_1^*}^{\overline{v}_1} \left[F_2(v_2^*)x + (1 - F_2(v_2^*))(px - c_1) - m^*\right] f_1(x)\,\mathrm{d}x,$$

where $m^* = F_2(v_2^*)v_1^* + (1 - F_2(v_2^*))(-a_1)$.

In the tying-hands game, no type that escalates in equilibrium has to pay the audience costs. Taking $v_1^*(m^*)$ from (2.5) with m^* defined in (A.5), the expected payoff in both cases is:

$$\mathcal{U}_1(\mathrm{TH}) = \int_{v_1^*(m^*)}^{\overline{v}_1} \left[F_2(v_2^*)x + (1 - F_2(v_2^*))(px - c_1)\right] f_1(x)\,\mathrm{d}x.$$

The expected payoff for the risk-generation model is equivalent to the one from the tying-hands model except the lowest type to escalate is $\hat{v}_1$ from (A.3). Although it is quite possible that this type may be a bluffer, this is irrelevant for the expected payoff: burning the bridge effectively commits him to fighting if S_2 resists.

Appendix D

Formalities for Chapter 5

Recall that S_2 initiates a challenge if $v_2 \geq t$, so upon observing such a challenge S_1 will infer that his opponent's valuation is distributed uniformly on the interval $[t, \overline{v}_2]$. The analysis essentially boils down to pinning down t such that this type is indifferent between the status quo (payoff of zero) and initiating a crisis in which behavior then follows according to the MTM equilibrium prescription. In the verbal escalation equilibrium from Lemma B.4, S_1 is expected to refuse appeasement regardless of valuation. This means that t's expected payoff from initiating a crisis will be $-a_2 < 0$, so she will never do so. On the other hand, in the assured appeasement equilibrium from Proposition B.7, S_1 is expected to concede with certainty, in which case her payoff would be $t \geq 0$, and she will initiate. In equilibria that involve mobilization for certain war, this type fill find herself at war if S_1 prepares for war but will capitulate if he mobilizes for coercion:

$$\begin{aligned}\mathcal{U}_2(\text{Prop. B.1}; t) &= F_1(v^{\text{wq}})t + \int_{v^{\text{wq}}}^{v^{\text{wa}}} [(1 - p(m^*(x)))t - c_2] f_1(x)\,\mathrm{d}x \\ &\quad + (1 - F_1(v^{\text{wa}}))(-a_2) \\ \mathcal{U}_2(\text{Prop. B.2}; t) &= F_1(v^{\text{wq}})t + \int_{v^{\text{wq}}}^{v^{\text{wc}}} [(1 - p(m^*(x)))t - c_2] f_1(x)\,\mathrm{d}x \\ &\quad + (1 - F_1(v^{\text{wc}}))(-a_2).\end{aligned}$$

In crises that involve no mobilizations preparatory for war, t faces only two possible outcomes: appeasement by S_1 or her own capitulation (because she is the lowest-valuation type of S_2 to participate and these never fight under the circumstances):

$$\begin{aligned}\mathcal{U}_2(\text{Prop. B.3}; t) &= F_1(v^{\text{cq}})t + (1 - F_1(v^{\text{cq}}))(-a_2) \\ \mathcal{U}_2(\text{Prop. B.4}; t) &= F_1(\underline{\tau})t + (1 - F_1(\underline{\tau}))(-a_2) \\ \mathcal{U}_2(\text{Prop. B.5}; t) &= F_1(v^{\text{aq}})t + (1 - F_1(v^{\text{aq}}))(-a_2) \\ \mathcal{U}_2(\text{Prop. B.6}; t) &= F_1(\widetilde{v}^{\text{aq}})t + (1 - F_1(\widetilde{v}^{\text{aq}}))(-a_2).\end{aligned}$$

Setting these expression equal to zero (the expected payoff from not initiating the crisis) produces equations that implicitly define the type, t^*, that would be indifferent between staying with the status quo and starting a crisis that would result in the particular continuation. But there can be at most one t^* for whom the expected continuation payoff is zero. Loosely speaking, this is so because the expected payoff of the lowest-valuation type to initiate a crisis must be strictly increasing in that type provided S_1 does not appease for sure. Uniqueness

of t^* also implies that the deterrence model has a unique equilibrium in which conditional on a challenge, S_1 updates to believe that v_2 is distributed uniformly on $[t^*, \overline{v}_2]$:

PROPOSITION D.1. *The militarized deterrence model has a unique equilibrium in which S_2 initiates if, and only if, $v_2 \geq t^*$. The game continues according to the equilibrium in MTM induced by beliefs $[t^*, \overline{v}_2]$, which S_1 obtains by Bayes' Rule upon observing a challenge.*

Given any continuation equilibrium, EQ, situational stability is:

$$W_S = (1 - F_2(t^*(\text{EQ})))W_C(\text{EQ}).$$

That is, it is the probability that S_2 issues a challenge, $\Pr[v_2 \geq t^*(\text{EQ})]$, multiplied by the probability that the crisis resulting in the equilibrium EQ of the MTM escalates to war, $W_C(\text{EQ})$ with $F_2(\cdot)$ truncated to the interval $[t^*(\text{EQ}), \overline{v}_2]$.

There are three possible ways in which war can be avoided. The probability that S_2 does not even challenge is the status quo is the probability that S_2's valuation is too low to initiate a crisis:

$$P_S = F_2(t^*(\text{EQ})).$$

The probability that the status quo is peacefully revised is the probability that S_2 initiates a crisis but S_1 appeases her: $(1 - P_S)P_C(\text{EQ})$. Finally, the probability that the status quo is protected by the forceful act of S_1 is the probability that S_2 initiates a crisis but is then forced to back down when S_1 mobilizes: $(1 - P_S)(1 - P_C(\text{EQ}))P_E(\text{EQ})$.

Bibliography

Achen, Christopher H., and Snidal, Duncan. 1989. Rational Deterrence Theory and Comparative Case Studies. *World Politics*, **41**(2), 143–169.

Acheson, Dean. 1987. *Present at the Creation: My Years at the State Department*. New York: W. W. Norton & Company.

Alt, James, Calvert, Randall, and Humes, Brian. 1988. Reputation and Hegemonic Stability: A Game-Theoretic Analysis. *American Political Science Review*, **82**(2), 445–466.

Altfeld, Michael F. 1983. Arms Races? — And Escalation? A Comment on Wallace. *International Studies Quarterly*, **27**(2), 225–231.

Appleman, Roy E. 1961. *South to the Naktong, North to the Yalu*. Washington, DC: U.S. Government Printing Office.

Arrian. 1976. *History of Alexander and Indica*. Cambridge: Harvard University Press. Two volumes.

Arrow, Kenneth J. 1970. *Social Choice and Individual Values*. 2nd edn. New Haven: Yale University Press.

Austen-Smith, David, and Banks, Jeffrey S. 2000. Cheap Talk and Burned Money. *Journal of Economic Theory*, **91**(1), 1–16.

Bajanov, Evgueni. 1995. Assessing the Politics of the Korean War, 1949-51. *Cold War International History Project Bulletin*, 54–91.

Ballard, John R. 1998. *Upholding Democracy: The United States Military Campaign in Haiti, 1994-1997*. Westport: Praeger.

Banks, Jeffrey S. 1990. Equilibrium Behavior in Crisis Bargaining Games. *American Journal of Political Science*, **34**(3), 599–614.

Bates, Darell. 1984. *The Fashoda Incident of 1898: Encounter on the Nile*. Oxford: Oxford University Press.

Beeching, Jack. 1975. *The Chinese Opium Wars*. New York: Harcourt Brace Jovanovich.

Beevor, Antony. 1998. *Stalingrad, The Fateful Siege: 1942–1943*. New York: Penguin.

Bennett, D. Scott, and Stam, Allan C. 1996. The Duration of Interstate Wars, 1816-1985. *American Political Science Review*, **90**(2), 239–57.

Bernheim, B. Douglas. 1984. Rationalizable strategic behavior. *Econometrica*, **52**(4), 10007–28.

Betts, Richard K. 1995. *Military Readiness: Concepts, Choices, and Consequences*. Washington, D.C.: The Brookings Institution.

Binmore, Ken. 2007. *Does Game Theory Work? The Bargaining Challenge*. Cambridge: The MIT Press.

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

Blechman, Barry M. C., and Kaplan, Stephen S. 1978. *Force Without War: U.S. Armed Forces as a Political Instrument*. Washington, DC: Brookings Institute.

Brandenburger, Adam. 1992. Knowledge and Equilibrium in Games. *The Journal of Economic Perspectives*, **6**(4), 83–101.

Brecher, Michael. 1993. *Crises in World Politics: Theory and Reality*. Oxford: Pergamon Press.

Brewer, John. 1990. *The Sinews of Power: War, Money, and the English State, 1688–1783*. Cambridge: Harvard University Press.

Brito, Dagobert L., and Intriligator, Michael D. 1985. Conflict, War, and Redistribution. *American Political Science Review*, **79**(4), 943–57.

Brodie, Bernard. 1959. *Strategy in the Missile Age*. Princeton: Princeton University Press.

Brune, Lester H. 1993. *America and the Iraqi Crisis, 1990–1992: Origins and Aftermath*. Claremont: Regina Books.

Brune, Lester H. (ed). 1996. *The Korean War: Handbook of the Literature and Research*. Westport: Greenwood Press.

Bucholz, Arden. 2001. *Moltke and the German Wars, 1864-1871*. Houndmills: Palgrave.

Bueno de Mesquita, Bruce. 1981. Risk, Power Distributions, and the Likelihood of War. *International Studies Quarterly*, **25**(4), 541–568.

Bueno de Mesquita, Bruce. 1990. Pride of Place: The Origins of German Hegemony. *World Politics*, **43**(1), 28–52.

Bueno de Mesquita, Bruce, and Lalman, David. 1986. Reason and War. *American Political Science Review*, **80**(4), 1113–1129.

Bueno de Mesquita, Bruce, and Lalman, David. 1992. *War and Reason*. New Haven: Yale University Press.

Bueno de Mesquita, Bruce, and Siverson, Randolph M. 1995. War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability. *American Political Science Review*, **89**, 841–855.

Bueno de Mesquita, Bruce, Morrow, James D., Siverson, Randolph M., and Smith, Alastair. 1999. An Institutional Explanation of the Democratic Peace. *American Political Science Review*, **89**(4), 841–855.

Bueno de Mesquita, Bruce, Morrow, James, Siverson, Randolph, and Smith, Alastair. 2003. *The Logic of Political Survival*. Cambridge: The M.I.T. Press.

Butow, R.J.C. 1961. *Tojo and the Coming of the War*. Stanford: Stanford University Press.

Calvin, James Barnard. 1984. *The China-India Border War (1962)*. Tech. rept. Marine Corps Command and Staff College.

Carr, Edward Hallett. 1939. *The Twenty Years Crisis, 1919-1939*. London: Macmillan.

Chen, Jian. 1992. China's Changing Aims during the Korean War, 1950–1951. *The Journal of American-East Asian Relations*, **1**(Spring), 8–41.

Chen, Jian. 1994. *China's Road to the Korean War: The Making of the Sino-American Confrontation*. New York: Columbia University Press.

Chen, Jian. 1996. Chinese Policy and the Korean War. *In:* Brune (1996).

Chen, Xiaolu. 1989. China's Policy toward the United States, 1949-1955. *In:* Harding and Yuan (1989).

Chiozza, Giacomo, and Goemans, Hein. 2003. Peace Through Insecurity: Tenure and International Conflict. *Journal of Conflict Resolution*, **47**(4), 443–467.

Chiozza, Giacomo, and Goemans, Hein. 2004a. Avoiding Diversionary Targets. *Journal of Peace Research*, **41**(4), 423–443.

Chiozza, Giacomo, and Goemans, Hein. 2004b. International Conflict and the Tenure of Leaders: Is War Still *Ex Post* Inefficient? *American Journal of Political Science*, **48**(3), 604–619.

Cho, In-Koo, and Kreps, David M. 1987. Signaling Games and Stable Equilibria. *Quarterly Journal of Economics*, **102**(May), 179–221.

Christensen, Thomas J. 1992. Threats, Assurances, and the Last Chance for Peace: The Lessons from Mao's Korean War Telegrams. *International Security*, **17**(1), 122–154.

Cimbala, Stephen J. 1994. *Military Persuasion: Deterrence and Provocation in Crisis and War*. University Park: The Pennsylvania State University Press.

Clarke, Kevin A., and Primo, David M. 2007. Modernizing Political Science: A Model-Based Approach. *Perspectives on Politics*, **5**(4), 741–753.

Claude, Inis. 1962. *Power and International Relations*. New York: Random House.

Costello, John. 1981. *The Pacific War*. New York: Perennial.

Crawford, Vincent P., and Sobel, Joel. 1982. Strategic Information Transmission. *Econometrica*, **50**(6), 1431–1451.

Danilovic, Vesna. 2001. Conceptual and Selection Bias Issues in Deterrence. *Journal of Conflict Resolution*, **45**(1), 97–125.

Dorussen, Han, and Mo, Jongryn. 2001. Ending Economic Sanctions: Audience Costs and Rent-Seeking as Commitment Strategies. *Journal of Conflict Resolution*, **45**(4), 395–426.

Downs, George W., and Rocke, David M. 1994. Conflict, Agency, and Gambling for Resurrection: The Principal-Agent Problem Goes to War. *American Journal of Political Science*, **38**(2), 362–80.

Drezner, Daniel W. 2003. The Hidden Hand of Economic Coercion. *International Organization*, **57**(Summer), 643–659.

Dupuy, R. Ernest, and Dupuy, Trevor N. 1985. *The Encyclopedia of Military History from 3500 B.C. to the Present*. 2nd rev edn. New York: Harper & Row Publishers.

Dupuy, Trevor N. 1978. *Elusive Victory: The Arab-Israeli Wars, 1947-1974*. New York: Harper & Row.

Ellsberg, Daniel. 1975. The Theory and Practice of Blackmail. In: Young, Oran R. (ed), *Bargaining: Formal Theories of Negotiation*. Urbana: University of Illinois Press. Original Santa Monica: Rand Coroporation P-3883, 1959.

Erickson, John. 1999. *The Road to Stalingrad: Stalin's War with Germany*. New Haven: Yale University Press.

Eyerman, Joe, and Hart, Jr., Robert A. 1996. An Empirical Test of the Audience Cost Proposition: Democracy Speaks Louder than Words. *Journal of Conflict Resolution*, **40**(4), 597–616.

Farkas, Andrew. 1998. *State Learning and International Change*. Ann Arbor: The University of Michigan Press.

Farrar, Peter N. 1983. Britain's Proposal for a Buffer Zone South of the Yalu in November 1950: Was It a Neglected Opportunity to End the Fighting in Korea? *Journal of Contemporary History*, **18**(2), 327–351.

Farrell, Joseph, and Gibbons, Robert. 1989. Cheap Talk Can Matter in Bargaining. *Journal of Economic Theory*, **48**(1), 221–237.

Fearon, James D. 1992. *Threats to Use Force: Costly Signals and Bargaining in International Crises.* Ph.D. thesis, University of California, Berkeley.

Fearon, James D. 1994a. Domestic Political Audiences and the Escalation of International Disputes. *American Political Science Review*, **88**(Sept.), 577–92.

Fearon, James D. 1994b. Signaling versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model. *Journal of Conflict Resolution*, **38**(2), 236–69.

Fearon, James D. 1995. Rationalist Explanations for War. *International Organization*, **49**(3), 379–414.

Fearon, James D. 1997. Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs. *Journal of Conflict Resolution*, **41**(1), 68–90.

Fearon, James D. 2002. Selection Effects and Deterrence. *International Interactions*, **28**(1), 5–29.

Ferris, Wayne H. 1973. *The Power Capabilities of Nation-States: International Conflict and War*. Lexington: Lexington Books.

Foot, Rosemary. 1991. Making Known the Unknown War: Policy Analysis of the Korean Conflict in the Last Decade. *Diplomatic History*, **15**(2), 411–432.

Frankel, Max. 2005. *High Noon in the Cold War: Kennedy, Khrushchev, and the Cuban Missile Crisis*. Presidio Press.

Freedman, Lawrence (ed). 1998. *Strategic Coercion: Concepts and Cases*. Oxford: Oxford University Press.

Freedman, Lawrence. 2003. *Evolution of Nuclear Strategy*. 3rd edn. New York: Palgrave Macmillan.

Freedman, Lawrence, and Karsh, Efraim. 1993. *The Gulf Conflict, 1990–1991: Diplomacy and War in the New World Order*. Princeton: Princeton University Press.

Frey, Bruno S., and Kucher, Marcel. 2000. History as Reflected in Capital Markets: The Case of World War II. *Journal of Economic History*, **60**(2), 468–496.

Fudenberg, Drew, and Levine, David K. 1992. Maintaining Reputation when Strategies are Imperfectly Observed. *The Review of Economics and Statistics*, **59**(3), 561–579.

Fudenberg, Drew, and Tirole, Jean. 1991a. *Game Theory*. Cambridge: The M.I.T. Press.

Fudenberg, Drew, and Tirole, Jean. 1991b. Perfect Bayesian and Sequential Equilibrium. *Journal of Economic Theory*, **53**(2), 236–260.

Gaddis, John Lewis. 1989. The American "Wedge" Strategy, 1949-1955. *In:* Harding and Yuan (1989).

Gaddis, John Lewis. 1998. *We Now Know: Rethinking Cold War History*. Oxford: Oxford University Press.

Garfinkel, Michelle, and Skaperdas, Stergios. 2000. Conflict Without Misperceptions or Incomplete Information: How Future Matters. *Journal of Conflict Resolution*, **44**(6), 793–807.

Garnham, David. 1976. Dyadic International War 1816-1965: The Role of Power Parity and Geographical Proximity. *The Western Political Quarterly*, **29**(2), 231–242.

Gaubatz, Kurt Taylor. 1996. Democratic States and Commitment in International Relations. *International Organization*, **50**(1), 109–139.

Geller, Daniel S. 1993. Power Differentials and War in Rival Dyads. *International Studies Quarterly*, **37**(2), 173–193.

Gelpi, Christopher, and Griesdorf, Michael. 2001. Winners or Losers: Democracies in International Crises. *American Political Science Review*, **95**(3), 633–648.

George, Alexander, and Smoke, Richard. 1974. *Deterrence in American Foreign Policy: Theory and Practice*. New York: Columbia University Press.

George, Alexander L. 1994. Coercive Diplomacy: Definition and Characteristics. *In:* George and Simons (1994).

George, Alexander L., and Simons, William E. (eds). 1994. *The Limits of Coercive Diplomacy*. 2nd edn. Boulder: Westview Press.

Gies, Joseph, and Gies, Frances. 1969. *Life in a Medieval City*. New York: Harper Perennial.

Gilpin, Robert. 1981. *War and Change in World Politics*. Princeton: Princeton University Press.

Glaser, Charles L. 2004. When Are Arms Races Dangerous? Rational versus Suboptimal Arming. *International Security*, **28**(4), 44–84.

Gochman, Charles S. 1990. Capability-Driven Disputes. In: Gochman, Charles S., and Sabrosky, Alan Ned (eds), *Prisoners of War? Nation-States in the Modern Era*. Lexington: Lexington Books.

Goemans, Hein E. 2000. Fighting for Survival: The Fate of Leaders and the Duration of War. *The Journal of Conflict Resolution*, **44**(5), 555–79.

Goldhamer, Herbert. 1979. *Reality and Belief in Military Affairs*. Santa Monica: Rand. Rand Report R-2448-NA, February 1979.

Goldsworthy, Adrian. 2002. *The Punic Wars*. London: Cassell.

Goncharov, Sergei N., Lewis, John W., and Xue, Litai. 1993. *Uncertain Partners: Stalin, Mao, and the Korean War*. Stanford: Stanford University Press.

Gowa, Joanne. 2001. *Who Yields? Democracies in International Crises*. Presented at the International Relations Speaker Series, Yale University, April 25, 2001.

Graff, David A. 2002. *Medieval Chinese Warfare, 300-900*. London: Routledge.

Green, Peter. 1996. *The Greco-Persian Wars*. Berkeley: University of California Press. Revised ed.

Gribkov, Anatoli I., and Smith, William Y. 1993. *Operation Anadyr: U.S. and Soviet Generals Recount the Cuban Missile Crisis*. Chicago: Edition Q.

Guisinger, Alexandra, and Smith, Alastair. 2002. Honest Threats: The Interaction of Reputation and Political Institutions in International Crises. *Journal of Conflict Resolution*, **46**(2), 175–200.

Gurtov, Melvin, and Hwang, Byong-Moo. 1980. *China Under Threat: The Politics of Strategy and Diplomacy*. Baltimore: The Johns Hopkins University Press.

Hall, Richard C. 2000. *The Balkan Wars, 1912–1913: Prelude to the First World War*. London: Routledge.

Hao, Yufan, and Zhai, Zhihai. 1990. China's Decision to Enter the Korean War: History Revisited. *The China Quarterly*, Mar., 94–115.

Harcave, Sidney (ed). 1990. *The Memoirs of Count Witte*. New York: M. E. Sharpe, Inc.

Harding, Harry, and Yuan, Ming (eds). 1989. *Sino-American Relations, 1945-1955: A Joint Reassessment of a Critical Decade*. Wilmington: Scholarly Resources.

Harsanyi, John C. 1968. Games with Incomplete Information Played by 'Bayesian' Players. *Management Science*, **14**(Mar.), 159–82, 320–34, 486–502.

Harvey, Frank P. 1995. Rational Deterrence Theory Revisited: A Progress Report. *Canadian Journal of Political Science*, **28**(3), 403–436.

Harvey, Frank P. 1998. Rigor Mortis or Rigor, More Tests: Necessity, Sufficiency, and Deterrence Logic. *International Studies Quarterly*, **42**(4), 675–707.

Heller, Francis H. (ed). 1997. *The Korean War: A 25-Year Perspective*. Lawrence: Regents Press of Kansas.

Herek, Gregory M., Janis, Irving, and Huth, Paul. 1987. Decision-Making During International Crises: Is Quality of Process Related to Outcome? *Journal of Conflict Resolution*, **31**(2), 203–226.

Hermann, Charles F. (ed). 1972. *International Crises: Insights from Behavioral Research*. New York: The Free Press.

Hershberg, James G. 2004. Russian Documents on the Korean War, 1950–53. *Cold War International History Project Bulletin*, 369–398.
Hibbert, Christopher. 1974. *The House of Medici: Its Rise and Fall*. New York: Harper Perennial.
Hickey, Donald R. 1989. *The War of 1812: A Forgotten Conflict*. Urbana: University of Illinois Press.
Hinton, Harold C. 1966. *Communist China in World Politics*. London: Macmillan.
Hirshleifer, Jack. 1988. The Analytics of Continuing Conflict. *Synthese*, **76**(2), 201–233.
Hirshleifer, Jack. 1991. The Technology of Conflict as an Economic Activity. *The American Economic Review*, **81**(2), 130–4.
Hirshleifer, Jack. 2000. The Macrotechnology of Conflict. *Journal of Conflict Resolution*, **44**(6), 773–792.
Hoffmann, Steven A. 1990. *India and the China Crisis*. Berkeley: University of California Press.
Homer. 1990. *The Iliad*. New York: Penguin. Tr. Robert Fagles.
Hopf, Ted. 1991. Polarity, The Offense-Defense Balance, and War. *American Political Science Review*, **85**(2), 475–93.
Hopf, Ted. 1994. *Peripheral Visions: Deterrence Theory and American Foreign Policy in the Third World, 1965-1990*. Ann Arbor: University of Michigan Press.
Hopple, Gerald W. 1984. Intelligence and Warning: Implications and Lessons of the Falkland Islands War. *World Politics*, **36**(3), 339–361.
Houweling, Henk W., and Siccama, Jan G. 1988. Power Transitions as a Cause of War. *The Journal of Conflict Resolution*, **32**, 87–102.
Howard, Michael. 1981. *The Franco-Prussian War: The German Invasion of France, 1870-1871*. London: Routledge.
Huth, Paul, and Russett, Bruce. 1984. What Makes Deterrence Work? Cases from 1900-1980. *World Politics*, **36**, 496–526.
Huth, Paul, and Russett, Bruce. 1988. Deterrence Failure and Crisis Escalation. *International Studies Quarterly*, **32**, 29–45.
Huth, Paul, and Russett, Bruce. 1990. Testing Deterrence Theory: Rigour Makes a Difference. *World Politics*, **42**, 466–501.
Huth, Paul, Gelpi, Christopher, and Bennett, D. Scott. 1993. The Escalation of Great Power Militarized Disputes: Testing Rational Deterrence Theory and Structural Realism. *American Political Science Review*, **87**(3), 609–623.
Huth, Paul K. 1988. *Extended Deterrence and the Prevention of War*. New Haven: Yale University Press.
Huth, Paul K., and Russett, Bruce. 1993. General Deterrence Between Enduring Rivals: Testing Three Competing Models. *American Political Science Review*, **87**(1), 61–73.
Ike, Nobutaka. 1967. *Japan's Decision for War: Records of the 1941 Policy Conferences*. Stanford: Stanford University Press.
Iklé, Fred Charles. 1964. *How Nations Negotiate*. New York: Harper & Row, Publishers.
India Ministry of External Affairs. 1963. *Notes, Memoranda, and Letters Exchanged and Agreements Signed Between the Governments of India and China: White Paper*. New Delhi: Ministry of External Affairs.
Jakobson, Max. 1961. *The Diplomacy of the Winter War: An Account of the Russo-Finnish War, 1939–1940*. Cambridge: Harvard University Press.
Jervis, Robert. 1970. *The Logic of Images in International Relations*. New York: Columbia University Press.
Jervis, Robert. 1979. Deterrence Theory Revisited. *World Politics*, **31**(2), 289–324.
Jervis, Robert, Lebow, Richard Ned, and Stein, Janice Gross. 1985. *Psychology and Deterrence*. Baltimore: The Johns Hopkins University Press.
Kagan, Donald. 1990. *The Archidamian War*. Ithaca: Cornell University Press.
Kagan, Donald. 2003. *The Peloponnesian War*. New York: Penguin Books.
Kahn, Herman. 1960. *On Thermonuclear War*. Princeton: Princeton University Press.
Kahn, Herman. 1965. *On Escalation: Metaphors and Scenarios*. New York: Praeger.
Kam, Ephraim. 2004. *Surprise Attack: The Victim's Perspective*. Cambridge: Harvard University Press.
Karsten, Peter, Howell, Peter D., and Allen, Artis Frances. 1984. *Military Threats: A Systematic Historical Analysis of the Determinants of Success*. Westport: Greenwood Press.

Kaufman, Burton I. 1997. *The Korean War: Challenges in Crisis, Credibility, and Command*. 2nd edn. New York: McGraw-Hill.

Kennedy, Paul M.1987. *The Rise and Fall of the Great Powers*. Cambridge: Cambridge University Press.

Kilgour, D. Marc, and Zagare, Frank C. 1991. Credibility, Uncertainty, and Deterrence. *American Journal of Political Science*, **35**(2), 305–334.

Kim, Woosang. 1989. Power, Alliance, and Major Wars, 1816-1975. *Journal of Conflict Resolution*, **33**(2), 255–273.

Kim, Woosang, and Morrow, James. 1992. When Do Power Shifts Lead to War? *American Journal of Political Science*, **36**(Nov.), 896–922.

Kissinger, Henry A. 1979. *The White House Years*. Boston: Little Brown.

Kreps, David, and Wilson, Robert. 1982. Reputation and Imperfect Information. *Journal of Economic Theory*, **27**(Aug.), 253–79.

Kreps, David, Milgrom, Paul R., Roberts, D. John, and Wilson, Robert. 1982. Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. *Journal of Economic Theory*, **27**(2), 245–252.

Kydd, Andrew H. 2000. Arms Races and Arms Control: Modeling the Hawk Perspective. *American Journal of Political Science*, **44**(2), 222–238.

Kydd, Andrew H. 2005. *Trust and Mistrust in International Relations*. Princeton: Princeton University Press.

Lai, Brian. 2004. The Effects of Different Types of Military Mobilization on the Outcome of International Crises. *Journal of Conflict Resolution*, **48**(2), 211–229.

Lake, David. 1992. Powerful Pacifists: Democratic States and War. *American Political Science Review*, **86**(1), 24–38.

Lalman, David. 1988. Conflict Resolution and Peace. *American Journal of Political Science*, **32**(3), 590–615.

Lamb, Alastair. 1964. *The China-India Border*. Oxford: Oxford University Press.

Langer, William L. 1935. *The Diplomacy of Imperialism, 1890-1902*. Cambridge: Harvard University Press. Two volumes.

Langlois, Catherine C., and Langlois, Jean-Pierre P. 2005. Fully Informed and On the Road to Ruin: The Perfect Failure of Asymmetric Deterrence. *International Studies Quarterly*, **49**(3), 503–527.

Lauren, Paul Gordon. 1994. Coercive Diplomacy and Ultimata: Theory and Practice in History. *In:* George and Simons (1994).

Lebow, Richard Ned. 1981. *Between Peace and War: The Nature of International Crisis*. Baltimore: The Johns Hopkins University Press.

Lebow, Richard Ned. 1987. Deterrence Failure Revisited. *International Security*, **12**(1), 197–213.

Lebow, Richard Ned, and Stein, Janice Gross. 1987. Beyond Deterrence. *Journal of Social Issues*, **43**, 5–71.

Lebow, Richard Ned, and Stein, Janice Gross. 1989. Rational Deterrence Theory: I Think, Therefore I Deter. *World Politics*, **41**, 208–224.

Lebow, Richard Ned, and Stein, Janice Gross. 1990. Deterrence: The Elusive Dependent Variable. *World Politics*, **42**, 336–369.

Leeds, Brett Ashley. 1999. Domestic Political Institutions, Credible Commitments, and International Cooperation. *American Journal of Political Science*, **43**(4), 979–1002.

Lemke, Douglas, and Werner, Suzanne. 1996. Power Parity, Commitment to Change, and War. *International Studies Quarterly*, **40**(2), 235–260.

Leng, Russell J. 1993. *Interstate Crisis Behavior, 1816-1980: Realism versus Reciprocity*. Cambridge: Cambridge University Press.

Leventoğlu, Bahar, and Slantchev, Branislav L. 2007. The Armed Peace: A Punctuated-Equilibrium Theory of War. *American Journal of Political Science*, **51**(4), 755–771.

Levy, Jack S. 1987. Declining Power and the Preventive Motivation for War. *World Politics*, **40**(1), 82–107.

Levy, Jack S. 1988. When Do Deterrent Threats Work. *British Journal of Political Science*, **18**, 485–512.

Levy, Jack S. 1989a. The Causes of War: A Review of Theories and Evidence. In: Jervis, Robert, Stern, Paul C., and Tilly, Charles (eds), *Behavior, Society, and Nuclear War*, vol. 1. New York: Oxford University Press.

Levy, Jack S. 1989b. The Diversionary Theory of War: A Critique. Pages 259–288 of: Midlarsky, Manus I. (ed), *Handbook of War Studies*. Ann Arbor: University of Michigan Press.

Levy, Jack S. 1994. Learning and Foreign Policy: Sweeping a Conceptual Minefield. *International Organization*, **48**(2), 279–312.

Li, Xiaobing, Millett, Allan R., and Yu, Bin (eds). 2001. *Mao's Generals Remember Korea*. Lawrence: University Press of Kansas.

Liddell Hart, B. H. 1991. *Strategy*. New York: Meridian.

Lipson, Charles. 2003. *Reliable Partners: How Democracies Have Made a Separate Peace*. Princeton: Princeton University Press.

Lohmann, Susanne. 2003. Why Do Institutions Matter? An Audience-Cost Theory of Institutional Commitment. *Governance*, **16**(1), 95–110.

Lynn, John A. 1999. *The Wars of Louis XIV, 1667–1714*. London: Longman.

Mahon, Lord. 2005. *The Life of Belisarius*. Yardley, PA: Westholme Publishing.

Mailath, George J., and Samuelson, Larry. 1998. *Your Reputation Is Who You're Not, Not Who You'd Like To Be*. Manuscript, Department of Economics, University of Pennsylvania.

Mailath, George J., and Samuelson, Larry. 2006. *Repeated Games and Reputations: Long-Run Relationships*. Oxford: Oxford University Press.

Manchester, William. 1978. *American Caesar: Douglas MacArthur, 1880–1964*. Boston: Little, Brown, and Company.

Mansfield, Edward, Milner, Helen, and Rosendorff, B. Peter. 2002. Why Democracies Cooperate More: Electoral Control and International Trade Agreements. *International Organization*, **56**(3), 477–513.

Mansourov, Alexandre Y. 1995. Stalin, Mao, Kim, and China's Decision to Enter the Korean War, September 16-October 15, 1950: New Evidence from the Russian Archives. *Cold War International History Project Bulletin*, 94–119.

Maoz, Zeev. 1983. Resolve, Capabilities, and the Outcome of Interstate Disputes, 1816-1976. *Journal of Conflict Resolution*, **27**(2), 195=229.

Maoz, Zeev. 1990. *Paradoxes of War: On the Art of National Self-Entrapment*. Boston: Unwin Hyman.

Maoz, Zeev, and Russett, Bruce. 1993. Normative and Structural Causes of Democratic Peace. *American Political Science Review*, **87**(Sept.), 624–638.

Mark, Eduard. 1997. The War Scare of 1946 and Its Consequences. *Diplomatic History*, **21**(3), 383–415.

Maxwell, Neville. 1970. *India's China War*. New York: Pantheon Books.

McCarty, Nolan, and Meirowitz, Adam. 2007. *Political Game Theory: An Introduction*. Cambridge: Cambridge University Press.

McClelland, Charles A. 1961. The Acute International Crisis. *World Politics*, **14**(1), 182–204.

McGinnis, Michael D. 1992. Bridging or Broadening the Gap? A Comment on Wagner's 'Rationality and Misperception in Deterrence Theory'. *Journal of Theoretical Politics*, **4**(4), 443–457.

Mearsheimer, John J. 1990. Back to the Future: Instability in Europe after the Cold War. *International Security*, **15**(1), 5–56.

Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. New York: W.W. Norton.

Meisner, Maurice. 1999. *Mao's China and After: A History of the People's Republic*. New York: The Free Press.

Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca: Cornell University Press.

Milgrom, Paul R., and Roberts, John D. 1982. Predation, Reputation and Entry Deterrence. *Journal of Economic Theory*, **27**(2), 280–312.

Morgan, Patrick M. 1977. *Deterrence: A Conceptual Analysis*. Beverly Hills: Sage Publications.

Morgan, Patrick M. 2003. *Deterrence Now*. Cambridge: Cambridge University Press.

Morgenthau, Hans J. 1948. *Politics among Nations: The Struggle for Power and Peace*. New York: Alfred A. Knopf.

Morris, Ivan. 1975. *The Nobility of Failure: Tragic Heroes in the History of Japan*. New York: The Noonday Press.

Morrow, James D. 1989a. Bargaining in Repeated Crisis: A Limited Information Model. Pages 207–228 of: Ordeshook, Peter C. (ed), *Models of Strategic Choice in Politics*. Ann Arbor: University of Michigan Press.

Morrow, James D. 1989b. Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining. *American Journal of Political Science*, **33**(4), 941–72.
Morrow, James D. 1989c. A Twist of Truth: A Reexamination of the Effects of Arms Races on the Occurrence of War. *The Journal of Conflict Resolution*, **33**(3), 500–529.
Morrow, James D. 1994a. Alliances, Credibility, and Peacetime Costs. *Journal of Conflict Resolution*, **38**(2), 270–297.
Morrow, James D. 1994b. *Game Theory for Political Scientists*. Princeton: Princeton University Press.
Moul, William Brian. 1988. Balances of Power and the Escalation to War of Serious Disputes among the European Great Powers, 1815-1939: Some Evidence. *American Journal of Political Science*, **32**(2), 241–275.
Mueller, John. 1995. *The Quiet Cataclysm: Reflections on the Recent Transformation of World Politics*. New York: Harper Collins.
Mullik, B. N. 1971. *My Years with Nehru: The Chinese Betrayal*. Bombay: Allied Publishers.
Myerson, Roger B. 1979. Incentive Compatibility and the Bargaining Problem. *Econometrica*, **47**(1), 61–74.
Myerson, Roger B., and Satterthwaite, Mark A. 1983. Efficient Mechanisms for Bilateral Trading. *Journal of Economic Theory*, **28**, 265–81.
Nalebuff, Barry. 1991. Rational Deterrence in an Imperfect World. *World Politics*, **43**(3), 313–335.
Naroll, Raoul, Bullough, Vern L., and Naroll, Frada. 1974. *Military Deterrence in History: A Pilot Cross-Historical Survey*. Albany, NY: State University of New York Press.
Nash, John. 1951. Non-cooperative Games. *Annals of Mathematics*, **54**, 286–295.
Norwich, John Julius. 2003. *A History of Venice*. London: Penguin Books.
Offner, Arnold A. 2002. *Another Such Victory: President Truman and the Cold War, 1945-1953*. Stanford: Stanford University Press.
O'Neill, Barry. 1991. Conflictual Moves in Bargaining: Warnings, Threats, Escalations, and Ultimatums. In: Young, H. Peyton (ed), *Negotiation Analysis*. Ann Arbor: The University of Michigan Press.
O'Neill, Barry. 1992. Are Game Models of Deterrence Biased Towards Arms-Building? Wagner on Rationality and Misperception. *Journal of Theoretical Politics*, **4**(4), 459–477.
O'Neill, Barry. 1999. *Honor, Symbols, and War*. Ann Arbor: The University of Michigan Press.
Oren, Michael B. 2002. *Six Days of War*. Oxford: Oxford University Press.
Organski, A. F. K. 1968. *World Politics*. 2nd edn. New York: Knopf.
Organski, A. F. K., and Kugler, Jacek. 1980. *The War Ledger*. Chicago: University of Chicago Press.
Orme, John David. 1987. Deterrence Failures: A Second Look. *International Security*, **11**(4), 96–124.
Orme, John David. 2004. *Paradox of Peace: Leaders, Decisions, and Conflict Resolution*. New York: Palgrave Macmillan.
Overy, Richard. 1995. *Why the Allies Won*. New York: W.W. Norton.
Palmer, Glenn, and Partell, Peter J. 1999. Audience Costs and Interstate Crises: An Empirical Assessment of Fearon's Model of Dispute Outcomes. *International Studies Quarterly*, **43**(June), 389–405.
Parker, Geoffrey. 1996. *The Military Revolution: Military Innovation and the Rise of the West, 1500–1800*. Cambridge: Cambridge University Press.
Paul, T. V. 1994. *Asymmetric Conflict: War Initiation by Weaker Powers*. Cambridge: Cambridge University Press.
Pearce, David G. 1984. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, **52**(4), 1029–50.
Petersen, Walter J. 1986. Deterrence and Compellence: A Critical Assessment of Conventional Wisdom. *International Studies Quarterly*, **30**(3), 269–294.
Pollack, Jonathan D. 1989. The Korean War and Sino-American Relations. *In:* Harding and Yuan (1989).
Powell, Robert. 1990. *Nuclear Deterrence Theory: The Search for Credibility*. Cambridge: Cambridge University Press.
Powell, Robert. 1993. Guns, Butter, and Anarchy. *American Political Science Review*, **87**(1), 115–32.
Powell, Robert. 1996a. Bargaining in the Shadow of Power. *Games and Economic Behavior*, **15**(2), 255–89.
Powell, Robert. 1996b. Stability and the Distribution of Power. *World Politics*, **48**(Jan.), 239–67.
Powell, Robert. 1999. *In the Shadow of Power*. Princeton: Princeton University Press.

Powell, Robert. 2004. The Inefficient Use of Power: Costly Conflict with Complete Information. *American Political Science Review*, **98**(2), 231–241.

Powell, Robert. 2006. War as a Commitment Problem. *International Organization*, **60**(1), 169–203.

Prescott, William H. 2001. *History of the Conquest of Mexico*. New York: The Modern Library.

Press, Daryl G. 2005. *Calculating Credibility*. Ithaca: Cornell University Press.

Reed, William. 2000. A Unified Statistical Model of Conflict Onset and Escalation. *American Journal of Political Science*, **44**(1), 84–93.

Richards, Diana, Morgan, T. Clifton, Wilson, Rick K., Schwebach, Valerie L., and Young, Gary D. 1993. Good Times, Bad Times, and the Diversionary Use of Force: A Tale of Some Not-So-Free Agents. *Journal of Conflict Resolution*, **37**(3), 504–535.

Richardson, Lewis F. 1960. *Arms and Insecurity: A Mathematical Study of the Causes and Origins of War*. Pittsburgh, PA: Homewood.

Robertson, E. M. 1967. *Hitler's Pre-War Policy and Military Plans, 1933-1939*. New York: The Citadel Press.

Roe, Patrick C. 2000. *The Dragon Strikes: China and the Korean War, June–December 1950*. Novato, CA: Presidio Press.

Roy, Arundhati (ed). 2006. *13 December: The Strange Case of the Attack on the Indian Parliament*. New Delhi: Penguin Books India.

Russett, Bruce. 1993. *Grasping the Democratic Peace*. Princeton: Princeton University Press.

Sagan, Scott D. 1994. From Deterrence to Coercion to War: The Road to Pearl Harbor. *In:* George and Simons (1994).

Sagan, Scott D. 2000. The Commitment Trap: Why the United States Should Not Use Nuclear Threats to Deter Biological and Chemical Weapons Attacks. *International Security*, **24**(4), 85–115.

Sample, Susan G. 1998. Military Buildups, War, and Realpolitik: A Multivariate Model. *Journal of Conflict Resolution*, **42**(2), 156–175.

Sample, Susan G. 2002. The Outcomes of Military Buildups: Minor State vs. Major Powers. *Journal of Peace Research*, **39**(6), 669–691.

Sartori, Anne E. 2002. The Might of the Pen: A Reputational Theory of Communication in International Disputes. *International Organization*, **56**(1), 121–149.

Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton: Princeton University Press.

Schaller, Michael. 2002. *The United States and China: Into the Twenty-First Century*. New York: Oxford University Press.

Schelling, Thomas C. 1956. Essay on Bargaining. *American Economic Review*, **46**(3), 281–306.

Schelling, Thomas C. 1958. The Strategy of Conflict: Prospectus for a Reorientation of Game Theory. *Journal of Conflict Resolution*, **2**(3), 203–264.

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

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven: Yale University Press.

Schelling, Thomas C. 1982. Establishing Credibility: Strategic Considerations. *American Economic Review*, **72**(2), 77–80.

Schultz, Kenneth A. 1999. Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War. *International Organization*, **53**(2), 233–266.

Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. Cambridge: Cambridge University Press.

Schultz, Kenneth A., and Weingast, Barry R. 2003. The Democratic Advantage: Institutional Foundations of Financial Power in International Competition. *International Organization*, **57**(Winter), 1–40.

Selten, Reinhard. 1975. Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International Journal of Game Theory*, **4**(1), 25–55.

Selten, Reinhard. 1978. The Chain-Store Paradox. *Theory and Decision*, **9**(2), 127–159.

Shen, Zhihua. 1996. The Discrepancy between the Russian and Chinese Versions of Mao's October 2 1950 Message to Stalin on Chinese Entry into the Korean War. *Cold War International History Project Bulletin*, 237–242.

Sheng, Michael M. 1995. China's Decision to Enter the Korean War. *Korea and World Affairs*, **19**(2).

Signorino, Curtis S. 1999. Strategic Interaction and the Statistical Analysis of International Conflict. *American Political Science Review*, **93**(2), 279–98.

Signorino, Curtis S. 2002. Strategy and Selection in International Relations. *International Interactions*, **28**(Mar.), 93–115.

Signorino, Curtis S., and Tarar, Ahmer. 2006. A Unified Theory and Test of Extended Immediate Deterrence. *American Journal of Political Science*, **50**(3), 586–605.

Singer, J. David. 1984. *Deterrence, Arms Control, and Disarmament*. 2nd edn. Lanham, MD: University Press of America.

Singer, J. David, Bremer, Stuart, and Stuckey, John. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. Pages 19–48 of: Russett, Bruce M. (ed), *Peace, War, and Numbers*. Beverly Hills: Sage Publications.

Siverson, Randolph M., and Sullivan, Michael P. 1983. The Distribution of Power and the Onset of War. *Journal of Conflict Resolution*, **27**(3), 473–494.

Siverson, Randolph M., and Tennefoss, Michael R. 1984. Power, Alliance, and the Escalation of International Conflict, 1815-1965. *American Political Science Review*, **78**(4), 1057–1069.

Slantchev, Branislav L. 2003a. The Power to Hurt: Costly Conflict with Completely Informed States. *American Political Science Review*, **47**(1), 123–33.

Slantchev, Branislav L. 2003b. The Principle of Convergence in Wartime Negotiations. *American Political Science Review*, **47**(4), 621–32.

Slantchev, Branislav L. 2004a. How Initiators End Their Wars: The Duration of Warfare and the Terms of Peace. *American Journal of Political Science*, **48**(4), 813–829.

Slantchev, Branislav L. 2004b. *Mobilization, Offense Dominance, and Preemptive War*. Manuscript, University of California, San Diego.

Slantchev, Branislav L. 2005. Military Coercion in Interstate Crises. *American Political Science Review*, **99**(4), 533–547.

Slantchev, Branislav L. 2006. Politicians, the Media, and Domestic Audience Costs. *International Studies Quarterly*, **50**(2), 445–477.

Slantchev, Branislav L. 2010. Feigning Weakness. *International Organization*.

Smith, Alastair. 1996. Diversionary Foreign Policy in Democratic Systems. *International Studies Quarterly*, **40**(1), 133–153.

Smith, Alastair. 1998a. Fighting Battles, Winning Wars. *The Journal of Conflict Resolution*, **42**(3), 301–20.

Smith, Alastair. 1998b. International Crises and Domestic Politics. *American Political Science Review*, **92**(3), 623–38.

Smith, Alastair. 1999. Testing Theories of Strategic Choice: The Example of Crisis Escalation. *American Journal of Political Science*, **43**(4), 1254–83.

Snyder, Glenn H. 1961. *Deterrence and Defense*. Princeton: Princeton University Press.

Snyder, Glenn H., and Diesing, Paul. 1977. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton: Princeton University Press.

Spence, Michael. 1973. Job Market Signaling. *Quarterly Journal of Economics*, **87**(3), 355–374.

Strauss, Barry. 2004. *The Battle of Salamis: The Naval Encounter that Saved Greece—and Western Civilization*. New York: Simon and Schuster.

Stueck, William. 1998. *The Korean War as History: David Rees'* Korea: The Limited War *in Retrospect*. Paper presented at the NARA Conference on Cold War documentation, September 25-26, 1998, at the University of Maryland, College Park.

Stueck, William. 2002. *Rethinking the Korean War*. Princeton: Princeton University Press.

Tallett, Frank. 1992. *War and Society in Early Modern Europe, 1495–1715*. London: Routledge.

Tarar, Ahmer. 2005. Diversionary Incentives and the Bargaining Model of War. *International Studies Quarterly*.

Taubman, William. 2003. *Khrushchev: The Man and His Era*. New York: W. W. Norton & Company.

Taylor, A. J. P. 1961. *The Origins of the Second World War*. New York: Touchstone.

Taylor, A.J.P. 1971. *The Struggle for Mastery in Europe, 1848-1918*. Oxford: Oxford University Press.

The Insight Team of the London *Sunday Times*. 1974. *The Yom Kippur War*. Garden City, NY: Doubleday.

Thomas, Hugh. 2005. *Conquest: Montezuma, Cortés, and the Fall of Old Mexico*. New York: Simon and Schuster.

Thucydides. 1996. *The Peloponnesian War*. New York: The Free Press. In Robert B. Strassler, ed. *The Landmark Thucydides: A Comprehensive Guide to the Peloponnesian War*.

Trachtenberg, Mark. 1991. *History and Strategy*. Princeton: Princeton University Press.

Trotter, William R. 1991. *A Frozen Hell: The Russo-Finnish Winter War of 1939–1940*. Chapel Hill: Algonquin Books.

United States Department of State. 1976. *Foreign Relations of the United States, 1950. Volume VII: Korea*. Washington, DC: Government Printing Office.

United States Department of State. 1996. *Foreign Relations of the United States, 1961–1963. Volume VI: Kennedy-Khrushchev Exchanges*. Washington, DC: Government Printing Office.

Van Creveld, Martin. 1977. *Supplying War: Logistics from Wallenstein to Patton*. Cambridge: Cambridge University Press.

Van Evera, Stephen. 1999. *Causes of War: Power and the Roots of Conflict*. Ithaca: Cornell University Press.

von Neumann, John, and Morgenstern, Oskar. 1947. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Wagner, R. Harrison. 1982. Deterrence and Bargaining. *Journal of Conflict Resolution*, **26**(2), 329–358.

Wagner, R. Harrison. 1992. Rationality and Misperception in Deterrence Theory. *Journal of Theoretical Politics*, **4**(2), 115–141.

Wagner, R. Harrison. 1994. Peace, War, and the Balance of Power. *American Political Science Review*, **88**(3), 593–607.

Wallace, Michael D. 1982. Armaments and Escalation: Two Competing Hypotheses. *International Studies Quarterly*, **26**(1), 37–56.

Walter, Barbara F. 2002. *Committing to Peace*. Princeton: Princeton University Press.

Waltz, Kenneth N. 1979. *Theory of International Politics*. New York: McGraw-Hill, Inc.

Wang, Jisi. 1989. The Origins of America's "Two China" Policy. *In:* Harding and Yuan (1989).

Weart, Spencer R. 1988. *Never At War: Why Democracies Will Not Fight One Another*. New Haven: Yale University Press.

Weeks, Jessica. 2008. Autocratic Audience Costs: Regime Type and Signaling Resolve. *International Organization*, **62**(1), 35–64.

Weidenmier, Marc D. 2002. Turning Points in the US Civil War: Views from the Grayback Market. *Southern Economic Journal*, **68**(4), 875–890.

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

Werner, Suzanne, and Kugler, Jacek. 1996. Power Transisions and Military Buildups: Resolving the Relationship between Arms Buildups and War. Pages 187–210 of: Kugler, Jacek, and Lemke, Douglas (eds), *Parity and War: Evaluations and Extensions of The War Ledger*. Ann Arbor: The University of Michigan Press.

Whiting, Allen S. 1960. *China Crosses the Yalu: The Decision to Enter the Korean War*. New York: The Macmillan Company.

Whiting, Allen S. 1975. *The Chinese Calculus of Deterrence: India and Indochina*. Ann Arbor: University of Michigan Press.

Wittman, Donald. 1979. How a War Ends: A Rational Model Approach. *The Journal of Conflict Resolution*, **23**(4), 743–63.

Wohlstetter, Roberta. 1962. *Pearl Harbor: Warning and Decision*. Stanford: Stanford University Press.

Wright, Quincy. 1965. *A Study of War*. Second edn. Chicago: University of Chicago Press.

Xenophon. 1950. *The Persian Expedition*. Revised edn. London: Penguin Classics. Tr. Rex Warner.

Young, Oran R. 1968. *The Politics of Force: Bargaining During International Crises*. Princeton: Princeton University Press.

Zagare, Frank C., and Kilgour, D. Marc. 2000. *Perfect Deterrence*. Cambridge: Cambridge University Press.

Zhang, Shu Guang. 1995. *Mao's Military Romanticism: China and the Korean War, 1950-1953*. Lawrence: Regents of Kansas Press.

Index