Red Lines and Faits Accomplis in Interstate Coercion and Crisis

by

Daniel W. Altman

Submitted to the Department of Political Science at the Massachusetts Institute of Technology on February 13, 2015 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Political Science

ABSTRACT

The International Relations literature has an established view of interstate crises that explains how states pursue victory in terms of signaling resolve. States make gains with credible coercive threats (compellence). In contrast, this dissertation conceives of each crisis as a strategic competition between a challenger seeking to make gains unilaterally by fait accompli and its adversary’s countervailing efforts to set red lines to deter these faits accomplis. After clarifying the neglected concepts of “red line” and “fait accompli,” the dissertation takes up two questions the literature has left unexplored: When are faits accomplis likely to occur? When are they likely to lead to war? The result is a theory of coercive conflict explaining why deterrent red lines that contain any of four weaknesses – types of gray areas, in essence – are especially vulnerable to faits accomplis.

This theory is tested with two case studies – the 1948-1949 Berlin Blockade Crisis and the 1962 Cuban Missile Crisis – and an analysis of gray areas and land grabs in territorial crises since 1918. Making extensive use of declassified documents, the case studies show that the “game” of crises need not be a matter of convincing the adversary of one’s willingness to fight. Instead, states pursue victory by finding gray areas and other weaknesses in deterrent red lines that they can exploit to unilaterally take as much as possible – often by fait accompli – without crossing the line and overtly firing on the other side. Crises, from this standpoint, are a game of finding ways to advance without attacking.

The analysis of territorial crises makes use of original data on all land grab faits accomplis since 1918. It shows first that states far more often make territorial gains by fait accompli than by coercing a territorial cession. It then focuses on the impact of geographical gray areas, which take two forms: islands located awkwardly between two core territories and border ambiguities. It finds that two-thirds of all land grabs since 1918 targeted a gray area. These gray areas render faits accomplis more effective at making a gain without provoking war and, consequently, more likely to occur.

Thesis Supervisor: Barry Posen
Title: Ford International Professor of Political Science
I would first like to thank my advisor, Barry Posen. I pursued working with Barry for three reasons. First, he would push me towards a more strategic, practical, and concrete framing of the project. Second, he would force me to fight my propensity toward dense theory-laden writing and instead pursue clarity. Third, and most importantly, he gives comments of both a quality and quantity that are unparalleled. He did all three and more. I thank him for his many, many contributions.

My committee as a whole contributed immensely to the project. Ken Oye saw promise in it all the way back during my first year at MIT – long before anyone else could comprehend what I was trying to say. His support at that early stage was invaluable. Vipin Narang has been an excellent source of not just substantive comments but also my go-to source of advice on all manner of career and practical issues. Phil Haun was always the first to read my drafts and send me insightful comments.

Anyone who has attended understands just how important the tight-knit grad student community is in the MIT Political Science Department and – especially – the Security Studies Program. Although I cannot possibly list everyone that I should, let me thank Noel Anderson, Lena Andrews, Mark Bell, Nathan Black, Christopher Clary, Fiona Cunningham, Jeremy Ferwerda, Tobias Harris, Chad Hazlett, David Jae, Sameer Lalwani, Marika Landau-Wells, Joyce Lawrence, Jon Lindsay, Krista Loose, Michele Margolis, Phil Martin, Andrew Miller, Nicholas Miller, Reid Pauly, Miranda Priebe, Kai Quek, Josh Shifrinson, Joseph Torigian, David Weinberg, Steve Wittels, Alec Worsnop, Ketian Zhang, and many others not named here. Among the faculty, I thank Fotini Christia, Taylor Fravel, Frank Gavin, Roger Petersen, and Steve Van Evera for their guidance and input during the process. In the broader IR community in Cambridge, I would also like to thank Monica Toft, Ivan Arreguin-Toft, Molly Roberts, Brandon Stewart, Walt Cooper, and Yuri Zhukov. I am particularly grateful to Jeffrey Friedman for being an exceptional sounding board for new ideas.

I owe a debt of gratitude to the Center of International Security and Cooperation at Stanford University. Given the size of his contributions, I have come to see Ken Schultz as a de facto member of my committee. I thank him for both his excellent comments and his generosity with his time. Scott Sagan has gone above and beyond in finding ways to help me develop the skills to have an impact with my research, both with respect to this project and in general. I would also like to thank Lynn Eden, Jim Fearon, David Holloway, Amy Zegart, and quite a few PhD students in the Stanford Political Science Department. Many of my fellow fellows at CISAC also contributed to the project, including Neil Narang, Brad Roberts, James Cameron, Erin Baggott, Edward Geist, Jonathan Hunt, David Traven, and Rebecca Slayton. The Stanton Foundation and MacArthur Foundation each funded a one-year fellowship at CISAC without which the project would not be where it is today.
Finally, I thank my family and friends for supporting me throughout this process and/or distracting me so that it took so long. My parents, Drew Altman and Pam Koch, supported me in too many ways to count, as they have throughout my life.
Table of Contents

Ch1: Introduction ......................................................... 7

Ch2: A Theory of Red Lines and Faits Accomplis ....................... 15

Ch3: The Conventional Wisdom: Crisis Strategy as Signaling ........... 43

Ch4: Gray Areas and Land Grabs in Territorial Crises ................... 59

Ch5: The Berlin Blockade Crisis of 1948-1949 .......................... 90

Ch6: The Cuban Missile Crisis of 1962 ................................ 134

Ch7: Conclusion .................................................................. 175
Ch. 1 Introduction

This book is about the interplay between two tools of statecraft: the red line and the \textit{fait accompli}. Although policymakers have turned to these tools throughout history, the academic field of International Relations has dedicated surprisingly little attention to them. No accepted concepts or definitions exist. Nor do theories of when each tool tends to succeed, tends to fail, or tends to lead to war. This book contributes to filling these gaps.

The importance of red lines starts with the question: When does deterrence succeed in forestalling aggression? The canonical answer is that deterrence is first and foremost about bringing to bear sufficient capability and resolve to bear to forestall an attack. This is so, but deterrence is also about drawing lines – red lines – that distinguish acceptable behavior from unacceptable transgressions that will be met with retaliation.

Setting strong red lines is no simple task. It may be clear that inaction is acceptable and that an all-out invasion is unacceptable, but what of the countless shades of gray in between? What the adversary challenges the deterrent threat over a small issue that is distinctly separate from what is truly important? Is it as fully protected? During the Cold War, the United States consistently had far greater confidence in its deterrent red line protecting its NATO allies in general than in its ability to deter a Soviet military operation to seize West Berlin. What if something does not fall clearly within the deterrent red line? For years the United States sought to deter North Korea from building nuclear weapons, but did this red line extend to operating a nuclear reactor? What if there is a way to get around the line and seize what it protects without violating it? The Soviet Union may have had a credible deterrent red line against an American attack, but this line failed to extend to deterring indirect attack via the arming of Afghan \textit{mujahideen} in the 1980s.

Just as the challenges of setting strong red lines are both considerable and under-appreciated, the strategy of the challenger seeking to undermine, bypass, or get away with violating these deterrent red lines requires renewed analysis. Some deterrent red lines fail entirely, like that of the Soviet Union in 1941 when Nazi Germany launched its surprise invasion. Others fail when the deterrer backs down under pressure, as when Britain, France, and Czechoslovakia gave up on deterring a German seizure of the Sudetenland in 1938. However, alongside these lies what may be the most common – and yet most neglected – form of deterrence failure: the \textit{fait accompli}. With this strategy, challengers take a gain unilaterally, putting the deterrer in the unenviable position of having to decide between relenting to the loss and escalating in retaliation.

At issue is a basic question about how states make gains in crises. Most thinking on interstate crises emphasizes coercive bargaining. For states that desire something from an adversary, the first step is to make a coercive threat demanding the concession. After issuing this threat, the second step is signaling resolve. Through actions like mobilizing forces and firm public statements of commitment,
states convince the adversary that their threat is credible. When the threat is credible, the adversary relents and grants a concession in order to avoid war.

This theory of victory in crisis is undeniably compelling. It fits important cases, including Nazi Germany’s successful effort to intimidate Czechoslovakia into relinquishing the Sudetenland and the Soviet Union’s withdrawal of nuclear missiles from Cuba under threat from the United States.

Despite its appeal, the problem with the conventional wisdom is a simple one: how often does this actually happen? Do states make gains by coercing concessions? Or instead by unilaterally taking those gains by fait accompli? I believe an exhaustive analysis of this question across all varieties of crises would show that although both occur at times, the fait accompli is by far the more common way for gains to be made. In this book, I establish that this is so in all territorial crises from 1918-2007 and in two of the most important crises of the Cold War: the Berlin Blockade Crisis (1948-1949) and the Cuban Missile Crisis (1962).

The literature has lavished attention on signaling and coercing concessions while giving short shrift to faits accomplis. This has left important questions unanswered and unaddressed. When are faits accomplis likely to occur? When and how do states succeed in setting red lines that successfully deter faits accomplis? When can states get away with faits accomplis without provoking war?


The starting point to answering these questions is to conceive of each crisis as a strategic competition between a challenger seeking to make gains by fait accompli and its adversary’s countervailing efforts to set red lines to deter them. Deterring faits accomplis seizing something of limited value is inherently difficult, because threatening war may not be credible when the issue in
dispute is not worth a war. This problem is also very general. All states have many small interests — including individual pieces of territory — to protect. With no solution, states would be vulnerable to being picked apart piece by piece in what Thomas Schelling labelled “salami tactics.”

In the next chapter, I argue that the best solution to the problem of deterring \textit{faits accomplis} is often to rely on what I refer to as a strong red line. In coercion, every threat contains a demand, and all demands set a red line distinguishing compliance from noncompliance. Strong deterrent red lines are those that make use of a focal point to encompass multiple smaller units into one larger whole that states can more credibly threaten to defend. The focal point makes for an unmistakable, stark red line. Borders are the most well-known example. Failing to uphold the deterrent red line protecting even a small part of this larger whole creates the problem that focal points help to resolve: “If not here, where?” Correspondingly, weak red lines leave any of four openings for \textit{faits accomplis}: arbitrariness (lacking a focal point), imprecision, unverifiability, and incompleteness (vulnerability to outflanking). These weaknesses in red lines often take the form of gray areas.

Two central predictions flow from this logic. The first: gray areas (and the other weaknesses) in deterrent red lines make \textit{faits accomplis} more likely to succeed in making gains without provoking war. Consequently, the second: gray areas make \textit{faits accomplis} more likely to occur.

India’s seizure of the Siachen region of Kashmir in 1984 offers an example what I have found to be a typical \textit{fait accompli}, i.e., one that targets a gray area and avoids war. The Simla Agreement ending the 1971 Indo-Pakistani War established the Line of Control dividing Kashmir. However, the negotiators left a gray area at the northern end, an inhospitable mountainous area centered on the Siachen Glacier with no population or economic value. The agreement read, “Thence north to the glaciers.” This phrase proved to several competing interpretations, notably straight north vs. in continuation of the northeasterly direction the border ran up to that point. The result was a gray area between these interpretations. By the 1980s, both sides claimed the region and the temptation to score a political victory by advancing territorially had grown stronger. Reacting to intelligence reports of an imminent Pakistani \textit{fait accompli} seizing Siachen, India moved first. The Indian Army advanced rapidly by helicopter and established military posts on the mountain peaks on the far side of the glacier before Pakistani forces could react. The result was a series of small battles, but India retained control of the disputed region and Pakistan forwent any immediate escalation elsewhere.\footnote{V. R. Raghavan, \textit{Siachen: Conflict without End} (Viking: Penguin Books India, 2002), pp. 22-23, 27, 36, 42, 50, 55.} India’s \textit{fait accompli} succeeded in seizing the gray area.

Conversely, in 1999, Pakistani forces operating disguised as Kashmiri insurgents occupied small but strategically significant positions on India’s side of the Line of Control in Kashmir. This \textit{fait accompli} seizing a small, unoccupied territory in Kashmir seems all too similar to what India did fifteen years earlier. Yet the case differs in one critical respect: the strength of the deterrent red line, i.e., the lack of a gray area in Kargil. The Line of Control provided a clear focal point, making for a strong Indian red line that Pakistan nonetheless transgressed. Like other \textit{faits accomplis}, Pakistan’s strategy relied on the
hope that India would relent to such a small loss rather than fight. This sort of seizure of a small, low-value territory across a clearly-defined border (i.e., despite a strong red line) could happen all the time. If borders are but lines on a map and if focal points matter little, then small border territories should be vulnerable to this sort of fait accompli regardless of how clearly the line was drawn. Yet the Kargil operation is a striking aberration. Attempting to seize a small piece of what is clearly a part of another state’s territorial core – violating a clear focal point in the process – is surprisingly rare. There are very few cases like Kargil.

The Kargil operation was an utter debacle for Pakistan, provoking Indian counterattacks that proved willing to sacrifice the lives of more than 500 Indian soldiers in order to retake the territory. Indian forces were careful not to cross the relevant focal point (the Line of Control), even in the air as they conducted bombing operations only within their own territory. The failure of the operation contributed to Pakistani Prime Minister Nawaz Sharif’s fall from power in a military coup. More telling than the disastrous outcome for the operation, though, is how rarely other states have taken similar actions, before or since.

Overall, this book lays out a different basic perspective on interstate crises. The traditional way to understand crises is to suppose that policymakers think about crisis strategy as if asking themselves, “How can we convince the other side that we are willing to fight in order to convince them to back down?” Although an undisputedly important question that I explore further, in this book I approach crises as if statesmen ask themselves, “What can we get away with unilaterally taking without starting a war?”

Research Design and Findings

To test this theory of red lines and faits accomplis, I conduct a statistical analysis of gray areas and ‘land grab’ faits accomplis in interstate crises over territory from 1918 to 2007 and two case studies: the Berlin Blockade Crisis of 1948-1949 and the Cuban Missile Crisis of 1962. For both lines of investigation, I selected cases that offered significant methodological advantages while still establishing the theory’s ability to shed light on some of the most important crises of the 20th Century.

The most credible statistical test of the theory requires a focus on interstate crises over territory. A broader analysis of red line characteristics, faits accomplis, and crisis outcomes in all the interstate crises regardless of issue type would suffer from serious comparability and inference problems. Is unilaterally fishing in disputed waters comparable to unilaterally occupying disputed territory? Is a gray area in a red line against seizing territory comparable to ambiguity in a red line against arming a rebel group? Focusing on territorial crises avoids these problems.

---

Analyzing territorial crises offers three additional advantages. First, territory is the most common issue area over which interstate crises arise. Second, territorial crises are particularly significant for the causes of war because states so often fight over territory. Third, territorial crises contain two unusually exogenous sources of gray areas in deterrent red lines: islands and border ambiguities.

The analysis of territorial crises (Chapter Four) relies on original data I have collected on all land grab faits accomplis since 1918. Russia’s 2014 annexation of the Crimean Peninsula offers a typical example. I focus on two questions: When do states try to get away with land grabs? When do they succeed without provoking war? I find that the nature of the deterrent red lines meant to stop these land grabs is an important part of the answers. There are two main sources of gray areas in territorial red lines. First, the disputed territory is an island or island-like peninsula located awkwardly between two states’ core territories but physically separate from both. Second, the territory falls in a border ambiguity due to poor delimitation or demarcation, a problem often rooted in colonial-era boundaries. Both types of gray areas constitute territories which fail to lie clearly and unmistakably within the core territory of a state; hence the term “gray area.” Chapter Four provides a more detailed summary of the research design for this line of inquiry.

This analysis shows the following. First, states are much more likely to acquire territory by fait accompli than by coercion. Since 1918, there have been 10 coerced territorial cessions and 88 land grabs. Second, approximately two-thirds of all land grabs since 1918 targeted one or the other type of gray area. Third, land grabs are approximately twice as common in crises over gray area territories than in crises over other territories. Finally, land grabs against gray areas are half as likely to provoke war as land grabs against other territories. States generally cannot get away with land grabs unless they target gray areas.

The case studies rely on declassified documents to show how characteristics of red lines shaped two of the Cold War’s most important crises. By comparing each side’s perceptions of what would and would not lead their adversary to escalate to war, I show that the United States and Soviet Union worried more about violating strong red lines than impinging on their adversaries’ interests. By comparing American policymakers’ perceptions of the various strategies available to them, I show how they identified and exploited weaknesses in Soviet red lines.

In the Berlin Blockade Crisis, the Western Powers believed that forcing transport aircraft through to Berlin would succeed, whereas attempting to force through a convoy of trucks and armed vehicles would end in disaster – this despite both actions producing the same strategic effect of undermining the Soviet blockade. Both sides sought to avoid crossing what I refer to as the firing-on-forces focal point. On land, the West would likely have to push through Soviet troops to reach Berlin. In contrast, the Soviets needed to fire first to stop the airlift. This simple difference proved decisive. Some American policymakers even anticipated that the Soviets would refrain from engaging a truck convoy with force at the border and would instead block roads and destroy their own bridges to strand the convoy.
The guiding logic of crisis strategy is taking unilateral actions designed to maneuver around strong red lines, pushing forward as far as possible without definitively violating them. The pivotal red line was against overt attack.

However, units of the Soviet military did fire on American transport aircraft flying food and fuel into Berlin, and these incidents scarcely caused a stir. The reason: the Soviets fired high-powered searchlights at the aircraft, trying to bring them down by blinding the pilots as they attempted to land in Berlin (unsuccessfully). Had Soviet forces fired bullets or shells for exactly the same purpose, the consequences would have been a dramatic increase in tensions and the likelihood of war. The Soviets were deterred from an overt attack, but not from a more ambiguous form of attack – searchlights – that targeted a gray area in the red line. American lives were lost to plane crashes during the Berlin Airlift and therefore due to Soviet policy, but none due to overt attack. Enduring the loss of a plane and its crew to an accident was acceptable. Losing them to Soviet searchlights would have sown controversy without risking war. Losing one to a Soviet fighter or anti-aircraft artillery battery, however, would have created a real risk of war. The only difference was the red line.

In contrast, traditional signals of resolve were less important and less credible in the Berlin Blockade Crisis than the literature has assumed. For instance, the American deployment of 60 B-29 bombers to Britain at the height of the crisis is cited as a prototypical example of “nuclear saber-rattling.” I show that it had little impact on the course of the crisis, and that it was never expected to do so.

The theory also explains important aspects of the Cuban Missile Crisis, the case which has inspired so much of the literature’s emphasis on signaling and coercion. The Soviet Union sparked the crisis by attempting to deploy ballistic missiles to Cuba in a covert fait accompli, initially falling under the radar of the American attempts to uphold red line. The efficacy of the American blockade of Cuba hinged not on the ability to draw an arbitrary – and therefore weak – red line at sea, but rather on which side could prevail through unilateral actions without overtly attacking the other during an intercept of a Soviet freighter by the U.S. Navy.

I do find that signaling was a critical part of how the United States pressured the Soviet Union to withdraw the missiles. The efficacy of American signals at establishing credibility came in part from the traditional methods – public statements, mobilization, etc. The conventional wisdom is vindicated in this regard. Yet even with this aspect of the case that the conventional wisdom best explains, fears concerning minor violations of strong red lines played a major role. Specifically, Khrushchev’s decision to promptly remove the missiles was galvanized in large part by the uncontrollable chance that the use-of-force red line would be crossed when Cuban or Soviet anti-aircraft batteries fired on American surveillance aircraft. The risk of crossing a strong red line proved vital to American brinksmanship.

I conduct the case studies with what I refer to as a “unique observable implications approach.” Observable implications are visible predictions of a theory that can be checked against evidence. For
example, the absence of wars among democracies is an observable implication of democratic peace theory. *Unique* observable implications are predictions of a theory that alternative theories – both formal and common sense – do not predict. *Certain* observable implications, conversely, are things that must be observed for the theory to plausibly explain the case.

The procedure is as follows. Think through the way the theories being tested apply to the case. Identify the observable implications of these theories. Establish through deductive logic which observable implications are certain or unique and to what extent. Evaluate them against all available evidence. When the process of researching the case uncovers additional unique or certain observable implications, exploit them as well. Divide the case study into sections for each of these informative observable implications, or for sets of related observable implications. 6

The unique observable implications approach also guides case selection. The objective is to maximize the number and quality of unique and certain observable implications. Because the approach relies on fine-grained details of cases, it is impossible to know in advance exactly which case will contain the most informative observable implications. But, it is possible to pick cases likely to be rich in them. For this project, particularly complex crises with multiple red lines and complex avenues for challenging them offered the best chance of yielding the most informative observable implications. I selected two crises featuring blockades on that basis. These cases are not meant to be representative, but rather to best test the theory. 7

Moreover, in order to conduct tests based on these especially-informative fine-grained details in cases, it is generally important to have access to extensive primary source accounts of decision-making deliberations. Language constraints come into play as well. The early Cold War era met these data requirements because vast collections of typed documents were created and most have now been declassified, at least in the United States. All of this reasoning points to a relatively small set of cases from which to choose. The Berlin Blockade Crisis and the Cuban Missile Crisis were among the most promising.

**Plan of the Book**

The book divides into three parts. Chapters Two and Three discuss the theory and its distinctiveness from the conventional wisdom understanding of crisis. Chapter Four presents the

---


7 Selecting two representative cases would not yield a sufficient sample size to claim generalizability on that basis, so I do not view this as a downside of this approach.
quantitative analysis of gray areas and land grabs in interstate crises from 1918 to 2007. Chapter Five and Six consist of the Berlin and Cuba case studies. Chapter Seven concludes.

Chapter Two lays out the theoretical rationale for expecting weaknesses in red lines such as gray areas to make *faits accomplis* more successful and therefore more common. It begins with clear definitions of the concepts. It then explains why strong red lines are important for deterring *faits accomplis*, following this discussion with a review of related arguments in the literature. The chapter also explores why weak red lines exist when they pose such significant disadvantages for the states setting them.

Chapter Three summarizes the conventional wisdom on interstate coercion and crisis. Part I focuses on the traditional structural variables, most important among them the balance of power and the balance of interests. Part II develops the logic of signaling resolve in crises. Part III builds on this to explore the policies through which states are thought to be able to signal resolve effectively.

Chapter Four encompasses the analysis of gray areas and land grabs in interstate crises over territory from 1918 to 2007. It discusses the newly-compiled data on gray areas in these crises and on all land grabs since 1918. It then examines the linked questions of what makes land grabs more likely to occur and what makes land grabs more likely to provoke war.

Chapters Five and Six conduct case studies of the Berlin Blockade Crisis of 1948-1949 and the Cuban Missile Crisis of 1962. The chapters ask questions such as: why did American policymakers believe that resupplying Berlin by air would succeed but trying to do so by truck would end in a debacle? Did the United States emphasize signaling resolve during the Berlin Crisis, for instance by deploying B-29 bombers to Britain in an act of “nuclear saber-rattling”? Did it work? In the Cuban Missile Crisis, why did the United States choose blockade over airstrike or invasion? Why did Soviet Union elect not to run the blockade? What made that red line hold? What led the Soviet Union to withdraw its missiles from Cuba?

Finally, Chapter Seven concludes with a more informal discussion of the different ways in which states can manage crises. It frames a set of options contrasting a strategy of exploiting weaknesses in red lines with *faits accomplis* against alternatives, including preemptive attack, brinksmanship, and proportional retaliation. In part because these alternatives foment instability and risk war to a greater extent, a strategy of judicious, targeted *faits accomplis* is often the most appealing way to make gains in crises.
Ch. 2: A Theory of Red Lines and Faits Accomplis

Deterrence is not an all or nothing enterprise. States must deter outright invasion and conquest, but they also must deter more limited unilateral violations, which often take the form of faits accomplis. These violations come in numerous forms, ranging from seizing a small piece of territory to supplying arms to a rebel group to building another centrifuge. Faits accomplis challenge states because making deterrent threats credible is difficult when the issue in dispute is not worth a war. This problem is quite general because all states have many small interests (e.g., individual pieces of territory) to protect. Without a solution, states would be vulnerable to a series of individually limited violations that could sum to much more. How do states manage this problem, and avoid being picked apart piece by piece?¹

The best solution is often to rely on a strong red line. In coercion, every threat contains a demand, and all demands set a red line dividing compliance from noncompliance. Strong red lines are those that make use of a focal point to encapsulate smaller units into one larger unit that states can more credibly threaten to defend, and do so without leaving any of several types of openings for faits accomplis.

Weak red lines suffer from one or more of four vulnerabilities: arbitrariness, imprecision, unverifiability, and incompleteness. First, arbitrary red lines are those that are not set on focal points. These focal points are unique, conspicuous, and clearly different from nearby alternatives. Second, imprecise red lines leave gray areas in which it is ambiguous whether certain actions would or would not violate them. Third, unverifiable red lines are set such that it will not be immediately clear whether or not they have been violated. Fourth, incomplete red lines are those that can be bypassed, the object they aim to secure taken without violating the red line meant to protect it. The history of interstate crises is replete with faits accomplis targeting these gray areas in red lines.

When the adversary calculates correctly, its fait accompli will often succeed in yielding a limited gain. When the adversary miscalculates the deterrer’s willingness to fight for whatever has been seized, the result will be war. Either way, vulnerable red lines tend to produce undesirable outcomes for the states that set them.

It was not by accident that as Iran and North Korea proceeded towards the capability to construct nuclear weapons, both did so with a long series of incremental and deniable steps rather than one overt public decision to get the bomb. That distinction is perhaps the easiest way to understand the

significance of the theory put forward in this chapter. Had Iran or North Korea been obliged to make a single, up-front, all-or-nothing, Rubicon-crossing public decision about whether or not to pursue nuclear weapons, this chapter explains why it would have been more difficult – and therefore less likely – for them to do so. However, as I discuss further below, each exploited a variety of red line vulnerabilities to inch closer and closer to the bomb.

When Italy defied Britain, France, and its legal obligations under the 1936 Non-Intervention Agreement by intervening in the Spanish Civil War, the policy was not implemented with an official announcement and the immediate large-scale deployment of Italian forces. Instead, Italy deployed its forces gradually and secretly, initially referring to them as “volunteers,” providing arms covertly, and steadily tightening a blockade on Republican-held ports for which it denied responsibility. The first contingent of Italian Black Shirts traveled to Spain unarmed and wearing civilian clothing. Even Italy’s extensive use of submarines for that blockade is attributable to the inability to conceal the nationality of surface warships. Italy escaped any meaningful retribution.²

The traditional way to understand crises is to suppose that statesmen approach them by asking themselves the question, “What can we do to convince the other side that we are willing to fight in order to get them to back down?” Crises, from this standpoint, revolve around signaling resolve. Although an undisputedly important perspective that I explore further in Chapter Three, in this chapter I approach crises as if statesmen ask themselves a different guiding question, “What can we get away with taking unilaterally without starting a war?” The result is a theory of coercive conflict which distinguishes strong and weak red lines, explaining how red lines with any of four weaknesses leave openings that adversaries can exploit with faits accomplis. This theory builds on neglected insights from Thomas Schelling’s classics Arms and Influence and The Strategy of Conflict.³

The structure of the chapter is as follows. It first explains and defines the basic concepts and variables of the study. The definitions of “red line” and “fait accompli” are particularly significant because the field lacks conventional definitions of these two underutilized concepts. It then lays out the core logic of the relationship between red lines, faits accomplis, and crisis outcomes. This discussion explains why ‘strong’ red lines matter and why vulnerabilities in red lines elicit faits accomplis. It next explores what the existing literature has to say about each vulnerability. Finally, the chapter defines the

² John F. Coverdale, Italian Intervention in the Spanish Civil War (Princeton: Princeton University Press, 1975), pp. 117, 169, 200, 204; Richard Smoke, War: Controlling Escalation (Cambridge: Harvard University Press, 1977), pp. 54-56, 71. Smoke concludes that British and French red lines (“response thresholds”) were imprecise in a way that produced uncertainty and thereby inspired caution, but it is clear from his discussion that Italy and Germany sought and utilized this ambiguity to aid the Nationalists with a series of unilateral interventions.
³ Virtually every major point in the first hundred or so pages of these books has become a mainstay of the literature on coercion, except for the lengthy sections on what I refer to as red lines. Schelling placed no small amount of emphasis on these sections, but despite that, the insights have gone largely unutilized. This chapter aims in part to reformulate and clarify these ideas in order to bring them to light. Thomas C. Schelling, Arms and Influence (New Haven: Yale University Press, 1966); Thomas Schelling, The Strategy of Conflict (Cambridge: Harvard University Press, 1960).
scope conditions of the theory and considers the question of why vulnerable red lines exist when they have such detrimental consequences for the states setting them.


This book addresses two broad research questions: What explains the actions taken in interstate coercive conflicts? What explains the outcomes of these conflicts? More informally, how do states go about trying to win in a crisis? When and how does crisis strategy work? The theory provided offers partial answers to each question by focusing on the effects of vulnerable red lines and the causes of *faits accomplis*. To lay the groundwork for this theory, it is first necessary to explain and define the major variables and concepts.

The domain of the project is interstate coercive bargaining. Coercive bargaining is defined as bargaining under the threat of punishment. This punishment often takes the form of starting a war, but can instead consist of more limited actions such as economic sanctions or punitive bombardments. In coercive bargaining, the coercer threatens to harm the interests of the target if a demand is not met. The target ultimately retains the decision between resistance and capitulation. By definition, the outcome is not simply imposed, however severe the level of duress. Ironically, for that reason this definition of coercive bargaining excludes *faits accomplis*, which are unilateral impositions. However, *faits accomplis* normally represent the failure of deterrent threats. The two are inherently linked within a strategic competition between *faits accomplis* and the red lines meant to deter them.

The term ‘coercive bargaining’ is intentionally redundant. Coercion is, after all, a form of bargaining that Schelling (1966) aptly described as “dirty bargaining.” The purpose of the redundancy is to make clear my intention to draw from and contribute to the two bodies of literature in International Relations that too often deal separately with the same phenomenon. I refer to these as the coercion literature, which primarily follows a realist approach, and the bargaining literature, which primarily follows a rational choice approach.

---

4 Although I refer to threats of punishment and generally discuss red lines in the context of punishment-type coercion, the theory in this chapter also applies to denial-type coercion. On this distinction, see Robert A. Pape, *Bombing to Win: Air Power and Coercion in War* (Ithaca: Cornell University Press, 1996).


In the generic threat, “If you X, then I Y,” demands are X. This definition suits both threats of the form “If you do X, I will do Y” (deterrence) and “Unless you do X, I will do Y” (compellence).9 These demands can be explicit or tacit, as long as they are conveyed.

The conventional wisdom on interstate coercive bargaining assumes that demands are simple. One side in a conflict demands a certain amount of whatever is under dispute and threatens a punishment if that demand is rejected. If the threat is credible and the demand is not too high, the other side agrees. The strategic calculation behind setting a demand is one of assessing how much can be obtained such that the other side still agrees to the terms.10 This simple view of demands is compelling, but it misses a series of ways in which properties of demands other than how much is demanded have powerful effects on interstate coercion and bargaining. These properties of demands are easiest to think of as characteristics of red lines.11

---


9 Schelling, Arms and Influence, p. 69.


11 I make no claim to have addressed all possible characteristics of demands beyond how much is demanded. One further characteristic of demands (but not red lines) that matters is attributability. If an adversary complies with a demand, how clear is it that this action was due to the demand rather than to other considerations? Attributability is critical for capitulations to affect reputations. Unattributable demands are more likely to succeed in eliciting concessions because they entail lower reputational costs. Thomas Schelling also discusses two other
I define “red line” as the part of a coercive demand which distinguishes compliance from violation. Facing a coercive demand, an adversary has a range of options it can take. The coercer must divide these into the categories of compliance and noncompliance. The coercer threatens punishment for noncompliant actions falling on one side of the red line, and declares (or implies) the intent not to punish for compliant actions falling on the other. Red lines are therefore a necessary component of all threats, even though the vast majority of threats do not use the term “red line” explicitly. States do not always set red lines, but red lines are an essential element of all coercive threats. A case without any red lines is a case of something other than coercive bargaining. These red lines are heavily constrained by the actions sought from the adversary and the context in which the demands are made; they are not mere semantics.

The term “red line” entered common usage in English through an overly literal translation of the Hebrew phrase “kav adom,” with the same meaning. This occurred through the Israeli government’s declaration of a “Red Line” against increased Syrian military involvement in Lebanon in 1976. Use of the term has steadily risen since then, particularly in the 21st century. Although there are many prior usages of the phrase “red line,” “redline,” and “thin red line,” they do not carry the same meaning. “Line in the sand” is an older phrase with a more similar meaning; “red line” has increasingly supplanted it in common usage, perhaps because it avoids the rhetorical problem that lines in the sand tend not to survive even a modest breeze.

Historically, most deterrent red lines were never referred to as “red lines,” and the term is still often not used when it could be today. For instance, it still appears disproportionately in Middle Eastern contexts. Statesmen, however, have always had ample vocabulary for setting an unmistakable red line without making use of the phrase, as when President John F. Kennedy said, “It shall be the policy of this nation to regard any nuclear missile launched from Cuba against any nation in the Western Hemisphere as an attack by the Soviet Union on the United States, requiring a full retaliatory response upon the Soviet Union.”

Even in International Relations contexts, “red line” is sometimes used in slightly different ways. The most different usage refers not to the dividing line between compliance and noncompliance but instead to the minimum level of concession needed from the other side such that a party will accept a

---

12 Schelling, Arms and Influence, pp. 4, 74. I make no assumption about the veracity of these claims; bluffs and aggression despite capitulation are important possibilities.
bargain rather than resort to punishment/force. These two overlap when states make sincere demands that the adversary not transgress past the point where the demander finds fighting preferable to enduring the *fait accompli*. Nonetheless, the two definitions will diverge in some cases. These differences are most often found in the form of transitory positions taken in the course of negotiations, which are not the focus of the book.

The term “red line” has also been increasingly associated with explicitness (i.e., excluding implicit threats) and threats to use force (i.e., excluding threats of lesser punishments). I view these as separate variables, essentially add-ons to the underlying concept of a divider between compliance and violation. Although these variables are important, imposing them as restrictions on the definition of “red line” risks artificially segmenting the phenomena. It also risks obscuring the significance of red lines through conflation with the analytically distinct questions concerning explicitness and what punishment is threatened. Moreover, neither of these variables has as simple an effect on outcomes as one might think. Lower magnitude threats may offset their reduced heft by being cheaper to implement, and therefore more credible. Explicitness need not imply intentions. The United States explicitly left South Korea outside its defensive perimeter prior to the Korean War, yet nonetheless decided it needed to defend South Korea when the North invaded in 1950. Many of the strongest red lines are never explicitly stated to the adversary. It is hardly necessary to tell a rival to refrain from sending its army across the border for that red line to be keenly perceived and understood. In the present, South Korea need not make explicit its deterrent red line against a North Korean invasion across for North Korea to perceive that deterrent threat. Because I reject including these added requirements in the definition of “red line,” I use the term more widely than it sometimes appears in common usage.

The explanatory variables are four potential vulnerabilities in the red line component of demands: arbitrariness, imprecision, incompleteness, and unverifiability. In the following section, I provide definitions and follow these with a set of examples from coercive nonproliferation. I sometimes refer to these vulnerabilities collectively as “gray areas,” although this term suits some vulnerabilities better than others.

*Faits accomplis* are the intermediate variable for the study, meaning that at times I examine the causes of *faits accomplis* (red line characteristics) and at times I examine the effects of *fait accomplis* (on outcomes). The overall theory connects red line characteristics to outcomes through *faits accomplis*.

---

17 Speed limits nicely illustrate this distinction. A speed limit of 65 miles per hour offers a precise red line at 65mph by the definition I use, albeit a somewhat arbitrary and partially unverifiable one (because speed sensors have a margin of error). In contrast, the line at which a police officer will actually pull a car over and issue a speeding ticket, which corresponds to the alternative definition, is higher and less clearly defined. Thanks to Christopher Clary for this example.
18 This is in part a side-effect of strong red lines being set on focal points, as these focal points allow the adversary to perceive the red line without needing to be told.
I define a *fait accompli* as “making a limited unilateral gain at an adversary’s expense in an attempt to get away with that gain when the adversary chooses to relent rather than escalate in retaliation.”¹⁹ Sometimes this strategy works, as with Russia’s 2014 annexation of the Crimean Peninsula. Other *faits accomplis* provoke a stronger response than had been hoped. Pakistan’s 1999 infiltration of forces to occupy positions on India’s side of the Line of Control in the Kargil district of Kashmir offers an example. The Soviet deployment of missiles to Cuba in 1961 and Germany’s remilitarization of the Rhineland in 1936 offer two more. *Faits accomplis* take many forms in addition to seizing territory, including the construction of a nuclear reactor in violation of red lines from the international community. Most military operations are not *faits accomplis*. Military operations which do qualify include land grabs seizing territory, hostage-rescue raids like Israel’s 1976 raid on Entebbe, and airstrikes to destroy weapons of mass destruction sites such as Israel’s 1981 destruction of the Osirak reactor.

The two most important words in the definition are “limited and “unilateral.” “Limited” does not mean unimportant, but rather only that the *fait accompli* does not aim to conquer the deterrer outright or change the regime.²⁰ This is crucial because the strategy is intended to aim for a gain small enough that the adversary will let it go rather than escalate. “Unilateral” means that the adversary is not a participant in the change in the status quo, which is instead imposed upon that adversary to its detriment. Coercion is usually thought of as leveraging a threat to earn a concession given by the other side, but the focus here is on simply taking a concession without it being given. Coercion and the *fait accompli* are two fundamentally different ways of acquiring a gain.

*Faits accomplis* are distinct in principle from demands or signals. For example, the Austrian ultimatum to Serbia in 1914 is not a *fait accompli* because Serbia retained the option to capitulate or attempt to negotiate; Austria did not simply seize a strip of Serbian territory at the outset. Unfortunately, it is often not possible to place crisis actions into neatly defined, mutually exclusive boxes. For instance, virtually any *fait accompli* has the potential to signal something, as – for that matter – does nearly any other action taken by a state during a crisis. It is therefore necessary to evaluate the primary purpose of actions and categorize accordingly. I use the term “*fait accompli*” far more frequently than is the norm in International Relations, but I believe I do so without broadening its definition beyond existing usage. This suggests that the concept has been under-utilized, and would benefit from more consideration.

---


²⁰ This creates a definitional oddity wherein a single action may be a *fait accompli* with respect to one adversary but not another. Iraq’s 1990 invasion and annexation of Kuwait was a *fait accompli* with respect to the United States, Saudi Arabia, and other states opposed to that action. However, it was not a *fait accompli* with respect to Kuwait, because it was not limited from the Kuwaiti standpoint. *Faits accomplis* draw their effectiveness in part from being limited, as discussed in a later section, so however awkward this definitional quirk may be, it does conform to the logic of the theory.
States can execute *faits accomplis* repeatedly in one form of what Schelling (1966) termed “salami tactics,” a reference to slicing off the objective piece by piece in limited increments.\(^{21}\) I define “salami tactics” as repeated *faits accomplis*. Usually, but not necessarily, these *faits accomplis* are of a similar nature to each other. Hitler’s advances into the Rhineland, Austria, the Sudetenland, the remainder of Czechoslovakia, and Poland are the standard example of salami tactics. Germany made some of these gains by *fait accompli*, others by coercing concessions. Territorial salami tactics of this sort have been rare, but salami tactics have been far more prevalent in other issue areas with weaker red lines. For example, demands to halt human rights abuses or nuclear programs are often met with a continuing series of acts of limited unilateral noncompliance.

The dependent variable is the outcome of the crisis (or, more broadly, the coercive conflict). This variable ranges from total success to total failure. Importantly, states can suffer failure in two distinct ways: they can lose either by backing down or through the imposition of punishment, especially when that punishment leads to escalation and war. When a threat is rebuffed, the choice for the coercer is often between enduring a called bluff and starting a war. Success is simpler; it requires the adversary to back down, either by abandoning its demands or capitulating to concessions. In the real world, outcomes often fall in between these possibilities. Stalemates are possible, and there is a difference between small and large successes. It is easiest to think of the dependent variable as a spectrum from State A backing down completely to State B backing down completely, with stalemate in the middle.\(^{22}\) Also possible, but lying outside this spectrum, is the second form of failure: the imposition of punishment. This often initiates a costly war that represents a failure for both sides.

**A Typology of Red Line Vulnerabilities**

Vulnerable red lines are any of the following: arbitrary, imprecise, unverifiable, and incomplete. These vary from not at all vulnerable to fully vulnerable, with degrees of partial vulnerability both possible and common. The U.S. red line against the Soviet introduction of nuclear missiles into Cuba was partially unverifiable, for example, because a violation *might* have been detected. Any combination of vulnerabilities is possible in principle, including none and all. The standard deterrence scenario in which a state seeks only to deter invasion across a well-defined border offers a common example of a red line with none of the vulnerabilities. Importantly, red lines fail for many reasons beyond the four vulnerabilities discussed here, including the absence of a sufficient national interest to uphold them. The violation of a red line with one of these vulnerabilities therefore cannot be assumed to be a consequence of the vulnerability simply because the vulnerability exists. Below I define the four

---


red line vulnerabilities and provide a set of examples drawn from efforts by the United States to coercively halt and reverse nuclear programs in North Korea and Iran between 2002 and 2012.

An arbitrary red line is one that has no conspicuous, qualitative differences from nearby alternatives. I use the term “arbitrary” as the antonym of “focal,” rather than its alternative common usage as the antonym of “deliberate.” Non-arbitrary red lines fall on focal points. Established national borders generally provide such focal points. Geography can also often provide focal points (e.g., rivers and shorelines), as can certain norms and principles. Arbitrary red lines in the form of gray areas can arise, for instance, when political geographic focal points (established borders) and physical geographic focal points fail to overlap. The United Kingdom’s red line against an Argentinean seizure of the Falklands offers an example. Physical geography alone can also produce gray areas within focal points. Even as rivers offer focal points for strong red lines along borders, river islands often constitute gray areas that are tempting targets for faits accomplis.

An imprecise red line does not specify whether certain possible actions do or do not violate it. The degree of imprecision is the size of the resultant gray area. Imprecision is greater as more questions of the form “does it include _____” cannot be answered unequivocally. Imprecise demands are ambiguous, though imprecision is not the only way in which demands can be ambiguous. The focus here is on the precision of the demand itself, not the clarity with which it is conveyed.23

An unverifiable red line is one for which it will not be immediately known whether the red line has been violated if the adversary does in fact violate it.

An incomplete red line does not fully encapsulate that which it aims to protect. An incomplete red line could also be called a “narrow” or “flankable” red line, because it is vulnerable to a fait accompli which seizes whatever the red line is protecting without actually violating that red line. More formally, an incomplete red line exists when a state demands A to get or keep B, but the adversary can attain B while still complying with A.

United States coercive nonproliferation efforts targeting Iran and North Korea between 2002 and 2012 offer examples of all four vulnerabilities, which seem to be endemic to coercive nonproliferation. These vulnerabilities were, by and large, not the result of policy errors by Western diplomats. Setting strong red lines was inherently difficult due to the nature of the issue, as a later section will revisit.

To begin with Iran, the overarching objective was that Iran not construct nuclear weapons. In the words of United States Secretary of Defense Leon Panetta, “Our red line to Iran is, do not develop a

23 If one thinks of a radio message, the distinction is clear. A message could be received clearly but still be unclear as to its meaning. Alternatively, a highly-precise message could be garbled in transmission. It is possible that clarity of transmission and clarity of intended meaning produce the same effects, but I do not take that for granted. The former is a promising area for future research, given the role of poor communication in causing several prominent wars.
nuclear weapon. That's a red line for us."  

The problem was that this red line was unverifiable and, to a certain extent, incomplete. Once a state such as Iran accumulates all of the necessary materials and know-how to construct a nuclear weapon, constructing the bomb itself is comparatively easy to accomplish and conceal. Several nuclear powers did just that for a time. Consequently, demanding no nuclear weapon left Iran the opening to develop a large-scale nuclear program that would be capable of rapidly and (potentially) covertly constructing weapons. These problems led to the need for at least one additional red line limiting the nature of Iran's nuclear program.

Consider, then, consider the potential red line from the United States to Iran that it have no nuclear weapons program. This red line suffers from its own problems. In particular, it is imprecise because it does not specify an answer to the question: what is a nuclear weapons program? Is it the manufacturing of bombs? Is it the production of fissile material, i.e., highly enriched uranium or plutonium? This is the critical technical hurdle to overcome in making a nuclear bomb, but fissile material has civilian applications and is a legal right under the Nuclear Nonproliferation Treaty. Even within an attempt to limit fissile material production, how about a pilot centrifuge facility? Or a few experiments at spent-fuel reprocessing? These gray areas leave obvious openings for Iran to progress towards a nuclear weapon without clearly violating the red line.

Given the imprecision of demanding no nuclear weapons program, the United States found itself for years demanding instead that Iran halt and/or abandon its uranium enrichment program. Doing so meant demands such as that Iran not build another centrifuge and that Iran not operate existing centrifuges. These demands were comparatively precise. However, even assuming the United States would have known whether its demands were violated, the problem was arbitrariness. In practice, this amounted to a red line of the form, “Do not build a 5,687th centrifuge.” Why the next centrifuge, and not the one after? Could the United States credibly threaten to abandon its coercive effort and impose punishment for one centrifuge, rather than to reinstate the very same demand for the next? If Iran built just one more, the temptation will be to demand a halt again, rather than to give up and impose punishment. The sudden emphasis on that centrifuge was arbitrary. The path was paved for an interminable series of threats, protracted but inconclusive negotiations, and gradual Iranian progress towards the bomb.

The history of coercive nonproliferation against Iran is a history of vulnerable red lines exploited with incremental steps towards the ability to construct a nuclear weapon. In 2005, the West set an explicit red line against Iran converting raw uranium into uranium hexafluoride gas, a necessary step for enriching uranium. When Iran did just this in 2006, the West responded by setting a new red line proscribing the use of this gas to enrich uranium. When Iran crossed that red line as well, a series of further red lines followed. At various times the Western powers demanded a halt to Iranian uranium enrichment, that Iran not resume enrichment after temporary halts, that Iran limit its enrichment to a

pilot scale rather than industrial scale, and that Iran not enrich past low levels suitable for nuclear power reactors.\textsuperscript{27} Eventually each of these red lines was crossed.

In addition to confronting many of the same difficulties when applying coercive nonproliferation to North Korea between 2002 and 2012, the United States suffered from certain specific fears over the potential incompleteness of its red line. In order to achieve denuclearization, the United States at times focused on demanding the closure or dismantlement of the Yongbyon reactor, North Korea’s source of fissile material (plutonium). This red line was incomplete due to the potential for a rumored secret North Korean uranium enrichment program. With such a program, North Korea could (and did) flank the Yongbyon red line and continue to expand its nuclear arsenal despite closing Yongbyon.\textsuperscript{28}

The logical solution to this incomplete red line was to broaden the demand by also requiring that North Korea forgo uranium enrichment and destroy all its extant nuclear devices, which was the course taken by the United States. The problem with this broader red line, however, was verifiability. Would the United States know for sure if covert uranium enrichment continued, or if a few bombs were concealed in a bunker?\textsuperscript{29} In bargaining with the North Koreans, the United States faced a tradeoff between completeness and verifiability. It opted for a complete demand with further verification demands to redress this unverifiability, but the North Koreans rejected these measures.\textsuperscript{30} It is possible that North Korea refused to agree to any terms that would leave the United States with a stronger red line against further steps towards a nuclear arsenal.

In sum, the fingerprints of vulnerable red lines run throughout the Iranian and North Korean nuclear programs. American red lines failed to do more than force the two states to proceed in limited, deniable increments toward the bomb. Iran and North Korea were able to pursue nuclear weapons without ever having to make anything like a single, public, all-or-nothing declaration of intent around which the United States could have set its red line. This would have been the ideal, strongest-possible red line, nearly the opposite of what actually transpired. Instead, the result in both cases was that the United States saw its vulnerable red lines picked apart with a series of 	extit{fait accompli} steps toward the bomb. The typology of red line vulnerabilities provides a theoretical lens and set of concepts for a discussion of these challenges.

**Why Red Lines Matter: The Problem of Faits Accomplis**

\textsuperscript{30} For the argument that the U.S. should initially accept incompleteness to achieve an initial deal, then work towards breadth, see: Joel S. Wit, “Enhancing U.S. Engagement with North Korea,” \textit{The Washington Quarterly}, Vol. 30, No. 2 (Spring 2007), p. 64.
Why do red lines matter? Why aren’t they mere window dressing at the margins of coercive conflicts shaped and determined by the power and interests of the states involved? If exploiting a vulnerable red line is analogous to slipping through a loophole in a contract, why does it work without a government to compel the other party to abide such technicalities? In international politics, wouldn’t this adversary instead insist on renegotiating, and wouldn’t that renegotiation be determined by the same considerations of power and interest that led to the status quo ante, implying the situation would return to where it stood initially? The importance of strong red lines is far from obvious.

Red lines matter because they frequently offer the best available solution to a pervasive problem confronting states: faits accomplis. When an adversary takes a limited unilateral action such as seizing territory or progressing further towards a nuclear weapon, states confront a difficult decision regarding the nature and extent of their response. How would India respond if Pakistani soldiers were to tomorrow seize territory in Kashmir? How would NATO have responded if the Soviet Union were to have seized West Berlin? How would Japan respond if China were to unilaterally seize disputed islands in the East China Sea? And if deterring the first fait accompli fails, how do states prevent being taken apart piece by piece with salami tactics? Making credible the threat to retaliate is essential for deterrence, but it is no easy task when the stakes are limited.\(^{31}\)

More often than not, the best option available to a state confronting this problem is to rely on a strong red line set on a focal point to encapsulate many small units of value that this state cannot credibly threaten to defend individually. Knowing it cannot mount a credible defense after abandoning this red line, that point becomes one from which it cannot retreat without greater cost. Any red lines behind it would be arbitrary, and easily pushed yet further back. Because this deterrer would be in dire

\(^{31}\) I focus my discussion on the importance of red lines for deterrence because the significance of red lines is simpler for deterrence than for compellence. Nonetheless, strong red lines are important for compellence. First, many compellent threats are recast in deterrent terms. Returning to the example of Iran’s nuclear program, what would seem to be a case of compellence in which the United States and others sought for Iran to dismantle its nuclear program was nearly always presented to Iran in deterrent terms. Whatever the eventual goal, the exigent demand was usually to halt enrichment or cease installing additional centrifuges. Figure 2.1 illustrates this nicely. Second, compellent red lines matter because they can lead to bargains and thereby create the contours of a future state of affairs, including its red lines. Demanding the adversary agree to a new state of affairs in which they will need to defend a vulnerable red line against fait accompli will heighten their fears that these faits accomplis are likely to take place. States capitulating to compellent threats will always be sensitive to renewed and increased demands following their concessions, but sticking these states with a vulnerable red line can exacerbate the problem. Finally, the distinction between deterrence and compellence itself may have more to do with red lines than has been appreciated. The fundamental distinction between the two lies in who must initiate the use of force (i.e., “move last”); deterrence is thought to have an advantage because the adversary must either initiate mutually-costly hostilities or relent to the status quo. But, consider the simple case in which a deterrer aims to discourage an aggressive adversary from invading. Even if this adversary does attack across the border with shots fired by both sides, the deterrer still retains the option to capitulate and avoid most of the costs of fighting. The advantage for the deterrer is less a matter of the last-move position than is usually believed. Instead, it has more to do with the importance of focal red lines (both borders and firing on forces) being taken for granted. Strong red lines based on focal points are why it is usually assumed that violating many deterrent red lines “automatically” means war, even though it is not in fact automatic. On deterrence vs. compellence, see Schelling, *Arms and Influence*, Ch. 2.
strait if it loses this red line, it can more credibly threaten to fight to defend it. National borders are the clearest example: if an adversary is allowed to cross a border without starting a war, then barriers to further predation are lost even if the initial invasion only occupied a small slice of territory. The critical question is, as Schelling put it, “If not here, where?”

Strong red lines set on focal points offer states a way to manage the problem of faits accomplis that is generally preferable to the alternatives. War always remains an option, but it is likely to be costly out of all proportion to the value of what the adversary has just taken. A strategy of proportional retaliation is often the most appealing alternative to a strategy predicated on strong red lines, especially for cases in which a strong red line is out of reach. However, a proportional response to a fait accompli may be insufficient to recover what has been lost, and it may entail grave open-ended risks without a clear path to an acceptable outcome. In many cases, a proportional response amounts to accepting a localized conflict on the adversary’s terms, often an unpromising proposition. Moreover, this sort of reciprocity strategy is prone to escalatory spirals in which each side retaliates for the last action by the other, resulting in a costly action-reaction cycle with no clear end point short of war. Strong red lines tend to provide a more stable solution to the problem of faits accomplis.

Poorly defined colonial borders provide many examples of the consequences of weak red lines. The Soviet-Japanese border running from Mongolia to Korea during the 1930s offers an extreme case, because the demarcation of this border was exceptionally poor. For instance, one 632 kilometer segment of the border was identified by only 35 markers, of which twelve had vanished and six seemed to have been moved. These conditions largely precluded strong red lines. Tensions ran high. The frontier was beset by numerous crises and small-scale faits accomplis in the form of occupying disputed land. The border saw 152 minor border disputes between 1932 and 1934. Even where the border seemed to be clearly and non-arbitrarily defined by the Amur River, fighting broke out in 1937 over river islands that fell ambiguously within what was otherwise a clear and conspicuous boundary between the two empires. Both sides perceive a serious chance of war breaking out on short notice. Border tensions eventually resulted in the Changkufeng War of 1938 and the larger Nomonhan war of 1939. This example illustrates the consequences of weak red lines between two long-standing rivals. Tensions and fears of war may have been the inevitable result of the collision of Japanese and Russian imperial ambitions, but the unusual pattern of fait accomplis, including some that boiled over into wars, was not.

As in the Soviet-Japanese case, the absence (or violation) of strong red lines does not automatically lead to war, but rather it leads to a particularly chaotic, war-prone state of disorder in which faits accomplis are unusually common and effective. From the desire to avoid this war-prone disorder comes the enhanced credibility of threats to uphold strong red lines.

32 Schelling, Arms and Influence, p. 159.
35 Ibid., p. 93.
36 Ibid., pp. 104-105.
All of this reasoning leads to the four hypotheses below. Taken as a group, the hypotheses claim that vulnerable red lines make *faits accomplis* more effective and therefore more likely. These hypotheses address the specific questions of when *faits accomplis* are likely to occur, when *faits accomplis* are likely to be successful, and when *faits accomplis* are likely to provoke war.

**H1: Vulnerable red lines make *faits accomplis* more likely.**

The theory makes one direct prediction about crisis behavior and strategy: states will tend to identify, target, and strike vulnerabilities in red lines. The remaining hypotheses turn to the implications of the theory for crisis outcomes.

**H2: Vulnerable red lines make deterrers more likely to endure losses to adversary *faits accomplis* without retaliating.**

This hypothesis claims that deterrence more often fails when the deterrent red line contains at least one of the four vulnerabilities. These losses take the form of a *fait accompli* by the adversary to which the deterrer ultimately relents. Nazi Germany’s sudden remilitarization of the Rhineland in violation of the Treaty of Versailles offers an example of a successful *fait accompli* that targeted an arbitrary red line. Acceding to gains following an adversary *fait accompli* becomes more likely both because 1) vulnerable red lines make *faits accomplis* more effective and 2) vulnerable red lines make *faits accomplis* more common. *Faits accomplis* become more effective because the deterrer is less likely to escalate to war in response to a *fait accompli* targets a vulnerability in the deterrent red line. The third hypotheses captures this claim.\(^{37}\)

**H3: *Faits accomplis* that strike vulnerable red lines are less likely to lead to war than *faits accomplis* that strike strong red lines.**

This third hypothesis is the basis for the first and second hypotheses. Because vulnerable red lines enable challengers to get away with *faits accomplis*, these vulnerabilities make *faits accomplis* more likely to occur and more likely to produce lasting gains. Conversely, *faits accomplis* striking strong red lines are more likely to provoke an escalatory response that can lead to war. Somalia’s 1977 attempt to seize the ethnically-Somali Ogaden Province from Ethiopia offers an example of a *fait accompli* attempt which violated a strong red line and led to war. Pakistan’s 1999 Kargil Operation seizing unoccupied strategic positions on India’s side of the Line of Control (the salient focal point) offers another, albeit one that led to less intense hostilities. India and Pakistan’s nuclear tests in 1998 crossed

---

\(^{37}\) It is possible that red line vulnerabilities are better at predicting where *faits accomplis* take place (gray areas) than whether they take place. Still, there is no compelling reason why red line characteristics would affect the location of *faits accomplis* if they do not also make those *faits accomplis* more effective. Why would states design *faits accomplis* around red line vulnerabilities if they accrue no advantage in doing so? And, it stands to reason that making an action more effective makes it more likely.
a focal point and led to more punishment by the United States than prior actions that included the actual construction of nuclear bombs.

**H4: Vulnerable red lines make war more likely by making faits accomplis more likely.**

Vulnerable red lines make *faits accomplis* more likely, and not all of those additional *faits accomplis* end peacefully. Whether a *fait accompli* results in a successful gain or a war depends on whether the state employing it has successfully gauged the level of loss the other side will accept before it prefers war. In that sense, this theory is not entirely divorced from traditional structural variables like power and interests (see next chapter). Miscalculating or gambling on an uncertain estimate of this difficult-to-observe aspect of the other side’s interests can cause states to attempt *faits accomplis* in pursuit of limited gains only to end up starting wars.\(^38\) Although the connection to red lines has received relatively little attention, the claim that *faits accomplis* are perilous and fraught with the potential to cause war is generally accepted in the literature.\(^39\) Argentina’s seizure of the Falkland Islands in 1982 offers an example of a *fait accompli* that backfired and led to a failed wider war against the United Kingdom.

Although the fourth hypothesis explains why vulnerable red lines can make war more likely, the third hypothesis explains the opposite, i.e., why vulnerable red lines can make war less likely. The root of this seeming contradiction lies in the fact that vulnerable red lines connect to the probability of war through two distinct mechanisms, pushing in opposite directions through each. Vulnerable red lines make *faits accomplis* more likely. Because *faits accomplis* are a risky tactic, this in itself implies an increase the likelihood of war. However, not all *faits accomplis* are equally war-prone. Although vulnerable red lines may make *faits accomplis* more common, these vulnerabilities also make the *faits accomplis* which target them less prone to causing war. In short, *faits accomplis* become more common, but each *fait accompli* is less dangerous.

Vulnerable red lines increase the likelihood of war through one mechanism but reduce it through another. The net effect on the likelihood of war depends on which of these mechanisms predominates. I return to this point in Chapter Four, where it has significant implications for the quantitative analysis of the relationships between red lines, *faits accomplis*, and war.

Many crises feature more complex strategic interactions than these hypotheses can capture. One reason why is that something gained through a *fait accompli* must be retained. This generally requires a new deterrent red line. The new red line can have vulnerabilities of its own. For instance, a river island remains a gray area regardless of which side occupies it. In many cases, challengers seize something in a *fait accompli* only to have the original deterrent retake it in a counter *fait accompli*. Both

---

\(^{38}\) On the miscalculation of an adversary’s resolve leading to war, see Fearon, “Rationalist Explanations for War”; Geoffrey Blainey, *The Causes of War* (New York: The Free Press, 1988), Ch. 3. More on this in the next chapter.

the Berlin and Cuba cases considered in later chapters feature an initial *fait accompli* targeting a red line vulnerability (as predicted) followed by a more complex strategic interaction involving the new red lines set during the crises. As explained in the previous chapter, this type of case is optimal for qualitative theory testing because it offers such a rich array of observable implications. Nonetheless, an effort to catalog each of these possibilities would quickly lose touch with parsimony, and so the above hypotheses represent only the simplest, most direct predictions of this theory of red lines and *faits accomplis*.

These hypotheses speak to the underlying research agenda of explaining the actions taken during crises and other coercive conflicts, as well as the outcomes of those conflicts. The hypotheses offer only one part of the complete answer to those larger questions. Although the dangers of *faits accomplis* are understood in the existing literature, the determinants of when *faits accomplis* take place and when they succeed have received virtually no attention, nor has the role of red lines in answering those two questions.

**Red Line Vulnerabilities in the Existing Literature**

This section reviews the existing literature in International Relations on red lines. Special emphasis is placed on Thomas Schelling’s insights in *Arms and Influence* and *The Strategy of Conflict*, from which the theory of red lines in this chapter draws heavily. The theory in this chapter builds on Schelling’s insights, refining them into a clearly-specified theory while making certain modifications and additions as necessary. I proceed through the literature’s discussion of each vulnerability in turn.

**EXPLOITING ARBITRARY RED LINES**

Red lines set on focal points are advantageous because once they are abandoned, there is no credible barrier against limited violations. Stern et al. (1989) conclude along these lines that

---

40 A final hypothesis also follows from this logic: vulnerable red lines make crises more likely. I exclude this from the set above for the sake of avoiding an additional dependent variable, and because this hypothesis is not adequately tested in the empirical chapters. It offers an opportunity for future research.

41 Schelling, *Arms and Influence*; Schelling, *The Strategy of Conflict*. Though undeniably brilliant, Schelling’s discussion of these dynamics is not always consistent, can be vague, and often conflates the four red line characteristics with each other and with other considerations involved in coercive signaling/communication. This likely explains why so much less has been made of these passages in these books than many other sections of those classic works, despite the emphasis Schelling himself placed on them. Partly because Schelling’s discussion of these matters tends towards vagueness and inconsistency, there is no clean way to catalog the differences between the theory in this chapter and the ideas in his books. I deal with this problem simply by citing any relevant portions for every major theoretical claim I make. Readers can judge for themselves where Schelling’s contribution ends and mine begins, but the underlying fact is that this theory is built on Schelling’s insights.
“concessions encourage further demands when ... a retreat takes a state past a salient point.”

Arbitrary red lines are easily pushed to alternative arbitrary lines with targeted *faits accomplis*. Schelling (1966) specifies the most important characteristic of non-arbitrary red lines as being different from alternative possible red lines in a way that is “qualitative rather than a matter of degree.” In other words, this focal point will make for a stronger red line if it is unique and singular, rather than one of several similar qualitative distinctions. Schelling (1960, 1966) also provides the one clear elaboration of why arbitrary red lines are so dangerous, and why focal points are therefore so valuable: “a focal point for agreement often owes its focal character to the fact that small concessions would be impossible, that small encroachments would lead to more and larger ones.”

The importance of focal points, salient points, conspicuous points, firebreaks, discontinuities, saliencies, and stark distinctions (all referencing the same idea) has a history in the literature. In International Relations, most of this attention came during the Cold War with regard to the risks of escalation from conventional war to nuclear war. Kahn (1965) famously emphasized the importance of the saliency of the nuclear-use threshold in limiting escalation. He quotes Deputy Assistant Secretary of Defense Alain Enthoven, “In efforts to limit violence, there is and will remain an important distinction, a ‘firebreak’ if you like ... a recognizable, qualitative distinction that both combatants can recognize and agree upon if they want to.” Yet, can’t conventional weapons in large quantities do as much damage as a nuclear bomb? Isn’t even this distinction arbitrary? Kahn defended the emphasis on the convention-nuclear focal point as follows,

The nuclear threshold has become a purely formal, nonlogical distinction between two types of war. Therefore, before it can be dismissed as too arbitrary, we have to ask whether or not there are, logically, functionally, or otherwise, more ‘objective’ and salient threshold, at least once the nuclear border has been crossed. If we do not find such thresholds, then we must ask why it is wrong to use arbitrary distinctions when they are widely accepted.

---

44 Schelling, *The Strategy of Conflict*, p. 111. I forgo further discussion here due to the treatment of the issue in the earlier section on why red lines matter.
The importance of focal points has also been recognized in studies of international law and agreements, with particular emphasis on their role in enabling agreements by making violations less likely. Huth et al (2011) find that international legal principles and agreements provide focal points that increase rate at which territorial conflicts get resolved.\(^{48}\) Carter and Goemans (2010) find that new borders tend to form based on pre-existing administrative boundaries because these boundaries provide a focal point that enables bargains.\(^{49}\) Schelling (1966) remarks on the intriguing fact that states very often comply with treaties they have not signed because these treaties provide focal points for all states, including non-signatories.\(^{50}\) Fortna (2003, 2004) makes the case for precise agreements that also provide “a focal point that can help prevent ‘salami tactic’ attempts to push the line to either side’s advantage.”\(^{51}\) She emphasizes the utility of demilitarized zones in establishing such focal points.\(^{52}\)

There have always been doubts about whether what is seen as a focal point by one side will also be seen that way by the other, perhaps due to cultural differences or some intervening psychological factor.\(^{53}\) This view relates to a broader skepticism about signaling discussed in the next chapter. This critique would only pose an existential threat to the theory in this chapter if mutually-recognized focal points are truly rare, but there are too many examples of mutually-perceived focal points such as rivers, shorelines, and longstanding borders to sustain that view. Still, one can reasonably suppose that when a deterrer perceives it has set a red line on a focal point that the adversary does not perceive in the same way, unexpected and particularly perilous \textit{faits accomplis} may be the result.

Finally, Schelling (1960) suggests that focal points may naturally strengthen over time, particularly when they are generally respected.\(^{54}\) This may explain Shimshoni’s (1988) observation that challengers sometimes use violence to prevent the status quo from becoming accepted and recognized, i.e. to prevent the formation or solidification of a focal point.\(^{55}\) Pakistan’s history of incursions and challenges to the Line of Control in Kashmir may be indicative of this strategy to try to minimize the


\(^{50}\) Schelling, \textit{Arms and Influence}, p. 140. In some cases, these ‘compliant’ actions would have been taken regardless, but in others states take highly specific actions that only make sense with respect to an existing treaty.


\(^{55}\) Jonathan Shimshoni, \textit{Israel and Conventional Deterrence: Border Warfare from 1953 to 1970} (Ithaca: Cornell University Press, 1988), p. 218. Strategic behavior designed to affect the strength of other states’ red lines is not the focus of this project even as it is consistent with it, and therefore offers an interesting area for further research.
focal-ness of a focal point which works in the adversary's favor. Strong Palestinian opposition to the Israeli security wall with segments running through the pre-1967 West Bank fits this mold, being motivated in part that this wall will become a focal point that serves as the basis for a future border. Egypt's motivation for the 1967-1970 War of Attrition with Israel has also been attributed to the desire to prevent the Suez Canal (and the adjacent Bar Lev Line) from becoming the presumed basis for an eventual border. The phenomena of aggressive actions designed to prevent the formation or solidification of focal points provides an interesting qualification to the broader claims in this chapter that strong focal points aid in deterring aggressive actions. There is no contradiction here. The logic mirrors that of the relationship between relative power and deterrence; greater relative power tends to deter attack, but favorable relative power shifts provide a window of opportunity for the adversary to strike that tends to promote an attack.

EXPLOITING IMPRECISE RED LINES

Imprecise red lines are vulnerable to faits accomplis in their gray areas. The temptation for the deterrer is to sacrifice the gray area in order to avoid an even more costly war. The hope is that the gray area can be lost without unduly calling into question the credibility of the deterrent threats to fight in response to future faits accomplis that unambiguously cross red lines.

According to Schelling (1966), imprecise red lines are quite prevalent: “most commitments are ultimately ambiguous in detail, ... [often] because of the plain impossibility of defining them in exact detail.” Ambiguity in detail is a common problem even for red lines set on undisputed focal points. Smoke (1978) observes, “Decision-makers may employ a gradualist strategy to try to smudge, or avoid as much as possible, the discreteness and obviousness of saliencies they are crossing.”

A contemporary example arises in the deterrence of a Chinese invasion of Taiwan, to which the United States has an ambiguous security commitment. It is particularly unclear whether or not the U.S. deterrent red line protecting Taiwan from Chinese attack extends to the Taiwanese-held islands of Jinmen (Quemoy) and/or Mazu (Matsu) from invasion. The United States would presumably be reluctant to fight over small islands within such close proximity to mainland China, given the military disadvantages of doing so and their lack of intrinsic worth. However, failing to fight might erode the

56 Raghavan, Siachen, p. 42.
60 Schelling, Arms and Influence, p. 67. Schelling further remarks, “There is always some threshold below which the commitment is just not operative, and even that threshold itself is usually unclear.” However inviolable a border is writ large, one soldier straying across inadvertently generally will not trigger large-scale retaliation.
61 Smoke, War: Controlling Escalation, p. 275.
credibility of the larger red line around Taiwan. It would also cede the islands. This could undermine Taiwan’s faith in U.S. security commitments and facilitate future Chinese coercion. Should tensions escalate in the future, a Chinese *fait accompli* targeting these islands is a problem that should be taken seriously.

But can’t “strategic ambiguity” be beneficial? A considerable amount has been written about both the virtues of ambiguity and the virtues of clarity in coercion, bargaining, and negotiation. There is no focused study that attempts to compile (let alone evaluate) these competing claims in the International Relations context. Nor is it clear that these claims are mutually exclusive, as they may each apply in differing circumstances. Imprecise red lines are only one type of ambiguity in coercive bargaining. Several of the main arguments in favor of ambiguity do not extend to setting imprecise red lines. Moreover, the theory in this chapter does not necessarily capture the net effect of ambiguity across all the ways it may influence outcomes. Instead, the theory explains why one important consequence of imprecise red lines is to enable adversary *fait accomplis* and thereby bring about undesirable outcomes for the states setting them. Below I briefly review the main claims the literature makes for and against imprecise demands.

The primary argument against ambiguity is inextricably linked to the primary argument for it. By providing states with the flexibility to respond to a violation by denying that a real red line was in fact crossed, ambiguity allows them an escape route to avoid enforcing the threat. This is advantageous in that it avoids entrapment into unwanted wars, but disadvantageous in that deterrence is less credible when backing down is less costly. Schelling (1960) makes the case as follows:

> It is essential … to leave as little room as possible for judgment or discretion in carrying out the threat. If one is committed to punish a certain type of behavior when it reaches certain limits, but the limits are not carefully and objectively defined, the party threatened will realize that when the time comes to decide whether the threat must be enforced or not, his interest and that of the threatening party coincide in an attempt to avoid the mutually unpleasant consequences. \(^{62}\)

According to Lebow (1977), “Flexible commitments, which appear to limit the would-be deterrer’s cost of disengagement, are hardly likely to be interpreted as impressive indications of resolve.”\(^{63}\) Snyder and Diesing (1977) argue along these lines that maximum explicitness and clarity produce maximum credibility, because explicitness engages the national reputation while imprecision allows states to evade some of the ignominy of backing down. However, they also find that states tend to choose this flexibility to avoid entrapment over the added credibility of specificity.\(^{64}\) Sagan (2000)

---

\(^{63}\) Lebow, *Between Peace and War*, p. 85.
\(^{64}\) Snyder and Diesing, *Conflict among Nations*, pp. 216-217. Note that it can be useful to draw an analytical distinction between explicitness and precision. How a demand is conveyed is distinct from its content.
articulates this same tradeoff clearly in the context of threats to retaliate with nuclear weapons against the use of chemical or biological weapons.  

These studies are far from alone in praising the benefits of clarity in making demands. Quinlan (2004) adds a useful qualifier: clarity about the demand component of threats may be more advantageous than clarity about the punishment component. For instance, telegraphing the exact form a punitive strike will take allows the adversary to take countermeasures to defeat it, or at least to better absorb the blow. Overall, the case for clarity is fairly consistent in placing its primary emphasis on the way that imprecision undermines credibility by providing a way to slip out of upholding red lines after violations.

The case for imprecision consists of several distinct lines of reasoning. Perhaps its greatest advantage is part and parcel of its main disadvantage; added flexibility reduces credibility but also eases the path for states to grant war-avoiding concessions without the same reputational and domestic political costs that might otherwise apply. Fisher (1971) argues that adversaries might more easily acquiesce to ambiguous demands because there is less political risk in doing so, i.e. less damage to their reputation and prestige. Ambiguity over one point of contention may also enable states to sidestep that issue and reach agreement on another matter. Particularly in disputes over symbolic issues, ambiguity may even enable both sides to claim victory.  

Another argument for imprecision is that it lets adversaries comply with the broader intent of the demand in whatever form is best for them, thereby reducing the costs of compliance. Fisher (1971) regards the most important drawback of specificity as that it puts a ceiling on what can be

---


68 Roger Fisher, *Basic Negotiating Strategy: International Conflict for Beginners* (London: Allen Lane, 1971), p. 75. However, I would argue that the attributability of an action to a demand is not solely or primarily a matter of that demand’s level of precision.


obtained. Chayes and Chayes (1990) go further, arguing that “increased detail and specificity, far from facilitating the resolution of compliance questions, may become a breeding ground for disputes.” However, although these arguments about negotiation are compelling in general, they may be less applicable to the intensely-competitive conflicts that this study examines. Many times, the only common interest the two parties share is avoiding the costs of war. Beyond that, the situation is zero-sum, e.g., when both sides want the same territory. If so, then any latitude left for the other party to unilaterally shape the final terms of the agreement can be assumed to mean they will shift the bargaining in their favor as far as possible, i.e. directly against the interests of the state with the imprecise demand. Similarly, in particularly confrontational settings, it is difficult to see how specificity breeds unnecessary disputes. Leaving these ambiguities in place only avoids disputes to the extent that the other side is freely permitted to take what it wants within the gray area.

**EXPLOITING UNVERIFIABLE RED LINES**

Unverifiable red lines are vulnerable to covert violation. The adversary can secretly violate the red line, and upon eventual discovery present its violation as a fait accompli. A classic example of a red line with verifiability problems (though not total unverifiability) is the Soviet Union’s placement of nuclear-tipped ballistic missiles in Cuba. In this case, the fait accompli was not implemented with full success; its discovery triggered the Cuban Missile Crisis.

That unverifiability makes violations more likely and prevents successful bargains (out of fear of covert violations) has long been believed. According to Smoke (1977), many escalatory actions are taken in the expectation that they can be carried out secretly or invisibly. Schultz (2010) shows that openings for unilateral and covert violations can lead to bargaining failure and conflict. Fortna (2003) finds that verification mechanisms lead to more durable peace agreements. Verification concerns have also shaped arms control negotiations. For example, there is a strong preference for and tendency towards agreements which ban a specific weapon system outright rather than agreements limiting the levels of a weapon that states can possess. The reason: the violation of an outright ban can be detected by spotting even one of the forbidden weapons, but determining how many of a weapon system a state possesses requires far more effort and information.

---

72 Ibid., p. 75.
EXPLOITING INCOMPLETE RED LINES

The vulnerability of incomplete red lines is analogous to that of deploying troops on too narrow a front: they can be flanked. Flanking seizes the object of the demand without violating the red line directly. Unlike the other red line characteristics, incompleteness has received scant attention in the International Relations literature. One exception is the conclusion in Stern et al (1989) that deterrent threats tend to fail when the adversary “believes it can ‘design around’ the defender’s commitment” (though how this is done is never explained). In part, the lack of attention to incompleteness may be because it is not as independent from the other red line vulnerabilities as it may seem. When the primary red line is vulnerable to flanking, more often than not that actor will also set an auxiliary red line to discourage that flanking. It is, then, a matter of whether this auxiliary red line is vulnerable on the other dimensions, usually arbitrariness. Nonetheless, treating incompleteness as a distinct characteristic is valuable. It is by far the simplest way of addressing the existence of multiple red lines in a single situation, and it calls attention to the important phenomenon of flanking red lines.

As I will explore in more detail, the 1948-1949 Berlin Crisis revolved around flanking incomplete red lines. First, the Soviet Union attempted to gain an advantage over the status quo ante by shutting off land supply routes into the city without engaging in hostilities. This was a flanking strategy, exploiting the fact that the U.S. red line protecting West Berlin from direct attack did not extend to the supply corridors. Provided that the United States was unwilling to allow the population of West Berlin to starve, the blockade was meant to leave the United States in the difficult position of having to choose between using force to break open land supply routes or backing down and handing West Berlin over to Soviet control.

However, the strength of the Soviet Union’s red line against resupplying Berlin relied on the focal point of attacking their forces blocking the transit corridors, and this focal point did not extend to the skies. Consequently, the United States vertically flanked this red line with the Berlin Airlift. In the air, it was the Soviets that would need to assail Western planes to stop resupply. With the unexpected demonstration of the viability of indefinite aerial supply of West Berlin, the Soviets’ effort to leave the United States with a choice between capitulation and war failed. The story of the Berlin Blockade Crisis centers on flanking incomplete red lines, as I will explore in more detail in Chapter Five.

Scope Conditions

Two scope conditions must be met in order to sustain the logic and predictions of the theory in this chapter. Both scope conditions are uniquely common in international security affairs, but both are met in certain other contexts.

---

First, both sides – but particularly the deterrer – must see war as costly. If the deterrer sees a potential war as likely to be short and cheap, little can be gained by a *fait accompli* which forces it to choose between war and backing down. One could argue that this scope condition was not met prior to World War One due to the false optimism with which so many states perceived their military prospects. Consequently, the hypotheses may have less explanatory power for the 1914 July Crisis than for others. In contrast, the expectation is significantly greater explanatory power in the crises preceding World War II and maximal explanatory power in Cold War coercive bargaining due to nuclear weapons raising the potential costs of a major war to extraordinary heights.

Second, bargaining that takes place within and is channeled by institutional rules might be expected to exhibit different dynamics. As a result, this study applies to institution-less bargaining, bargaining under anarchy. It is not necessary that no institutions exist at all, but rather that existing institutions do not directly channel the particular bargaining interaction.

Why Are Red Lines Ever Vulnerable?

Why are red lines ever arbitrary, imprecise, incomplete, or unverifiable? Why would states, as strategic actors, ever make a suboptimal demand? Is the culprit some combination of ignorance and inaptitude, or do states face constraints that limit the quality of the red lines that they can set? Because these questions seek to explain the independent variables of this study, they are, strictly speaking, beyond the scope of the project. Nonetheless, their importance is such that I provide a brief discussion here. I also commend these questions as a worthy basis for future research.

A first broad generalization from the cases and examples I have surveyed is that vulnerable red lines usually result from binding constraints, not sheer carelessness. Despite being able to choose from many potential red lines, states often cannot escape one or even several red line vulnerabilities while still demanding what they wish to demand. The nature of these constraints is consistent. States are rarely or never locked in to a particular red line, at least in the minimal sense that they can change to (inferior) alternatives if they choose to do so. Instead, states are often locked out of setting red lines that are strong on one (or more) dimension.

---


81 The empirical analysis of territorial crises does not quite align with this expectation. The theory appears to fare better after 1980, but there is no clear difference between nuclear and non-nuclear cases.

82 Note that the hypothesized effects of vulnerable red lines are the same regardless of whether the vulnerable exists due to carelessness or constraint. One would, however, be inclined to expect that such carelessness would also affect the outcome through other channels.

83 For instance, compellent threats may have an inherent arbitrariness problem because they must set deadlines. Deadlines can be made quite precise, but it is rare that a single focal point exists to explain why the deadline is one week rather than two, twelve hours rather than twenty four. For more on compellent threats and deadlines, see Schelling, *Arms and Influence*, p. 72.
Although these constraints are numerous and diverse, depending greatly on the intricacies of particular situations, they appear to vary most consistently by issue type. First, arbitrary red lines are more common for types of issues in which focal points are rare. The absence of a single viable focal point on which to set a red line is largely beyond control, and can impose the burden of an arbitrary demand. Conversely, when there are only one or a few focal points present, their importance may constrain states to select one of those few possibilities even when a red line on that focal point will suffer from some of the other vulnerabilities. Second, imprecision is more common for issues which cannot be quantified or measured easily. Consider, for example, a demand to tamp down nationalist saber-rattling in the media. Third, the incompleteness of a red line often depends on the number of ways in which a goal can be accomplished. Fears of a secret North Korean uranium enrichment program and undetected Soviet underground nuclear tests fit the bill. Plutonium production is not the only path to the bomb, and atmospheric testing is not the only way to test nuclear weapons. Fourth, the ease with which certain actions can be observed affects verifiability, as those same two examples also illustrate. Schelling (1966) offers the demands conditioning foreign aid as an example of an issue type afflicted by frequent red line vulnerabilities: “Whenever a recipient of foreign aid … is told that it must eliminate domestic corruption, improve its balance of payments, or raise the quality of its civil service, the results tend to be uncertain, protracted, and hard to attribute.” Red lines against nuclear proliferation and human rights violations seem to be far more prone to vulnerabilities than red lines against territorial conquest.

Constraints obstructing strong red lines often take the form of difficult tradeoffs among the vulnerabilities. Although ‘invulnerable’ red lines are feasible in some situations, in many others invulnerability on one dimension precludes it on another. In setting red lines against North Korea’s nuclear program, the United States confronted an incompleteness problem if it did not set an additional red line against a uranium enrichment program, but faced a verifiability problem if it did set that red line. In setting red lines against Iran’s nuclear program, the United States confronted an imprecision problem if it simply demanded the end to its nuclear weapons program, but faced an arbitrariness problem if it attempted to make that red line more specific, such as by demanding that not one additional centrifuge be installed (see earlier in the chapter for a more in-depth discussion of these examples). In the most direct discussion of these tradeoffs in the literature, Schelling (1960) argues,

> In order to make a threat precise ... it may be necessary to introduce some arbitrary elements. The threat must involve overt acts rather than intentions; it must be attached to the visible deeds, not invisible ones; it may have to attach itself to certain ancillary actions that are of no consequence in themselves to the threatening party.

Such problems can be found in a wide variety of issue areas. Red lines meant to deter human rights violations could be made specific to a village or even an individual, but only at the price of becoming arbitrary and, consequently, detached from the larger issue of human rights in that state writ.

---

85 Schelling, *The Strategy of Conflict*, p. 40
large.\textsuperscript{86} If something does befall that specific village, why not set a new red line for the next village rather than accept failure and begin inflicting mutually-costly punishment? And by focusing on one village, what about others where information is scarce, or that simply have not been emphasized to the same degree? Imprecision is a problem, but making a red line more specific can create problems of its own. Although I observe tradeoffs among red line vulnerabilities to be relatively common, I offer no theory which attempts to predict when and explain why certain tradeoffs will exist for certain types of issues. Future research might explore this possibility.

Some of the constraints that lead to vulnerable red lines – and especially imprecise red lines – are political in origin, including domestic and bureaucratic politics. For instance, imprecise demands might enable broader domestic political coalitions. According to Lebow (1977), “Ambiguous commitments often represent compromises between competing and contradictory political demands.”\textsuperscript{87} Massie (2004) describes German demands immediately prior to World War I as follows, “Specifics were precisely what the German government did not wish to reveal, because such revelations would arouse in Germany a bitter struggle over what German war aims actually were.”\textsuperscript{88} Alternatively, politicians who are evaluated relative to stated demands may prefer to keep demands imprecise so that they can call more outcomes victories. Trager and Vavrek (2011) conclude that vague threats incur smaller audience costs if the leader who made them eventually backs down than do specific threats.\textsuperscript{89} Bureaucracies may not be adept at generating precise positions on seemingly secondary issue areas. In the Berlin Blockade Crisis, bureaucratic infighting between the War Department and State Department during the Second World War prevented the United States from securing a written legal obligation from the Soviet Union to permit free transportation and supply of West Berlin.\textsuperscript{90}

Diplomatic considerations can also constrain red lines. For example, Stein (1991) concludes that the United States was prevented from setting a precise red line to deter even limited Iraqi seizures of disputed Kuwaiti-held territory by the intricacies of Middle Eastern diplomacy and politics. Other Arab states were only likely to support a U.S. intervention on Kuwait’s behalf (or the public threat of one) if it became clear that Iraq was seeking the conquest of Kuwait as a whole rather than the seizure of oil-rich disputed areas along the Iraq-Kuwait border.\textsuperscript{91} This precluded a precise U.S. red line protecting these territories. The United States adopted an ambiguous position on Quemoy and Matsu during the 1958 crisis partly in order to avoid criticism from congress and third-party allies.\textsuperscript{92} Britain’s oft-maligned

\textsuperscript{86} This often happens when media attention focuses on the plight of an individual activist imprisoned by an authoritarian regime.
\textsuperscript{87} Lebow, Between Peace and War, p. 85.
\textsuperscript{92} Snyder and Diesing, Conflict among Nations, p. 223.
imprecise commitment to defend France in the event of a war with Germany in 1914 can be attributed
to its nearly contradictory goals of deterring Germany without emboldening France.  

States may deliberately choose to set vulnerable red lines when they could avoid doing so for at
least two distinct reasons.  First, a state which desires war might seek to entice an adversary into a *fait accompli* that would serve as a *casus belli* and thereby aid in legitimizing the use of force.  Such cases fall outside the scope conditions for this theory, because war is perceived as desirable rather than costly, but are worth considering nonetheless.  Second, states might feel compelled by some domestic or
foreign audience to set a red line over an issue that is not truly of real importance to them.  In such
cases, a loophole might seem to provide a way to appease that audience without risking war over
someone else’s interests.  This strategy comes with a powerful drawback; states generally suffer
embarrassment when another state exploits a vulnerability in their red line, thereby succeeding with a
*fait accompli* in spite of the state making the threat.  

Finally, one important explanation for weak red lines that recurs in the historical record is long-
term unintended consequences.  The Treaty of Versailles and the terms by the agreements that ended
the Second World War locked in place numerous red line vulnerabilities that gave rise to decades of
tension.  The Berlin crises emerged due to wartime decisions which became extremely difficult to
change by 1948 but that might have been revised with only modest difficulty prior to the end of the war.  
Looking back on the Siachen Conflict, Indian General V. R. Raghavan commented, “The negotiators [of
the Simla agreement] had just been through a war.  They could not have imagined that another war
would be fought to interpret their words.”  The rarity of vulnerable red lines due to sheer carelessness
does not extend to decisions made in the past that affect red lines years in the future.  As a result, this
study should have important implications for the formulation of international agreements, especially
what will seem like minor details in those agreements which hold the potential to lie at the epicenter of
future crises. 

**Conclusion**

This theory of red lines and *faits accomplis* provides a framework for understanding crises and
other coercive conflicts that bears surprisingly little resemblance to the conventional wisdom.  This is a
view of crises as contests in which states seek out and identify weaknesses in deterrent red lines in
order to outmaneuver adversaries by exploiting those openings for *faits accomplis*.  The next chapter
summarizes the conventional wisdom, with its focus on crises as contests in competitive risk-taking with

---

94 My suspicion is that this drawback generally outweighs the benefits of this strategy, because trying and failing
due to what often looks like ineptitude is likely more damaging to a nation’s prestige than not trying at all.
However, one can imagine circumstances in which that would not be the case.  For instance, politicians facing
elections or term limits might make a demand with a vulnerable red line to appear strong in a critical time period,
knowing that the adversary’s *fait accompli* will likely come after they have been re-elected or left office.  This
hypothesis offers one potential explanation for the weakness of U.S. nonproliferation red lines.  
a primary emphasis on signaling resolve. My claim is not that the conventional wisdom is wrong, but rather merely that it is incomplete. The empirical chapters will evaluate both the theory of red lines in this chapter and the conventional wisdom to start to examine how much each can explain.

Although this project focuses on interstate coercive bargaining, this theory of red lines and faits accomplis applies to a diverse range of issue areas. For example, it would seem to shed light on the question of when and why countries violate treaties, which, much like threats, generally set red lines that distinguish compliance and noncompliance. Some of the strongest proponents of the constraining power of international treaties argue that countries rarely violate treaties, and that most violations occur due to ambiguities in the treaties’ terms, i.e., due to inadvertent misunderstandings of obligations. Therefore, these violations should not be seen as violations at all. The two leading proponents of this viewpoint also criticize the determination on the part of the United States to set “bright lines” in arms control treaties with the Soviet Union during the Cold War on the grounds that they created unnecessary tensions and disputes. The theory in this chapter suggests a starkly different conclusion: far from being accidental, these “misunderstandings” were likely deliberate strategic attempts to get away with beneficial treaty violations where possible while minimizing the prospects for punishment.

Even the general pattern of tension during the Cold War can plausibly be explained with this theory of red lines. To oversimplify, Cold war tensions peaked during the period 1948-1962, a span which saw a series of crises in the strategically vital regions of Central Europe and (to a lesser extent) Northeast Asia. Why did these tensions ease? It is striking that this same period saw the red lines in those two regions improve as both sides expectations converged around certain focal points like the Demilitarized Zone in Korea and the solidification of the initially-temporary zones of occupation in Germany. Perhaps as a result, two things happened. First, the level of tension and the apparent probability of war declined. The situation evolved from one crisis following another to an uneasy détente. Second, the location of these tensions shifted, with an increasing share of falling in more peripheral regions where clear red lines had not yet been established. There is a compelling story to be told about how increasingly strong red lines shaped and curbed Cold War tensions. In the words of Alexander George and Richard Smoke (1974), “It can be said on behalf of American policy of containment and its handmaiden, deterrence, that they did at least establish some firm lines in the postwar world, when so much was fluid.”

---


97 Chayes and Chayes, “From Law Enforcement to Dispute Settlement,” p. 155

98 George and Smoke, Deterrence in American Foreign Policy, p. 4.
Ch.3: The Conventional Wisdom: Crisis Strategy as Signaling

The conventional wisdom on interstate coercive bargaining follows the intuition that outcomes are driven first and foremost by power and interests. Success belongs to the stronger side, or the side which wants it more. Because power and interests are perceived only with considerable uncertainty, the actions taken during crises are meant to signal high levels of power and interest to the adversary. Signaling resolve is, therefore, the cornerstone of how the conventional wisdom explains crisis behavior and crisis strategy. This view contrasts sharply with that developed in the previous chapter, which attempts to understand crises as efforts to outmaneuver adversaries using *fait accomplis* that exploit vulnerabilities in red lines.

At one level, the level of crisis behavior and tactics, the question concerns the relative importance of signaling vs. *faits accomplis*. At the strategic level, the question is: how do states make gains in crises? The conventional wisdom emphasizes compellent threats. If a state wants something from its adversaries, the strategy works in the steps: 1) compellent threat, 2) signals of resolve, 3) receive the concession from the adversary if and when the threat is credible. In contrast, this project explores a different answer to this basic question of how states make gains in crisis: getting away with *faits accomplis* without provoking war. At this level, the difference is between pressuring the adversary to grant a concession vs. unilaterally taking it.

This chapter serves two main purposes. First, it demonstrates that a clear conventional wisdom governs how the field understands crises and coercive bargaining more generally. This conventional wisdom is unrelated to red lines and *faits accomplis*. The use of the term “conventional wisdom” is not intended to be pejorative. There is a great deal of conceptual and empirical merit to this conventional wisdom; I merely argue that it is incomplete. In the interest of distilling the literature to a coherent alternative perspective on coercion, this chapter focuses on structural and strategic determinants of crisis outcomes. Other important classes of explanatory variables such as bureaucratic politics, organizational processes, and leadership will be considered when relevant in the case studies, but are not reviewed here.

Second, this chapter develops and reviews the most important alternative explanations for crisis actions and outcomes in order to lay the groundwork for the empirical chapters to follow. Although much of the discussion consists of criticisms of existing scholarship, the purpose is always to extract the strongest possible set of alternative explanations from the conventional wisdom rather than test alternative formulations of these hypotheses that are little more than straw men. I reject no major theories or variables on purely theoretical grounds, but instead I try to distinguish stronger formulations of each from weaker substitutes that are all too easily falsified.
The chapter is divided into two parts. The first reviews structural explanations for crisis outcomes, starting with the two most fundamental: power and interest. The second reviews the literature on how states signal resolve in crises. These signals are the centerpiece of how literature understands states’ behavior during crises and their impact on outcomes.

Most importantly, this chapter combines with the last to lay the groundwork for a simple question: Are crisis actions primarily intended to send signals or make gains via *faits accomplis*? The answer I reach over the rest of the dissertation is that both are important, but signaling resolve is more difficult than is generally appreciated, and red lines are considerably more important than has been recognized.

**Structural Explanations for Crisis Outcomes**

This section summarizes the most important structural explanations for crisis outcomes that lie at the center of the conventional wisdom. The term “structure” is used here to refer to the structure of the interaction, not the structure of the international system. In particular, I consider relative power, relative interest, relative resolve, the costliness of war, alignment incentives, and multiple coercers acting as a coalition. I then discuss the limitations of these explanations before proceeding to explanations that give explanatory weight to the actions taken during crises and other coercive conflicts.

**RELATIVE POWER AND RELATIVE INTEREST**

Which side wants it more and which is more able to take it if push comes to shove are the two classic structural determinants of bargaining advantage and crisis outcomes. This distinction between will and capability provides the two core variables emphasized by the conventional wisdom, relative power and relative interest in the stakes. For instance, Betts (1979) provides a clear statement of this

---

1 Relative power may be somewhat more latent and implicit in rational choice bargaining theories, but it is inevitably a primary component of explicit variables such as the probability of victory in war.

perspective by explicitly distinguishing and testing a “balance of interests” explanation for nuclear coercion effectiveness against a “balance of power” explanation.³

The baseline hypothesis: possessing greater relative power elevates the probability of victory in coercive bargaining. Because the more powerful side of the dispute would tend to fare better should bargaining fail, it can more credibly threaten to stand firm, and the adversary has more to lose from refusing to back down. Although power is more concrete and observable than interest, as an explanatory variable it is afflicted by areas of ambiguity.⁴ One is the question of whether and how allies count towards relative power in crises with two sides but more than two states (more on this below).⁵ Another is relative power may be less significant in the nuclear age, because additional weapons may simultaneously offer little defense and little ability to do more than inflict redundant destruction on the adversary.⁶ And how to weigh the relative strength of nuclear arsenals against relative strength in conventional forces?

Perhaps the most pertinent ambiguity lies in the distinction between local relative power and overall relative power. For example, in the Cuban Missile Crisis, does the overall nuclear balance take precedence, or does the extreme conventional advantage held by the United States in the Caribbean region matter more? Does it matter at all? Similarly, did the conventional balance of power in Central Europe matter during the Cold War, or was its importance superseded by the total balance including all forces regardless of location? More generally, does local relative power matter, and should it be a separate variable? This issue is sometimes addressed with the related term “escalation dominance,” which a state possesses when it has the advantage at the various possible levels of escalation (i.e. both in a limited/localized conflict and a full-scale war).⁷ In discussions of relative power in the case studies to follow, it will prove necessary to consider these distinct ways of assessing relative power: monadic vs. coalitional, nuclear vs. conventional, local vs. total.

Greater relative interest is also thought to confer important advantages in coercion. What are interests? “Interests” to realists are essentially the same as “preferences” to rational choice theorists. Both determine how much positive or negative utility an actor receives from each possible outcome. Interests give rise to goals and objectives, which in turn are pursued with strategies.

³ Betts, Nuclear Blackmail and Nuclear Balance, pp. 14-16.
Is it advantageous to want to prevail more, or less? The general view, embodied is that an actor with greater interest will be more willing to absorb punishment should bargaining fail, yielding an edge. But, in some contexts, wanting something more can be a disadvantage that the other side can exploit to wrangle a better deal. This seeming paradox is readily resolved: wanting to win the stakes more is advantageous, whereas wanting to reach a bargain more (i.e., wanting to avoid war) is disadvantageous.\(^8\)

**RESOLVE**

Resolve is routinely placed in the same category as relative power and relative interest, but its meaning and proper usage has been murky and inconsistent. Given the central role of resolve in the conventional wisdom, some resolution is required. With that in mind, I argue here that resolve is best used as a purely conceptual device rather than as an explanatory variable, and that treating it as an explanatory variable often leads to measuring it tautologically. I define resolve as the “inclination to stand firm” in a conflict, which I believe best captures its most common usage.\(^9\)

Resolve has posed a definitional muddle for the study of coercion for decades. The crux of the problem with resolve lies in the lack of a clear definition of resolve. A country with a strong leader may be more resolute, as may a country with an authoritarian (or democratic) political system. A country with stronger interests at stake may be more resolute, as might a country with a more established reputation to uphold. Even an advantage in relative power might lead a state to take a firmer line, increasing its resolve by some uses of the term. The list could continue indefinitely, because resolve is a function of every source of motivation, including every source of advantage; any advantage can lead states to take a tougher line on an issue.\(^10\)

Resolve is commonly used in three ways: 1) as an all-inclusive catchall for sources of advantage and motivation, 2) as a residual umbrella created by excluding only relative power from this all-inclusive catchall, and 3) as a synonym for interest. All three are common, and all three are problematic, but the second and third are more so. The everything-but-power conceptualization (#2) contains a contradiction, because there are compelling reasons to believe that capabilities are a source of resolve. A more powerful state might be more resolute because it knows it will have the advantage should bargaining fail and war begin.\(^11\) This definition tries to separate the inseparable. Alternatively, if resolve is defined narrowly to be the same as interest (#3), why use the term at all?

---

9. This definition is quite similar to those of others, e.g., resolve as “the degree of motivation to stand firm.” Snyder and Diesing, *Conflict among Nations*, p. 190. This definition is best understood in terms of intentions.
10. The term “bargaining power” suffers from a similar problem. Consequently, I avoid using it.
11. Alternatively, some have suggested that weak states tend to have greater resolve because only their survival is threatened. Andrew Mack, “Why Big Nations Lose Small Wars: The Politics of Asymmetric Conflict,” *World Politics*, Vol. 27, No. 2 (Jan., 1975), pp. 175-200; Ivan Arreguin-Toft, “How the Weak Win Wars: A Theory of Asymmetric Conflict,” *International Security*, Vol. 26, No. 1 (Summer 2001), pp. 93-128. This inductive conclusion may be due to a selection effect, if only highly resolute weaker states are likely to enter crises with stronger adversaries.
The use of resolve as a catchall for all sources of motivation (or all but power) in a coercive conflict results in an insoluble problem of measurement. Resolve is not directly observable. Consequently, it must be inferred from its constituent elements or its effects. The former requires an aggregating theory which does not exist, and the latter risks tautology. This aggregating theory would answer the question: in any particular situation, how heavily should each constituent element of a state’s resolve be weighed relative to the others? Is strong leadership more or less important than geographic proximity or the existence of a public commitment? By how much? The lack of an adequate aggregating theory pushes the researcher to infer resolve from its effects on behavior. Too often this means that resolve is measured by looking at the strength of the actions taken, or even the outcome of the conflict. This procedure is tautological. Resolve explains victory, but resolve is measured from the firm actions which produced the victory.

Rather than conceive resolve as an independent variable perennially mired in opacity, it is better understood as a conceptual device serving an aggregating role. A broad array of independent variables collectively determine a state’s resolve. Resolve offers a terminological shorthand to refer to a state’s inclination to stand firm as determined by a wide array of factors such as relative interests, relative power, leadership, and reputation. It is not useful analytically to treat resolve as an explanatory variable, but one can still speak of trying to signal resolve to the adversary.

THE COSTLINESS OF WAR

Though it is too rarely included alongside relative power and relative interests, a variable of comparable importance is the costliness of bargaining failure, i.e. how well the participants fare when one opts to exit bargaining and reach a non-cooperative conclusion to the dispute. In security contexts, these costs are generally the costs to both sides of implementing the threatened punishment, which often results in war. The costliness of war is among the most common variables in rational choice models of bargaining and war, but it would seem to be mentioned considerably less often in the realist literature on coercion. There may be an element of truth to this difference in emphasis, but for the most part the difference is illusory. The academic literature on deterrence, for example, centers on the costliness of war. It merely refers to it as mutually assured destruction, or by way of its main determinant post-1945, the presence or absence of nuclear weapons. Second strike capabilities serve ensure that nuclear war will in fact be costly for the adversary. For another example, Mearsheimer’s (1983) theory of conventional deterrence argues that deterrence fails when the aggressor can hope to prevail with a low-cost blitzkrieg strategy rather than a costly attrition strategy. The costliness of war can and often does vary independently of relative power and relative interests. For instance, two states

---

12 E.g., Snyder and Diesing, Conflict among Nations, p. 450.
with symmetrical power and interests would likely expect far greater costs from a war if both are nuclear powers than if both are not.  

The costliness of war matters in two distinct ways. First, increasing costs of war tend to promote bargains and discourage bargaining failure (war). Reacting to the advent of nuclear weapons, Bernard Brodie famously wrote, “Thus far the chief purpose of our military establishment has been to win wars. From now on its chief purpose must be to avert them. It can have almost no other useful purpose.” Second, the side of a dispute with lower costs of war than its adversary often has an advantage. This second point is partially, but not wholly, a function of relative power. Because the side with lower expected costs from war can more readily threaten to go to war, its threats may be more credible. The limits of this advantage have been the subject of considerable debate in the nuclear strategy literature, because it speaks to the question of whether nuclear superiority matters when both sides already have the capacity to utterly devastate each other.

ALLIES

Although crises are often break out between longstanding rivals, standoffs sometimes develop between states with closer relations, even allies. In some of these situations, one or both states may have an interest in an alignment with the other that exceeds the interest in prevailing in the dispute. When this happens, war becomes less likely, and that state is more likely to back down. For example, in the Fashoda Crisis of 1898, Britain and France contested control of territory in Sudan. Despite gaining an initial positioning advantage when its forces arrived first in the disputed region, France ultimately relented in part because of its need for England’s support in the event of a war with Germany. Similarly, Britain adopted a relatively tolerant stance towards Italy’s 1939 invasion of Albania and recognized its conquest of Abyssinia because it sought to entice Italy to align with it rather than Germany. Snyder and Diesing (1977) argue that alignment incentives of this nature are likely to be

---


more powerful in multipolar systems, because in bipolar systems the overall balance cannot be changed through alignments.\textsuperscript{20}

**COERCION BY COALITION**

For simplicity, most theories of coercion are framed in terms of a single coercer and a single target or two comparable states coercing and bargaining with each other. Although crises almost always feature only two sides, one or both of the sides often consists of multiple states. The inclusion of partners in coercion both provides important advantages and creates significant challenges, as has been explored most fully in the context of economic sanctions. In general, working with an ally increases relative power. In the context of sanctions, it adds to the economic costs of refusing to capitate to the demands. Moreover, it prevents the sanctioned state from substituting added trade with a third party for lost trade with a coercer. For example, a U.S. ban on Iranian oil imports may do negligible harm to Iran if China or Japan continues to purchase Iranian oil at the world price. However, conducting multilateral coercion effectively is no easy task. First are problems of divergent interests. Some states may want stronger demands or punitive measures than others, but consensus may limit the measures taken to the preferences of the most recalcitrant member of the coercing coalition. Even so, individual states may have incentives to free ride on their coalition partners, escaping the costs of punishing the target state themselves while reaping the benefits if it capitulates. When the approval of coercion means requires the approval of a certain group of states, as in the United Nations Security Council, a third level to this problem may emerge where the least-interested members can extort side-payments for their assent to providing a collective good. On top of these problems come the procedural difficulties of coordinating the foreign policies of multiple states with diverse interests. The bottom line is that multilateral coercion is often necessary but fraught with difficulties that are unlikely to be completely overcome.\textsuperscript{21}

**THE MUDDLED STATE OF THE EVIDENCE FOR THE STRUCTURAL EXPLANATIONS**

These simple structural explanations for crisis outcomes lie at the heart of the conventional wisdom. However, the extensive empirical debate over the relative explanatory power of the structural explanations has shown few signs of reaching a resolution, in large part because selection effects bias quantitative analysis of the requisite “how much?” questions. These variables rarely vary within cases

\textsuperscript{20} However, presumably states other than the two poles sometime have an intense interest in recruiting a superpower ally if they find themselves at odds with the other superpower (or even a regional power). Snyder and Diesing, *Conflict among Nations*, pp. 429-450.

sufficiently for conclusive qualitative analysis (else they would not be structural), and cross-case qualitative analysis suffers from the same methodological afflictions as quantitative studies.\footnote{I refer here to both nonrandom selection into the treatment and nonrandom selection into the universe of cases. See James D. Fearon, “Signaling versus the Balance of Power and Interests,” \textit{Journal of Conflict Resolution}, vol. 38, No. 2 (Jun., 1994), pp. 236-269; Achen and Snidal, “Rational Deterrence Theory and Comparative Case Studies,” pp. 143-169. See also Bear F. Braumoeller, “Causal Complexity and the Study of Politics,” \textit{Political Analysis}, Vol. 11, No. 3 (2003), pp. 209-233.}

The key to the selection effects problem is that if a state has a relative power or relative interest advantage, both it and its adversary will take this into account in making the decisions to make explicit threats or instigate crises. Consequently, a state with a relative power disadvantage may tend to avoid a crisis, whereas a state with an advantage may pursue one. China might threaten war with the United States under far different circumstances than, for example, it would with Vietnam. A straightforward empirical test would expect a less powerful actor to tend to lose coercive conflicts. However, it is quite plausible that less powerful states only tend to take actions that lead to conflicts when they have compensatory advantages, often on dimensions that are unobserved.\footnote{Fearon, “Signaling versus the Balance of Power and Interests.”} These selection effects would bias any simple empirical test of the benefits of relative power, including regression analysis and comparative case studies. The multiplicity of plausible selection effects renders it difficult, perhaps impossible, to even have a valid prior expectation as to the net direction of the bias. Disadvantaged states will try to select out of a conflict, and advantaged states to select in, but which effect is stronger is difficult (perhaps impossible) to deduce. The selection effects problem is quite general, and no satisfying solution has yet been found.

The implication is that one should not expect to observe a correlation between these structural variables that corresponds to their causal effect, because strategic selection into crises is a complicated process that masks these effects in a variety of ways. This requires great caution in testing and claiming to falsify these simple structural hypotheses than has been the case in some past research.\footnote{Paul Huth and Bruce Russett, “What Makes Deterrence Work? Cases from 1900 to 1980,” \textit{World Politics}, Vol. 36, No. 4 (Jul., 1984), pp. 496-526; Richard Ned Lebow and Janice Gross Stein, “Deterrence: The Elusive Dependent Variable,” \textit{World Politics}, Vol. 42, No. 3 (Apr., 1990), pp. 336-369; Paul Huth and Bruce Russett, “Testing Deterrence Theory: Rigor Makes a Difference,” \textit{World Politics}, Vol. 42, No. 4 (Jul., 1990), pp. 466-501; Zeev Maoz, “Resolve, Capabilities, and the Outcomes of Interstate Disputes, 1816-1976,” \textit{Journal of Conflict Resolution}, Vol. 27, No. 2 (Jun., 1983), pp. 195-229.} This means avoiding the assertion that relative power does not matter simply because it is uncorrelated with certain outcomes. The paucity of empirical support for the benefits of relative power and relative interest advantages likely reveals more about the shortcomings of current methodologies than the irrelevance of the variables.

\textbf{ARE THE STRUCTURAL EXPLANATIONS ENOUGH?}

There are compelling reasons to go beyond this predominantly structural approach. I briefly discuss these here in order to make the case that it is important to look beneath the broad structural
factors which shape crisis outcomes. First, and quite simply, few doubt that more power and more resolve are advantageous. Second, structural variables are rarely manipulable through policy (especially in the short term, i.e., during a crisis). Third, and most important, I will show that this approach cannot explain many of the actions taken in interstate coercive bargaining; there is wide latitude within these structural constraints for other variables to determine actions and outcomes. Holding structure constant, crises can still be won or lost depending on strategic choices. These strategic choices are neither irrelevant nor epiphenomenal. These ‘usual suspects’ structural variables are powerful determinants of coercive bargaining outcomes, but it aims to show that a great deal about the courses and outcomes of these conflicts cannot be understood with reference to these variables alone.

Having developed a set of structural explanations for crisis outcomes in this section, the remainder of the chapter lays out the conventional wisdom explanations for crisis actions, which in turn affect outcomes. Because the most important concept running through these explanations is signaling, the next section develops and clarifies the logic of signaling in coercive bargaining. The third and final section builds on this discussion by focusing on how to signal credibly with concrete actions, summarizing a series of ways in which the literature suggests this difficult task can be accomplished.

**Implementing Signaling Strategies in Crises**

Troop deployments, mobilizations, shows of force, public statements of resolve, and other actions of this general nature are the currency of crisis for the conventional wisdom. All fit the mold as signals of resolve. This section makes the connections between these crisis actions and the political objective of crisis victory. In particular, I consider 1) taking escalatory actions which incur a risk of war, 2) staking or investing in reputation, 3) staking audience costs, 4) mobilization and exploiting first-strike incentives, 5) making actions automatic, and 6) deploying forces in use-it-or-lose-it positions.

These signals are thought to be more than “cheap talk” because they stake commitments, something which provides a means of making signals credible. When a state stakes its reputation on standing firm in a particular dispute, (in principle) it commits to its position because it has increased the costs of backing down. By voluntarily altering its own incentives, a state can use commitment to elevate its signal of resolve from cheap talk to a signal that states not intending to back down would be more likely to send. Commitments therefore have two effects: the change in incentives itself, and the signal this change sends to the adversary.25

---

25 Commitments thought to undergird costly signals come in two forms, tying hands and sinking costs. Tying hands works by raising the costs of backing down, while sinking costs entails incurring costs that only a resolute actor would be willing to bear. James D. Fearon, “Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs,” *Journal of Conflict Resolution*, Vol. 41, No. 1 (Feb 1997), pp. 68-90.
BRINKSMANSHIP

Perhaps the single most prevalent explanation for crisis behavior is that states take escalatory actions in order to deliberately risk – but not start – war. This brinksmanship is intended to demonstrate the willingness to risk war to the adversary. It is credible if and when a more resolute actor is more willing to risk war to win in the crisis, while a bluffer would be unwilling to take the gamble. The convincingness of the signal therefore increases with the gravity of the risk of war that is incurred by it. Importantly, demonstrative actions such as shows of force that do not significantly increase the risk of war would not be expected to have this same effect. This strategy can also be used to threaten escalation during a limited war, for example by using conventional forces to target the adversary’s nuclear forces. Schelling (1960) aptly labeled this strategy “the threat which leaves something to chance.” Powell (1990) shows that even a perfectly rational actor desiring to avoid war might take these actions in order prevail in a crisis. Bracken (1988) considers the controversial proposition that some chance of accidental nuclear war can be beneficial because it enables the deterrence of limited wars. The threat which leaves something to chance ranks among the most prevalent explanations for limited escalatory behavior in crises. Snyder and Diesing (1977) offer an elaboration:

The balance of resolve cannot always be read off directly from the ‘balance of interests’ without the intervention of communicative and risk-taking activity that clarifies the interests of the parties. Thus one function, possible the most important function, of the expanded ensemble of crisis maneuvers in the nuclear age is to clarify interests. He who is most willing to run risks would appear to have the most at stake.

The brinksmanship strategy relies on the assumption that there is some uncontrollable risk of accidental war, but that the magnitude of this risk is subject to control. What this means is that the state in question can control whether the probability of accidental war is, say, 1% vs. 10%, but cannot control which outcome will result from the random draw governed by that probability. This might be

30 See also Lebow, Between War and Peace, ch. 4; Snyder and Diesing, Conflict among Nations, pp. 210-211,234-243; Kahn, On Escalation, pp. 63-64; Slantchev, Military Threats, p. 29,36-42; Jervis, The Logic of Images, p. 242. Shimshoni critiques this hypothesis on the grounds that most of these limited escalatory actions are carefully managed by policymakers and therefore “antithetical to the loss of control.” Shimshoni, Israel and Conventional Deterrence, p. 21.
31 Note that the signaler could be revealing its perception of a favorable balance of power instead of or in addition to a high level of interest in the stakes. Snyder and Diesing, Conflict among Nations, p. 456.
done by forward deploying forces in such a way as to increase the risks of incidents with the adversary. In extreme cases, brinksmanship might even take the form of deliberate limited uses of force.

However, the process through which an accidental incident escalates to full-scale war remains opaque in many accounts of brinksmanship. As long as war remains costly, the question of how and why to take the last step off the precipice looms large. In a study of sixteen crises, Snyder and Diesing (1977) conclude that “in all cases, elements of deliberateness predominate over elements of inadvertence, even though the deliberateness may be premised in very inaccurate perceptions of the situation.” They further find that decision-makers generally seek to minimize and mitigate risks of escalation, rather than deliberately incur them. Per this view, brinksmanship is available as a policy tool in crises, but seldom used.

**STAKING REPUTATION**

Reputation offers two related explanations for crisis behavior. First, states may take actions designed to stake their reputation as a means of tying their hands. For instance, it may be more damaging to a state’s reputation if it backs down immediately after proclaiming an inviolable commitment to stand firm than it would have been without that statement. If so, publicly declaring that commitment can credibly signal to the adversary that this state plans to escalate rather than capitulate. Second, states may take actions intended to invest in a desired reputation. For instance, standing firm even over secondary matters of relatively low importance has been justified on the grounds that through reputation that action is beneficial in future conflicts of greater importance. These ideas parallel the logic of audience costs (next section), save that the primary concern is the leader’s reputation with the domestic political audience.

The question of why actions such as public statements of commitment and troop deployments serve to stake reputations has received surprisingly little attention. The longstanding assumption has been that a retreat after a public statement pledging to stand firm imposes greater reputation costs than to retreat without one. Yet, this should not be taken for granted. By backing down, the state would seem to have risked revealing something about its interests and resolve regardless of its explicit statements. For example, American policymakers anticipated considerable reputational damage were they to have passively acquiesced to North Korea’s invasion of South Korea in 1950, despite the lack of an explicit commitment. Hopf (1994) similarly concludes that “deterrence theory needs to be expanded to capture the deterrent value in losses that affect the challenger, despite inaction on the part of the

---


33 Snyder and Diesing, *Conflict among Nations*, p. 16. In context, this quotation could be construed as applying only to crisis onset, though I do not believe this to be the case.

34 Ibid., p. 243. However, the notion of using the threat which leaves something to chance in order to signal resolve is not fully developed, and instead is considered mainly in terms of its direct effect on incentives (pp. 234-235).

35 Snyder and Diesing, *Conflict among Nations*, p. 251; Jervis, *The Logic of Images*, p. 82; etc.
defender.” Snyder and Diesing (1977) find no examples in which make a commitment explicit successfully committed a state to do something it was not already committed to do by its interests. If reputation is on the line regardless of public statements, those statements lose much of their effectiveness as signals.

States seek to be perceived as having more than one signal trait. Although a reputation for resolve is perhaps the most common goal, it is not alone. In many cases, states must convey that they are status quo powers willing to ultimately sign and keep a war-avoiding agreement. If a state believes its adversary to be revisionist, set on war, undeterred, etc., it has little reason to avoid war and, potentially, strong incentives to strike first.

These two desirable reputations usually call for diametrically opposed behavior. To appear as a status quo power, a state will generally want to take conciliatory actions that a state dead set on war would not have taken. To appear resolute, a state will generally want to take aggressive actions and avoid concessions. Virtually any action taken during a crisis can be interpreted as a signal balancing these two objectives. This renders the empirical testing and falsification of signaling quite difficult.

**AUDIENCE COSTS**

The theory of audience costs has risen rapidly to rival other forms of signaling as a leading explanatory theory of crisis behavior. The central idea is that leaders leverage the specter of being punished by their domestic audience (e.g., at the ballot box) if they back down. Because this fear ties the hands of leaders to stand firm, the adversary must back down.

Audience costs can be understood as a strategy or as a constraint. The constraint view is simpler. Backing down, failing, accepting too little, or otherwise losing an interstate conflict is unpopular domestically, and tends to result in the punishment and/or removal of the leaders who presided over the debacle in question. Because the public is less informed than policymakers about

---

what outcomes were achievable, they punish leaders for negative outcomes even when granting concessions was in the national interest. Consequently, leaders are constrained to stand firm, to refuse to compromise, to take risks, and even to escalate in order to avoid not just the strategic costs of concessions, but also the domestic political costs. This constraint variant of audience costs is relatively uncontroversial, although its importance to decision-makers relative to considerations of national interest remains an open question.

The strategy of audience costs relies on a much stronger claim. It builds on this constraint to argue that leaders deliberately take escalatory actions that increase the domestic political costs of backing down. This leads the adversary to recognize these incentives and back down rather than further challenge an adversary newly committed to holding firm. This is Fearon’s (1994) original definition of audience costs. Later work of note expanded on the premise by integrating the domestic political opposition and the media.40

There is no consensus on the exact relationship between crisis signals and audience costs. Fearon’s (1994) assumption that that audience costs steadily mount with escalatory actions in crisis is tenuous. What escalatory actions qualify? When a policymaker backs down after publicly declaring that his nation would not do so, it is possible that the increase in embarrassment is so large as to override the national interest considerations that otherwise shape policy decisions. But, it is also possible that the audience costs depend only on the final outcome, and not on the actions taken.41 A public concerned with a particular foreign policy issue might punish leaders for failure in that area regardless of whether or not a public commitment to stand firm was made. Snyder and Borghard (2011), for instance, conclude that publics judge and punish political leaders based on their performance in achieving popular goals independent on whether they have made explicit commitments.42 If so, audience costs could act as a potent constraint on state policy even as the types of signaling actions that create them (public declarations of resolve, shows of force, etc.) do not carry much weight.

MOBILIZATION

Another option available to states in crises is to mobilize or forward-deploy forces so as to improve their prospects should the situation escalate to the limited or unlimited use of force. This military advantage can itself be the objective of these actions, which is important to bear in mind, but to what extent can policymakers use these measures to credibly signal resolve and thereby prevail without

41 Alternatively, pushing a crisis to the brink of war with additional escalatory actions might serve to inform the domestic audience that an outright coercive victory is out of reach without intolerable risks. This assumption of steadily mounting audience costs is stated explicitly in Fearon’s original formulation, but has received little scrutiny since. Fearon, “Domestic Political Audiences,” pp. 577, 580, 582.
the costly recourse to force?\textsuperscript{43} Particularly for large conscript armies which much disrupt the civilian economy by calling up reserves in order to mobilize, mobilization exacts a nontrivial economic cost that mounts with time. States with less interest at stake have an incentive to avoid this cost.

Moreover, military mobilization can be understood as a form of brinksmanship when it intersects with first-strike advantages. The July Crisis preceding World War One offers an example of acts of mobilization that credibly signaled a willingness to fight by opening windows of opportunity that elevated the likelihood of war. In particular, Russia’s decision to mobilize risked war by presenting Germany with a window of opportunity to strike before Russian forces reached peak strength. This action could plausibly be expected to increase German and Austrian perceptions of Russia’s willingness to fight for Serbia. Of course, in this instance the result was not success due to effective signaling, but war.\textsuperscript{44} Yet it is this very same increase in the probability of war which endows the decision to mobilize with credibility as a signal, because only states more willing to fight are more likely to take actions that elevate the probability of war. Wielding the preemptive motivation for war for signaling purposes follows the general logic of deliberately incurring a risk of war in order to signal resolve, and as such it is inherently perilous.

**AUTOMATICITY AND USE-IT-OR-LOSE-IT DEPLOYMENTS**

If an action can be made automatic conditional on some adversary action or inaction, the commitment to take that action is made credible. The prevalent analogy is between a crisis and a game of chicken played by two teenage drivers. In a test of nerves, the two drive directly towards one another, and the first to swerve loses. Both drivers would prefer the other to swerve first, but both also prefer to back down first to a fatal collision. One strategy which on paper assures victory in the game of chicken is to remove one’s steering wheel and throw it out the window, in full view of the adversary. Once the adversary knows turning is impossible, his only remaining choice is between a deadly collision and backing down by swerving. By sacrificing control over steering and making a straight course automatic, a driver can credibly commit to not backing down, something that would otherwise be impossible given his interest in avoiding collision.\textsuperscript{45}

But do automatic commitments exist in real-world crises? Schelling (1966) treats automaticity as one of the two major sources of commitment, along with reputation. Blechman and Kaplan (1978) conclude that, at least in some circumstances, “Forces deployed abroad will become automatically involved in outbreaks of conflict.”\textsuperscript{46} However, it is difficult to find ways in which threats can be made

\textsuperscript{43} These are two distinct mechanisms through which mobilization can affect outcomes. On the distinction see Slantchev, *Military Threats*.

\textsuperscript{44} For the theory that first-strike opportunities make war more likely, see, e.g., Van Evera, *The Causes of War*.


The classic example is form a work of fiction. In the movie *Dr. Strangelove*, the Soviet Union builds a doomsday device designed to destroy the world if a nuclear weapon detonates on Soviet territory. Snyder and Diesing (1977) offer four examples of creating “irrevocable commitments” during a crisis: the blockade of Berlin in 1948, the airlift to relieve it, the U.S. deployment of troops to Lebanon in 1958, and the construction of the Berlin Wall in 1961. On reflection, however, none of these were irrevocable. Each could have been undone with the appropriate order. Each can also be better explained as a targeted *fait accompli* exploiting a vulnerable red line.

One possible method for making punishment automatic is to exploit the dynamics of organizations. Even if a state’s leaders cannot credibly threaten to take a particular punitive action, they might be able to so deeply embed a standard operating procedure to launch in response to certain adversary behavior that the officers in charge of implementing the attack would do so regardless of the consequences. Using allies whose vital interests are more directly threatened in a similar manner offers an analogous approach. Yet, even if standard operating procedures were emplaced so as to render a nuclear launch automatic given some adversary action, it is difficult to see how this could be credibly conveyed to the adversary. Nor is this notion easily linked to crisis signaling. Automaticity may be a mechanism that is far more compelling in theory than in practice.

In certain situations, a state may be able to deploy its forces to a vulnerable forward location from which it must either use them even in what might otherwise be a limited conflict or risk losing them. If the loss of these forces cannot be afforded, this strategy can tie hands so as to make credible the threat to stand firm. AN example of this strategy is Chiang Kai-shek’s deployment of a large fraction of his ground forces to the small islands of Kinmen and Mazu during a crisis in which an invasion of those islands from mainland China appeared imminent. By increasing the costs of backing down, this action tied his hands in a way that only an actor not expecting to back down was likely to accept. Although Schelling (1966) describes this strategy as one of making the commitment to the islands automatic, at least as a technicality it is worth noting that resistance by the forces was not quite automatic. Instead, the Nationalist forces were deployed such that they (and likely additional forces) must be used to defend Kinmen and Mazu if it was deemed necessary to preserve the forces unable to withdraw. In that sense, use-it-or-lose-it deployments can be a seen as a lesser but still related cousin of automaticity.

**Conclusion**

---

48 This credible commitment fails because the Soviets fail to convey its existence to the United States.
49 Snyder and Diesing, *Conflict among Nations*, p. 227. This problem seems to be noted on p. 230.
51 Ibid., p. 200.
52 This possibility may also contribute to fears of accidental war, thereby potentially enabling the “threat which leaves something to chance.”
53 Schelling, *Arms and Influence*, p. 43.
This chapter has characterized a dominant conventional wisdom on interstate coercion, emphasizing signaling strategies when explaining crisis actions and structural explanations for crisis outcomes. Taken together, this conventional wisdom paints a picture of crises as contests in signaling by competitive risk-taking and commitment-staking shaped and constrained by relative power and relative interest. But, how does this crisis bargaining work at the tactical level? One possibility is that states signal resolve in order to intimidate the adversary into concessions or inaction. Another possibility is that states search out and exploit opportunities – vulnerabilities in red lines – to make unilateral gains, often by *fait accompli*. Although ultimately consistent with the broader theoretical framework through which the conventional wisdom approaches crisis tactics, this offers a distinctly different perspective on those tactics.
Ch. 4 Gray Areas and Land Grabs in Territorial Crises

How often do faits accomplis strike vulnerabilities in deterrent red lines such as gray areas? Do gray areas make faits accomplis more likely? Are states better able to unilaterally seize gains while avoiding war by targeting gray areas? More fundamentally, is it common for states to make gains by fait accompli, or are gains by coercing concessions more prevalent? This chapter explores each of these questions in the domain of interstate crises over territory between 1918 and 2007.

Territorial crises provide the best available pool of cases within which to evaluate the relationships between weak red lines, faits accomplis, and crisis outcomes. Analyzing crises over diverse issue types would create serious problems of comparability and inference. Is ambiguity in a red line against building nuclear weapons truly comparable to an imprecisely-demarcated border? Is a fait accompli seizure of territory comparable to unilaterally fishing in disputed waters? Limiting the scope of the analysis to territorial crises ensures comparability across cases. Red lines against nuclear proliferation and support for rebels are riddled with weaknesses, but using the tendency of these red lines to fail as evidence is problematic; perhaps deterrence is harder in these issue areas for other reasons? Moreover, territorial crises are intrinsically important. Territory is the most common issue over which crises arise, and wars are particularly likely to break out over territory.

This chapter tests three of the four hypotheses drawn from the theory in Chapter Two. These hypotheses read as follows when made specific to territorial crises: H1: Gray areas in red lines make land grabs more likely.² H3: Land grabs that strike gray areas are less likely to lead to war than land grabs that strike other territories. H4: Gray areas make war more likely by making land grabs more likely. Chapter Two underscored the tension between the third and fourth hypotheses, a point to which I return below.

The cornerstone for this chapter is new data on all ‘land grab’ faits accomplis since 1918. A land grab takes place when one state uses its military to unilaterally seize a disputed piece of territory with the intention to assume control of that territory. In general, the objective is to make a territorial gain without enduring the costs of a major war. For that reason, this chapter examines crises over pieces of a state’s territory, not disputes over the full territory of a state. Land grabs are frequently the signature event of territorial crises and play a major role in determining whether these crises escalate to war.

¹ I omit the second hypothesis because it deals with the gain-loss dimension of crisis outcomes. In this chapter, I focus on a more readily quantifiable dimension of outcomes: whether or not the crisis escalates to war. The case studies provide empirical tests for the second hypothesis.
Gray areas in territorial deterrent red line come in two forms. These gray areas exist when the disputed territory is either an island located awkwardly between two core territories or a region where the border is ambiguous. Although any study of crises is prone to selection effects – a problem I return to in later sections – these offer two comparatively exogenous sources of gray areas in deterrent red lines.²

This chapter shows the following: First, states acquire territory far more often by fait accompli than by coercing a concession. Second, two-thirds of all land grabs since 1918 struck gray area territories (and the vast majority since 1980). Third, land grabs are twice as common in crises over gray area territories versus crises over other territories. States are less likely to seize a disputed territory without a gray area to exploit. Fourth, land grabs striking gray areas lead to war half as often as land grabs against other territories. States consistently cannot “get away” with a land grab without provoking war unless the territory is a gray area.

Russia’s annexation of the Crimean Peninsula from Ukraine in 2014 provides a clear illustration of how these findings apply to a particular case. Note that although Crimea is not an island, it has the same physical geographic separation from the core territory of Ukraine that makes for a gray area. The patterns in the history of land grabs suggest the following: First, a land grab in the Crimean Peninsula was relatively likely to take place in the event of a Russia-Ukraine territorial crisis. Second, war was relatively unlikely to ensue from a land grab in Crimea. Third, a land grab in other regions of Ukraine such as Donetsk and Luhansk would be less likely. Fourth, a land grab in those regions within Ukraine’s core territory – even one seizing territory no larger than Crimea – would be considerably more likely to start a war between Russia and Ukraine.³ These predictions are by no means certain or deterministic, but if one applies the theory and the corresponding patterns from the history of interstate crises, that set of predictions is the result.

The chapter is structured as follows. The first section establishes that – at least since 1918 – states have far more often made gains by fait accompli than by coercion. It presents new data on all land grabs since 1918 as the anchor for this comparison. The second section summarizes the broader data on territorial crises and the measurement of each variable. The third provides evidence for a strong empirical association between gray areas and land grabs. The fourth turns to and evaluates the impact of both gray areas and land grabs on the likelihood of war with special emphasis on the interaction effect between them. The results suggest that land grabs more often strike gray areas because states are more able to get away with faits accomplis when they target gray areas.

---


³ Of course, red lines considerations are far from the only factors shaping this or any other particular case. In this instance, the relative power disparity between the Russian military and the Ukrainian military was unusually severe. Political disorder in Ukraine could have further hindered an effective military response to a second land grab. For those reasons and others, it is possible that Russia could have taken more without starting a war.
How States Make Territorial Gains in Crises

Do states make gains by coercing concessions? Or instead by unilaterally taking those gains by *fait accompli*? The territorial issue area offers a pool of comparable cases with which to approach this question. To assess how often states gain territory by coercion during crises, I searched through every crisis in the Interstate Crisis Behavior (ICB) dataset, which runs from 1918 to 2007. I then expanded the analysis beyond crises by searching through the Militarized Compellent Threats (MCT) dataset and all instances of territorial “cession” in the Territorial Change dataset since 1918.

The result was the surprising finding that not once in the last 50 years has a state successfully coerced another into ceding territory under threat without using its military to seize the territory first. Since 1945, a coerced territorial cession has happened only once. Indonesia pressured the Netherlands into relinquishing what is now Indonesia’s half of New Guinea in 1963. In the full period, 1918-2014, the number of coerced territorial cessions grows to ten. Table 4.1 lists these ten. The nine additional cases all occurred amid the unique international climate that existed between 1937 and 1940. Most took place in Eastern Europe. Since 1918 – and especially since 1945 – coercive threats demanding territory were common and signals of resolve were frequent; neither consistently produced territorial gains.

In contrast, states have unilaterally deployed military forces to seize territory at least 88 times since 1918. 63 of the 88 occurred since 1945. 44 of the 88 produced lasting territorial gains. Many of those that did not fail to do so because the adversary immediately retaliated with a *fait accompli* retaking the territory. 66 of the 88 occurred during ICB crises, as did all ten of the coerced territorial cessions. 27 of the 88 took place in crises that erupted into wars. None of the ten coerced cessions led to immediate war, but most were overturned by the Second World War within a decade. It is possible to draw the comparison in a variety of ways, but the bottom line is clear: states gain territory by *fait accompli* far more often than by coercion.

Table 4.2 lists these land grabs. The list contains 66 distinct cases. The asterisks indicate land grabs that provoked an immediate retaliatory land grab retaking the same territory which was just

---

5 Arguments can be made for the Spanish decision to cede Western Sahara in 1975 to Morocco and Mauritania and the Soviet withdrawal from portions of Iran in 1946. I code the former as a land grab due to the Green March and the latter as non-coercive.
6 Note that these adversaries could have responded to a land grab by coercing a pullback (instead of imposing one by *fait accompli*). This succeeded only once (Japan-USSR 1937).
7 Most of these cases appear to be missing from the ICB dataset because they are comparatively obscure crises rather than because no crisis was underway.
8 Unfortunately, I lack data on failed attempts to coerce concessions. Some of these cases may have led to an immediate war. Failed coercion does precede many land grabs.
seized during the same crisis. These cases contain two land grabs. Table 4.2 also includes data that become important in later sections. Columns 4 and 5 indicate the presence of gray areas using the two predominant sources of them in territorial conflicts: islands (including island-like peninsulas) and border ambiguities. Column 6 indicates whether the case ended in “full-scale war” by the ICB definition. Figure 4.1 displays the chronology of land grabs since 1918 against that for coerced territorial cessions, underscoring again the discrepancy between the two.

---

9 A few land grabs provoke retaliatory land grabs seizing other territories. I include both sides of such cases in Table 4.2.
10 None contain three. Cases with retaliation and counter-retaliation have already escalated to war.
<table>
<thead>
<tr>
<th>Year</th>
<th>Target</th>
<th>Coercer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1937</td>
<td>France (Syria)</td>
<td>Turkey</td>
</tr>
<tr>
<td>1937</td>
<td>Soviet Union</td>
<td>Japan</td>
</tr>
<tr>
<td>1938</td>
<td>Czechoslovakia</td>
<td>Germany</td>
</tr>
<tr>
<td>1938</td>
<td>Czechoslovakia</td>
<td>Hungary</td>
</tr>
<tr>
<td>1938</td>
<td>Czechoslovakia</td>
<td>Poland</td>
</tr>
<tr>
<td>1939</td>
<td>Lithuania</td>
<td>Germany</td>
</tr>
<tr>
<td>1940</td>
<td>Romania</td>
<td>Soviet Union</td>
</tr>
<tr>
<td>1940</td>
<td>Romania</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>1940</td>
<td>Romania</td>
<td>Hungary</td>
</tr>
<tr>
<td>1963</td>
<td>Netherlands</td>
<td>Indonesia</td>
</tr>
<tr>
<td>Year</td>
<td>Land Grab AGAINST</td>
<td>Land Grab BY</td>
</tr>
<tr>
<td>------</td>
<td>-------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>1919</td>
<td>Russia</td>
<td>Romania</td>
</tr>
<tr>
<td>1919</td>
<td>Russia</td>
<td>Finland*</td>
</tr>
<tr>
<td>1920</td>
<td>Russia</td>
<td>Poland</td>
</tr>
<tr>
<td>1920</td>
<td>Lithuania</td>
<td>Poland</td>
</tr>
<tr>
<td>1921</td>
<td>Panama*</td>
<td>Costa Rica</td>
</tr>
<tr>
<td>1921</td>
<td>Albania*</td>
<td>Yugoslavia</td>
</tr>
<tr>
<td>1923</td>
<td>Greece</td>
<td>Italy</td>
</tr>
<tr>
<td>1928</td>
<td>Bolivia*</td>
<td>Paraguay</td>
</tr>
<tr>
<td>1931</td>
<td>China</td>
<td>Japan</td>
</tr>
<tr>
<td>1932</td>
<td>Paraguay*</td>
<td>Bolivia</td>
</tr>
<tr>
<td>1932</td>
<td>Colombia*</td>
<td>Peru</td>
</tr>
<tr>
<td>1933</td>
<td>Saudi Arabia</td>
<td>Yemen</td>
</tr>
<tr>
<td>1937</td>
<td>Japan</td>
<td>Russia</td>
</tr>
<tr>
<td>1938</td>
<td>Japan</td>
<td>Russia</td>
</tr>
<tr>
<td>1939</td>
<td>Finland</td>
<td>Russia</td>
</tr>
<tr>
<td>1954</td>
<td>Japan</td>
<td>South Korea</td>
</tr>
<tr>
<td>1956</td>
<td>Egypt</td>
<td>Israel</td>
</tr>
<tr>
<td>1958</td>
<td>Sudan</td>
<td>Egypt</td>
</tr>
<tr>
<td>1959</td>
<td>China*</td>
<td>India</td>
</tr>
<tr>
<td>1959</td>
<td>India</td>
<td>China</td>
</tr>
<tr>
<td>1961</td>
<td>Portugal</td>
<td>India</td>
</tr>
<tr>
<td>1962</td>
<td>China*</td>
<td>India</td>
</tr>
<tr>
<td>1963</td>
<td>Algeria</td>
<td>Morocco</td>
</tr>
<tr>
<td>1965</td>
<td>India</td>
<td>Pakistan</td>
</tr>
<tr>
<td>1965</td>
<td>Pakistan</td>
<td>India</td>
</tr>
<tr>
<td>1966</td>
<td>Guyana</td>
<td>Venezuela</td>
</tr>
<tr>
<td>1967</td>
<td>Egypt</td>
<td>Israel</td>
</tr>
<tr>
<td>1967</td>
<td>Jordan</td>
<td>Israel</td>
</tr>
<tr>
<td>1967</td>
<td>Syria</td>
<td>Israel</td>
</tr>
<tr>
<td>1969</td>
<td>Russia</td>
<td>China</td>
</tr>
<tr>
<td>1971</td>
<td>UAE</td>
<td>Iran</td>
</tr>
<tr>
<td>1973</td>
<td>Israel</td>
<td>Egypt</td>
</tr>
<tr>
<td>1973</td>
<td>Israel</td>
<td>Syria</td>
</tr>
<tr>
<td>1974</td>
<td>South Vietnam</td>
<td>China</td>
</tr>
<tr>
<td>1974</td>
<td>China</td>
<td>South Vietnam</td>
</tr>
<tr>
<td>1974</td>
<td>Cyprus</td>
<td>Turkey</td>
</tr>
<tr>
<td>1975</td>
<td>Spain</td>
<td>Morocco</td>
</tr>
<tr>
<td>1977</td>
<td>Ethiopia*</td>
<td>Somalia</td>
</tr>
<tr>
<td>1978</td>
<td>Tanzania</td>
<td>Uganda</td>
</tr>
<tr>
<td>1980</td>
<td>Iran</td>
<td>Iraq</td>
</tr>
</tbody>
</table>

* This land grab provoked an immediate land grab retaking the contested territory.
Table 4.2 Land Grabs, 1918-2014

<table>
<thead>
<tr>
<th>Year</th>
<th>Territorial Fait Accompli AGAINST</th>
<th>Territorial Fait Accompli BY</th>
<th>Island or Peninsula</th>
<th>Ambiguity in Land Border</th>
<th>War</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>Peru*</td>
<td>Ecuador</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>Britain*</td>
<td>Argentina</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>Chad*</td>
<td>Nigeria</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>China, Vietnam</td>
<td>Malaysia</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>Pakistan</td>
<td>India</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>Thailand</td>
<td>Laos</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>China</td>
<td>India</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>Bahrain</td>
<td>Qatar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>India</td>
<td>China</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>Laos*</td>
<td>Thailand</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>Vietnam</td>
<td>China</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>China</td>
<td>Vietnam</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>Azerbaijan</td>
<td>Armenia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>Qatar</td>
<td>Saudi Arabia</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>Nigeria</td>
<td>Cameroon</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>Philippines</td>
<td>China</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>Peru*</td>
<td>Ecuador</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>Eritrea*</td>
<td>Yemen</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>Turkey</td>
<td>Greece</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>Greece</td>
<td>Turkey</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>Peru</td>
<td>Ecuador</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>Ethiopia*</td>
<td>Eritrea</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>India*</td>
<td>Pakistan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Spain*</td>
<td>Morocco</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>Eritrea</td>
<td>Ethiopia</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>Ukraine</td>
<td>Russia</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* This land grab provoked an immediate land grab retaking the contested territory
Figure 4.1: When and How States Acquire Territory from Adversaries
When Russia occupied and annexed the Crimean Peninsula in 2014, it did so via *fait accompli* rather than via coercion. When Argentina took the Falkland Islands from Britain in 1981, it did so in a surprise *fait accompli*. When Britain resolved to regain the islands, it demanded their return. Argentina refused. British forces then retook the islands in an invasion of its own while carefully avoiding attacks on Argentina itself, even on Argentinean naval vessels in port. When India confronted Portugal with demands for the enclave of Goa — demands backed by an overwhelming advantage in military power in the region — the Portuguese government refused, holding its ground and ordering its forces to fight to the bitter end. After giving up on threats and negotiations, India occupied Goa in 1961. These examples fit the broader pattern.

A land grab is defined as occurring when one state uses its military to unilaterally seize a disputed piece of territory with the intention to assume control of that territory. Usually — but not always — this intention to assume control involves formal annexation or the intent to eventually annex. This criterion excludes cross-border military operations for other purposes like interventions in civil wars, peacekeeping missions, raids on rebel bases, etc. To qualify as a land grab, the challenger need only hold the newly acquired territory momentarily. Land grabs are defined as a type of crisis behavior, so the outcome — i.e., whether the challenger retains lasting control of the territory — does not factor into the definition. Nor does whether the case then escalates to war.

For both the land grabs and the coerced territorial cessions, I exclude all cases like Iraq’s 1990 invasion of Kuwait where the disputed territory was the full territory of a crisis actor. Neither coercion nor the *fait accompli* is well-suited to this unlimited objective. For coercion, why surrender without a fight if there is any hope for victory? Similarly, *faits accomplis* by definition seize something of limited value in an attempt to get away with the gain without provoking war. Unlimited invasions imply a strategy more aggressive than a *fait accompli* and leave the adversary no choice but to fight or lose everything. Nonetheless, all results in the chapter are robust to including these full-territory cases.

---


12 It is difficult to apply to a small number of cases, with Turkey’s 1974 invasion of Cyprus and Israel’s 1956 seizure of the Sinai Peninsula and Gaza Strip among the most challenging cases; both are included. In contrast, India’s 1971 role in separating West Pakistan (Bangladesh) from East Pakistan (Pakistan) is not considered a land grab as India did not seek control of the territory.

13 States attempting to acquire territory generally aim for a small piece (far less than half) or the full territory of the adversary. Cases around 50% are virtually nonexistent. This makes strategic sense. Challengers either try to get away with a limited gain or accept that the adversary will resist fully. If the adversary will resist fully, why limit war aims? In a few cases, a coalition occupies the full territory of a state, but one or more members of the coalition receive only a smaller piece (e.g., the Soviet Union’s share of Poland in 1939). I exclude these cases.

14 Future studies using these data on land grabs may wish to include these full-territory invasion cases as land grabs. To facilitate this, I code all invasions where the dispute was over the full territory of a state as land grabs in the data. However, I mark all crises where the full territory of a crisis actor is disputed with the variable *fullterr* and use it to exclude them from analysis. There are also cases of all-out invasion where territorial change, overt conquest, border adjustment, etc. were not on the table (e.g., Grenada 1983, Panama 1989, Iraq 2003). I code these as non-territorial issues, although for the purposes of other studies they might reasonably be considered territorial.
To locate all land grabs since 1918, I examined every crisis in the Interstate Crisis Behavior dataset. I then combed through territorial Militarized Interstate Disputes data, Territorial Dispute data, Territorial Change data, Correlates of War data on interstate wars, and any information on a land grab I came across from other sources. More specifically, the list aims to encompass all land grabs associated with a dispute, crisis, or war, because it is too difficult to identify cases in which states occupy small territories without any public controversy or complaint. Data on these land grabs will be made available for use in future studies.

Why new data? The Militarized Interstate Dispute (MID) dataset and the Territorial Change dataset contain variables which come closest to identifying land grabs, but neither can be used to generate something which resembles the list in Table 4.2. The MID dataset includes data for the “highest action” (in terms of escalation) taken by each actor during the dispute. Although the 14th level of this variable is “occupation of territory,” cases only enter this category when no higher escalation occurs. Any case that escalates to, for instance, “attack” (level 17), “clash” (level 18) or war (levels 20 and 21) is not coded as an occupation of territory. This leaves out many land grabs and precludes evaluating relationships involving land grabs and war. Moreover, the majority of MID “occupations of territory” are not land grabs, because any cross-border incursion for any purpose qualifies (e.g., hot pursuit of criminals). Similarly, the Territorial Change dataset divides territorial changes into seven categories, three of which include land grabs: “conquest,” “annexation,” and (more rarely) “cession.” Along with many other differences, these three categories contain numerous cases of events other than a land grab. They together include 225 cases since 1918, far more than the number of land grabs in that period. In sum, neither dataset was designed to identify land grabs.

Coercion and the fait accompli are not the only way territory changes hands. The two are, however, the primary adversarial means by which states acquire territory at each other’s expense short of war. Territory also changes hands at the end of wars and – quite frequently – through negotiated agreements. Could these negotiated agreements be coercion? In examining all territorial cessions since 1918, I searched for cases where latent coercion may have factored into a negotiated outcome despite the absence of a crisis (by ICB criteria) or a compellent threat (by MCT criteria). I found that most of these agreements tend to occur years after a crisis – if a crisis occurred at all – and without anything resembling an explicit compellent threat. Many involve a compromise, rather than a one-sided outcome. Of those that do not, one side often prevails due to a favorable ruling by an international legal institution or arbitrator. The three strongest candidates for latent coercion are the British return of

---

15 I found most of the cases absent from the ICB data in the MID data and no other source. These tend to be relatively obscure and lower-intensity cases. I would like to thank Ken Schultz for proving the case narratives on territorial MIDs that made this undertaking far more efficient. Because these narratives focused on homeland territorial disputes and end in 2001, missing land grabs are more likely outside of those bounds. On these datasets: Daniel Jones, M. Stuart, A. Bremer, and J. David Singer, “Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns,” Conflict Management and Peace Science, Vol. 15, No .2 (1996), pp. 163-213; Huth, Standing Your Ground; Diehl and Goertz, Territorial Changes and International Conflict.

16 Both codebooks and datasets are available from the Correlates of War webpage. Also see Diehl and Goertz, Territorial Change and International Conflict, pp. 53-54.
Hong Kong to China, the Israeli return of the Sinai to Egypt, and the Spanish cession of Ifni to Morocco. Beyond those three, it was difficult to find cases where a strong argument exists for coercion driving the outcome.

Overall, the disparity between gains by coercion and gains by fait accompli is striking, all the more so because the literature has lavished attention on crisis signaling and ‘compellence’ while giving short shrift to faits accomplis. This has left four important questions unanswered. 1) When are faits accomplis likely to occur? 2) When are faits accomplis more likely to provoke war? These questions are the subject of the remainder of the chapter. The disparity also raises two questions for future studies to address: 3) Why are coerced concessions so rare in comparison to faits accomplis? 4) What does the rarity of coerced concessions mean for existing theories of the causes of war?

Data and Measurement

This section summarizes the data and the measurement decisions that went into creating it. An online appendix associated with these data will contain brief explaining the main coding decisions for each case. The analysis began with the ICB data on territorial crises between 1918 and 2007. For each case, I began by looking closely at the ICB case summary for that crisis. Where it was possible to code the variables on that basis, I did so. Where it was not, I consulted additional secondary sources as cited in the appendix.¹⁷

The unit of analysis is the crisis actor, i.e. a state involved in a crisis. The ICB project defines crisis as existing for a state when its highest level decision-makers perceive all three of the following: “a threat to one or more basic values, an awareness of finite time for response to the value threat, and a heightened probability of involvement in military hostilities.”¹⁸ The analysis assesses only primary crisis actors directly involved in territorial disputes.¹⁹

¹⁷ This appendix and the data will be made available on the author’s website after publication.
¹⁸ One significant drawback of the definition for this study is its propensity to exclude aggressors on the grounds that they did not view a particular confrontation as threatening. For example, Germany is not included as a crisis actor in the 1939 crisis over its invasion of Poland. When this definitional criterion excludes an otherwise usable primary crisis actor, it is noted in the case description in the appendix. Per this definition, the ICB also includes intra-war crises, i.e. crises that start during wars that were already well underway. Because this chapter focuses on peacetime crises, I exclude these intra-war cases along with crises involving external interveners participating in ongoing civil wars. Brecher and Wilkenfeld, A Study of Crisis, p. 3.
¹⁹ Secondary crisis actors excluded by this criterion fall into two categories. First are actors with an interest in the outcome of the crisis but no major role in the course of events. Japan in the 1939 crisis surrounding Germany’s invasion of Poland offers an example. Second are allies playing a subordinate role alongside a primary crisis actor. These states merely endorse a more involved ally’s territorial red line, and including them would amount to double-counting the case more than incorporating new information. Even where I include an ally as a primary actor due to their prominent role in the crisis, I mark that their territorial red line is extended deterrence and exclude them from the analysis.
Because this is a study of deterrent red lines, the analysis is set up to explain whether each crisis actor is or is not struck by a land grab. The exclusive focus is on the red lines of states which stand to lose territory.\textsuperscript{20} The data do, however, also include whether each crisis actor commits a land grab.

A territory can enter a crisis 1) occupied by the crisis actor, 2) occupied by the adversary, or 3) unoccupied. States cannot deter the seizure of a territory that their adversary already occupies. It is therefore necessary to distinguish crisis actors whose “pure compellent” red lines are not open to violation by land grab because the adversary occupied the territory prior to the start of the crisis. I exclude these cases from the analysis to follow, but the results are robust to including them.\textsuperscript{21}

The data on territorial crises contain 131 crisis actors that set deterrent red lines protecting pieces of territory in 102 interstate crises. Of these red lines, land grabs strike 59 (45%), a reminder that territorial crises often reach crisis status due to a land grab. Gray areas occur in 62 cases (47%). This high rate of gray areas may indicate that crises are more likely to break out over territorial gray areas than other disputed border areas.

Gray areas in territorial red lines are regions that do not fit unmistakably within the territorial core of a state. A “strong” territorial deterrent red line, as I use the term in this chapter, is a red line without any gray areas. Strong red lines are precisely delineated such that it is clear whether each piece of territory falls inside or outside the line. Strong red lines are also set on a focal point (see Chapter Two). In territorial conflicts, both physical and political geography provide these focal points in the forms of major waterways and established borders, respectively.\textsuperscript{22}

Gray areas in territorial conflicts come in four specific varieties. Per the typology in Chapter Two, the first pair qualifies as arbitrariness and the second pair as imprecision. First, a gray area exists when two focal points leave a small awkward area in between them. River islands fit this description. Each branch of the river is a focal point. The line against invasion across the river is stark, but what of islands in its midst? Damansky (Zhenbao) Island, for instance, figured prominently in the Sino-Soviet border conflict of 1969. Second, a gray area exists when a small area falls awkwardly outside the salient

\textsuperscript{20} Examining territorial crises requires determining whether territory is a main issue at stake in each crisis. Although the ICB data contains two variables which pertain to whether the issue at stake was territory, neither is optimized for the present analysis. First, the ICB data for the “gravity” of a crisis for a crisis actor include the option “territorial threat.” However, “territorial threat” includes cases such as support for a separatist rebel group where the challenger seeks no territory from the deterrer. Second, the ICB data for the “issue of the crisis” include the option “military-security” (defined to include territory). This variable includes many other types of issues, e.g., issues of freedom of navigation at sea. Consequently, I manually coded each ICB crisis for whether the primary issue was territory. The appendix includes justifications for excluded (non-territorial) crises when the determination seemed open to question.

\textsuperscript{21} However, some land grabs do in fact strike states which set pure compellent red lines. This happens when the crisis actor seizes the territory first, enabling the adversary to retake it.

\textsuperscript{22} I believe one other focal point to be salient in these conflicts: what can be called the “firing-on-forces” focal point. This is the logic of tripwires. Each side takes what it can without having to overtly attack the other. Although this chapter focuses on geographical sources of vulnerability in red lines, the case studies take up this question directly.
focal point. Vietnam’s deterrent red lines against Chinese occupation of features in the Spratly Islands offer an example. Third, a gray area exists when the focal point (the established border) is imprecise with respect to a particular area. Ethiopia and Eritrea have experienced several crises and land grabs over the Badme region, an area with respect to which colonial-era treaties were ambiguous. Fourth, a gray area exists when a state is imprecise about its territorial claims in its public declarations. I have found this fourth possibility to be rare in territorial crises, unlike the third.23

The explanatory variable gray area is coded as [1] if the contested territory is a small island or border ambiguity, [0] otherwise. The island cases include islands at sea, river islands, and islands in large lakes. Island-like peninsulas also qualify for inclusion. The underlying criterion here is obvious physical geographic separateness from the territorial core. Small islands often fall within border ambiguities, but I include them here and elsewhere only in the island category due to the superior degree of exogeneity of physical geography relative to political geography.

Second, gray areas take the form of preexisting ambiguities in the borders between states. These ambiguities occur due to either inadequate delimitation (lack of agreement on paper) or inadequate demarcation (lack of agreement on the ground as to how the broader agreement on paper applies to actual border markers). To reduce potential endogeneity, this ambiguity must predate the crisis by at least a decade. This arbitrary threshold is almost never determinative – most border ambiguities existed for many decades before the crises under examination, frequently originating in colonial-era boundaries.

Disputed borders, even those over ill-conceived colonial borders drawn with little regard for local demographics, are not in themselves ambiguous. All of the territories in the data were disputed, but only some were ambiguous. In the Siachen example discussed previously (Chapter One), for instance, all of Kashmir was disputed, but the fait accompli struck the one small gray area within that larger disputed region. Some of history’s most deeply resented borders were nonetheless precisely drawn and, I argue, less prone to land grabs because of it. France abhorred the inclusion of Alsace and Lorraine in Germany following the Franco-Prussian War, but German sovereignty over those territories was unambiguous. Similarly, Somalia never accepted the line serving as its border with Ethiopia, but the line was clearly-drawn nonetheless. This distinction between an ambiguous border and a disputed border is essential in reducing the potential endogeneity problem that states might emphasize or invent ambiguities when they wish to legitimize their claim to a desired piece of territory. I return to this potential problem in the next section.

I do not include in the gray area variable two rare types of gray areas in territorial red lines found in the data. The first are territorial enclaves, i.e., West Berlin. The second are overseas colonies

23 I have observed that ambiguity in deterrence overwhelmingly takes the form of ambiguity about the punishment if a red line is violated, not ambiguity about the red line itself. For example, when U.S. President Barack Obama declared, “There will be costs” for a Russian annexation of the Crimean Peninsula, he was relatively clear about the red line but quite ambiguous about the punishment for its violation.

24 Crises tend to arise only after states make conflicting territorial claims.
like Goa (for Portugal), Western Sahara (for Spain), West Irian (for the Netherlands). I leave these out because military indefensibility offers a powerful alternative explanation for why land grabs might be more common in these cases. These are sets of cases in which the adversary tends to have far greater access to the disputed territory than the deterrer due to geographic adjacency/proximity. Islands and ambiguous regions in land borders, in contrast, provide no consistent military advantage to challengers over deterrers. There are, in any case, only a handful of enclave and colony cases among territorial crises since 1918. The results are robust to including them as gray areas.

Land grabs are coded as a dichotomous [0,1] variable, where [1] corresponds to at least one land grab striking the crisis actor. Within any single crisis, I do not count land grabs or weight them by the size of the territory seized. The variable therefore does not distinguish, for instance, between seizing one island and seizing a group of islands within one crisis.

In territorial conflicts, the fait accompli strategy takes the form of the land grab. However, not all land grabs are faits accomplis. The definition of fait accompli requires that a land grab occur with a particular strategic logic: “in an attempt to get away with that gain when the adversary chooses to relent rather than escalate in retaliation.” Of necessity, the measurement of land grabs does not take into account the strategic intent of the challenger. The connection between land grabs and faits accomplis therefore requires an assumption about strategic intent. It seems reasonable to suppose that states generally prefer to make a gain without war if possible. These states presumably narrowed their aims by seizing only a limited territory in order to do so. If so, the assumption is fulfilled. Nonetheless, this assumption may not be correct for every land grab. I return to this point below in the context of the minority of land grabs that strike non-gray area territories (and, usually, provoke war). Some of these cases may be land grabs but not faits accomplis because the challenger embarked on the land grab knowing and accepting that war would ensue.

A case culminates in war [1] rather than peace [0] if the ICB dataset codes that case as a “full-scale war.” “Serious clashes” or “minor clashes” do not make the cut. This choice is necessary to minimize the possibility of tautology. Because many land grabs involve limited acts of violence, tautology would afflic a model in which land grabs are the explanatory variable and outcomes including limited force as war the dependent variable. To minimize tautology, war is defined strictly as the highest level of violence among the options in the ICB data.

The control variables fall into three categories: power, possession of nuclear weapons, and regime type. Data for each of these comes from the ICB dataset in order to maintain comparability with other ICB-based studies. Where necessary, I filled in gaps in by collecting the missing data. For each

---

25 Faits accomplis take many other forms in other issue areas.
26 Only data on adversary characteristics were missing. This happened when the adversary was not included as a crisis actor because, as the aggressor, it experienced no threat to its basic values.
control variable, the data includes the value for the crisis actor (the deterrer), its primary adversary (the challenger), and an interaction between them.\textsuperscript{27}

The power variable codes for power status as follows: small power (1), middle power (2), great power (3), and superpower (4). The interaction term indicates whether both are great powers, including superpowers. Nuclear weapons possession is a dichotomous variable for each crisis actor. An interaction term expresses whether both possess these weapons. The regime type variable records whether each crisis actor is a democracy or a dictatorship in a dichotomous variable.\textsuperscript{28} An interaction term marks cases of joint democracy.

**Gray Areas and Land Grabs**

Two-thirds of all land grabs since 1918 struck gray areas, and a territory disputed in a crisis is twice as likely to be seized in a land grab if that territory is a gray area. These results support the first hypothesis: gray areas make land grabs more likely. This section reviews the evidence on these two findings.

Of the 66 distinct land grab events in Table 4.2, 24 struck islands or island-like peninsulas and 20 struck border ambiguities, for a total of 44 (67%). This tendency is particularly pronounced since the early 1980s; almost every land grab since then has targeted a gray area.\textsuperscript{29} Table 4.3 summarizes the overall pattern in territorial crises between 1918 and 2007. Note the change in scope from all land grabs (Table 4.2) to all territorial crises with or without a land grab (Table 4.3).

\textsuperscript{27} It was usually straightforward to assign the primary adversary as the state with which the crisis actor has the territorial dispute. Where multiple territorial disputes collide (e.g., Israel vs. Egypt and Syria in 1973), I selected the more powerful adversary.

\textsuperscript{28} The ICB dataset divides states into five categories: 1) democratic, 2) civil authoritarian, 3) military-direct rule, 4) military-indirect rule, military-dual authority. I recode the four variants of autocracy as dictatorships.

\textsuperscript{29} The reason for this change remains elusive, but it may be concomitant with the general decline in the rate of interstate war over this period. Because these land grabs so often lead to war, the same explanations for why states have shown less inclination to fight each other over time may apply to an increased reticence towards war-prone land grabs.
### Table 4.3: Gray Areas and Land Grabs in Territorial Crises

<table>
<thead>
<tr>
<th></th>
<th>No Land Grab</th>
<th>Land Grab</th>
<th>% Land Grab</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Gray Area</td>
<td>50</td>
<td>19</td>
<td>28%</td>
</tr>
<tr>
<td>Gray Area</td>
<td>22</td>
<td>40</td>
<td>65%</td>
</tr>
<tr>
<td>% Gray Area</td>
<td>31%</td>
<td>68%</td>
<td></td>
</tr>
</tbody>
</table>
Since 1918, states have rarely tried to slice off even a very small piece of what is clearly another state’s core territory. This lack of land grabs assailing strong red lines comes despite the existence of vast numbers of small regions that could be seized in this way. Pakistan’s 1999 operation seizing Kargil despite overtly crossing the Line of Control exemplifies what could happen frequently yet hardly ever does. The hypotheses imply that operations like this are likely to backfire and therefore unlikely to take place.

Statistical analysis confirms what Table 4.3 suggests: the existence of a strongly positive correlation between gray areas and land grabs in territorial crises. Table 4.4 presents coefficients, standard errors, and significance levels from logit models with standard errors clustered by crisis.
Table 4.4: Regression Results: Gray Areas and Land grabs

<table>
<thead>
<tr>
<th></th>
<th>(1) Bivariate</th>
<th>(2) Deterrer Characteristics</th>
<th>(3) Dyadic Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gray Area</strong></td>
<td>1.58</td>
<td>1.45</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>(.46)***</td>
<td>(.46)***</td>
<td>(.51)**</td>
</tr>
<tr>
<td><strong>Power</strong></td>
<td>-.00</td>
<td>-.01</td>
<td>(.25)</td>
</tr>
<tr>
<td></td>
<td>(.25)</td>
<td>(.33)</td>
<td></td>
</tr>
<tr>
<td><strong>Challenger Power</strong></td>
<td>-.26</td>
<td></td>
<td>(.27)</td>
</tr>
<tr>
<td><strong>Both Great Powers</strong></td>
<td>1.71</td>
<td></td>
<td>(.90)*</td>
</tr>
<tr>
<td><strong>Nuclear</strong></td>
<td>.05</td>
<td>-1.16</td>
<td>(.66)</td>
</tr>
<tr>
<td></td>
<td>(.66)</td>
<td>(.86)</td>
<td></td>
</tr>
<tr>
<td><strong>Challenger Nuclear</strong></td>
<td>-2.46</td>
<td></td>
<td>(1.19)**</td>
</tr>
<tr>
<td><strong>Both Nuclear</strong></td>
<td>3.8</td>
<td></td>
<td>(1.94)*</td>
</tr>
<tr>
<td><strong>Dictatorship</strong></td>
<td>.06</td>
<td>.03</td>
<td>(.45)</td>
</tr>
<tr>
<td></td>
<td>(.45)</td>
<td>(.58)</td>
<td></td>
</tr>
<tr>
<td><strong>Challenger Dictatorship</strong></td>
<td>-1.30</td>
<td></td>
<td>(.71)*</td>
</tr>
<tr>
<td><strong>Both Democracies</strong></td>
<td>-.43</td>
<td></td>
<td>(1.17)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-.86</td>
<td>-.90</td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>(.29)***</td>
<td>(.61)</td>
<td>(1.27)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>132</td>
<td>127</td>
<td>127</td>
</tr>
</tbody>
</table>

* p<.10; ** p<.05; *** p<.01
Figure 4.2: The Effect of Gray Areas

Results with 95% confidence intervals
In Table 4.4, Column 1 presents the relationship between gray areas and land grabs in isolation. Column 2 adds all of the monadic control variables, i.e. the characteristics of the crisis actor (the deterrer). Column 3 adds these same control variables for the crisis actor’s primary adversary (the challenger) along with interaction terms between the two. The correlation between gray areas and land grabs remains similar, substantively strong, and statistically significant across the three models.

Figure 4.2 illustrates the main results of each model more intuitively using predicted probabilities. It displays the substantive effect of changing the strength of the deterrent red line by moving from a non-gray-area territory to a gray area on the predicted probability of a land grab, along with a 95% confidence interval. All other variables are held at their median values. The three results correspond to the statistical models in each column in Table 4.4. As in Table 4.4, the results are strongly significant for the first two models. Particularly striking is the substantive effect of a gray area. Moving from a territory that is not a gray area to one that is a gray area shifts the predicted probability of a land grab from 31% to 64% in the Deterrer Characteristics model.30

Including the full array of control variables (the Dyadic Characteristics model) does result in a decrease in the relationship. In that model, a gray area increases the probability of a land grab by 25%. The control variable primarily responsible for the difference is adversary regime type. Most challengers are autocracies, but democratic challengers are particularly likely to commit a land grab (doing so more than 60% of the time) and particularly likely to strike gray areas. There is no certainty as to why this is so, but one possibility is that democracies are following the strategic logic of the theory more consistently than autocracies.

In the study of crises, causal inference is always a challenge. One potential obstacle to inference here is that states claim borders are ambiguous when they want to dispute them. If so, the dispute creates the ambiguity rather than results from it. States that wish to contest a border certainly have an incentive to contrive an ambiguity in the border they are seeking to change. This potential endogeneity is both a measurement problem and an inference problem. Proper measurement minimizes the problem by distinguishing between preexisting ambiguities and contrived claims of ambiguities (e.g., Libya’s claims about the Aouzou Strip in Chad). In many cases in the data, states resent and contest a border without claiming it is poorly delimited or demarcated.

Moreover, even after dropping the ambiguity cases from the data and using the island variable alone, the result is still statistically and substantively significant. Because it is a matter of physical

30 In all models and as discussed previously, I omit crisis actors for the following reasons: 1) secondary crisis actors, 2) intra-war crises, 3) non-territorial issues, 4) extended deterrence, 5) territorial conflicts over the full territory of one of the crisis actors, and 6) pure compellent demands (cases where the adversary already occupies the disputed territory). For the fourth, fifth, and sixth, it is possible to re-run the base model (Column Two in Table 4.4) without each exclusion in turn. In all three instances, the results strengthen slightly. The results are also robust to dropping all retaliatory land grabs, not just the ones already excluded by the pure compellence restriction.
geography rather than political geography, the island cases are a comparatively exogenous source of weakness in territorial deterrent red lines.\textsuperscript{31}

Cases that resulted in land grabs are presumably more likely to enter the data as crises than similar cases without a land grab. These cases often are crises precisely because of the tensions surrounding a land grab. This potential selection effect is important because it means that gray areas likely lead to land grab much less often in general than in crises. Perhaps the vast majority of gray areas do not lead to a land grabs, but most of these cases are excluded from the data. For instance, there are innumerable small islands. Most crises over small islands include a land grabs, but most small islands never become the object of a crisis. Absent further analysis on the effect of gray areas in general deterrence (i.e., not just in crises), the implications of this analysis must be read with this the caveat that the results only apply to territorial crises.

However, this selection effect offers no alternative explanation for the correlation between gray areas and land grabs in crises. Even if the exclusive focus on crises is enough to exaggerate the frequency with which \textit{fait accomplis} occur, it cannot explain why they occur so much more often when the disputed territory is a gray area.

More worrisome threats to valid inference are likely rooted in the representativeness of the cases in which land grabs do not occur. Because land grabs tend to ensure the case enters the data as a crisis, the data should contain most land grabs. It does, indeed, contain most of those I have identified from all other sources. In contrast, crises without land grabs are unlikely to be representative of the broader set of cases in which land grabs do not occur. Although these cases would seem to approximate near-misses for land grabs and therefore offer the best available comparison group, it is possible that they diverge in some important way from the ideal counterfactual.

One alternative to relying on these cases as a comparison group is to evaluate the observed data on land grabs against an ad hoc assumption that seems reasonable: no more than half of the territories that could be seized in a land grab are islands or ambiguities. Although this number is arbitrary, 50\% would seem to represent a gross overestimate of how common these gray areas are relative to the vast number of other border territories. The null hypothesis remains that the strength of the red line makes no difference, which implies that land grabs should strike non-gray territories as often as gray area territories. Per Table 4.2, 67\% of land grabs strike gray areas. Can this be due to chance? Putting these two pieces together, the question is whether significantly more than half of land grabs strike a gray area. A one-sample, 95\% confidence level binomial test using the list of land grabs in Table 4.2 rejects the null hypothesis that 50\% of land grabs strike strong red lines (p<.01).\textsuperscript{32} The result indicates that more than 50\% of land grabs do so. There are two candidate explanations for this result, i.e., for the prevalence of

\footnote{An even more exogenous variant for the island variable would look only at whether any islands fall in waterways between two states (rather than whether a dispute arises over such islands). However, islands are too plentiful for this approach, as using Google Earth to zoom in on any large river or coastal region will confirm.}

\footnote{By analogy, this is the test that one would use to check if flipping a coin produces outcomes balanced evenly between heads and tails.}
gray areas among the set of land grabs. The first is that the assumption is wrong and the majority of pieces of land that could potentially be grabbed are either islands or ambiguities. This seems unlikely. Indeed, less than half of territories contested in an interstate crisis are gray areas (47%, as Table 4.3 shows). The remaining explanation is that land grabs more often strike gray areas.

Overall, the results provide strong confirmation of the pattern observed in the data: gray areas are associated with a sharp increase in the probability of being struck by a land grab in territorial crises from 1918 to 2007.

Gray Areas, Land Grabs, and War

This section evaluates the relationships between gray areas, land grabs, and the onset of war. The basic results are the presence of a correlation between land grabs and war but the absence of a correlation between gray areas and war. However, the main finding comes when assessing the interaction effect between gray areas and land grabs. Land grabs that strike gray areas are correlated with a lower risk of war; states are more able to get away with these land grabs. This finding directly supports Hypothesis Three. Having already shown that states tend to avoid land grabs when confronted with strong red lines, this section shows that when states do proceed with a land grab that violates a strong red line, the result is usually war.

Table 4.5 summarizes the relationships between gray areas, land grabs, and the onset of full-scale wars, with war as the dependent variable. I again use a logit model with standard errors clustered by crisis. The first column evaluates the effect of a gray area in isolation. Due to posttreatment bias once land grabs are included, it is the best model for evaluating that explanatory variable alone. The second column adds land grabs into the model; it enables the evaluation of the average unconditional effect of land grabs on the probability of war. The third column incorporates an interaction term between gray areas and land grabs, which allows for the main result. The fourth column presents the same model with added controls taking into account adversary characteristics.

The results are as follows. The first column shows that there is no statistical association between gray areas and war. I discuss how to interpret this result – and whether it weighs against the theory – later in this section. It is at odds with Hypothesis Four, which predicts that gray areas make war more likely by making land grabs more likely. The second column supports the literature’s existing assumption – and my assertion as well – that *faits accomplis* are a risky crisis tactic strongly associated with an increased likelihood of war. Recall that Van Evera labeled the *fait accompli* as a “halfway step to war.” A land grab increases the probability of war in a territorial crisis from under 10% to approximately 50%.

33 It was necessary to exclude control variables relating to adversary possession of nuclear weapons to avoid multicollinearity.
However, just as the literature has not previously explored when \textit{faits accomplis} are more vs. less likely to occur, it also has not yet explained when \textit{faits accomplis} are more vs. less likely to provoke war. The third and fourth columns contain an interaction term between gray areas and land grabs, which allows me to distinguish the land grabs striking gray areas from land grabs striking other territories. For ease of interpretation, I evaluate the implications using the predicted probabilities in Figure 4.3.
Table 4.5: Gray Areas, Land Grabs and the Likelihood of War

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray Area</td>
<td>-1.16</td>
<td>-1.23</td>
<td>0.20</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>(.62)</td>
<td>(.73)*</td>
<td>(1.26)</td>
<td>(1.08)</td>
</tr>
<tr>
<td>Land Grab Against</td>
<td>3.16</td>
<td>3.60</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.63)***</td>
<td>(.66)***</td>
<td>(.81)***</td>
<td></td>
</tr>
<tr>
<td>Gray Area &amp; Land Grab</td>
<td>-1.68</td>
<td>-2.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.40)*</td>
<td>(1.38)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td>0.17</td>
<td>0.24</td>
<td>0.25</td>
<td>-0.42</td>
</tr>
<tr>
<td></td>
<td>(.30)</td>
<td>(.35)</td>
<td>(.37)</td>
<td>(.46)</td>
</tr>
<tr>
<td>Challenger Power</td>
<td>-0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.42)</td>
<td></td>
</tr>
<tr>
<td>Both Great Powers</td>
<td>1.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.30)</td>
<td></td>
</tr>
<tr>
<td>Nuclear</td>
<td>-1.30</td>
<td>-1.32</td>
<td>-1.27</td>
<td>-1.10</td>
</tr>
<tr>
<td></td>
<td>(1.18)</td>
<td>(1.25)</td>
<td>(1.30)</td>
<td>(1.03)</td>
</tr>
<tr>
<td>Dictatorship</td>
<td>-0.05</td>
<td>-0.20</td>
<td>-0.26</td>
<td>-1.66</td>
</tr>
<tr>
<td></td>
<td>(.43)</td>
<td>(.55)</td>
<td>(.60)</td>
<td>(.95)*</td>
</tr>
<tr>
<td>Challenger Dictatorship</td>
<td></td>
<td></td>
<td></td>
<td>-0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(.81)</td>
</tr>
<tr>
<td>Both Democracies</td>
<td></td>
<td></td>
<td></td>
<td>-4.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.71)**</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.55</td>
<td>-3.12</td>
<td>-3.41</td>
<td>-1.12</td>
</tr>
<tr>
<td></td>
<td>(.72)**</td>
<td>(.90)***</td>
<td>(1.02)***</td>
<td>(1.78)*</td>
</tr>
<tr>
<td>N</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
</tr>
</tbody>
</table>

*p < .10; ** p < .05; *** p < .01
Figure 4.3: Gray Areas, Land Grabs, and the Predicted Probability of War

Results with 95% confidence intervals
Figure 4.3 displays the predicted probability of war for each possible combination of the presence vs. absence of gray areas and land grabs, holding all other variables at their median values. As Figure 4.3 illustrates, the predicted probability that a land grab leads to war approximately doubles from 28% to 60% depending only on whether the seized territory was a gray area. This is the main result of the section. Without a land grab, the predicted probability of war in territorial crises is less than 10%. Land grabs are an essential part of whether territorial crises escalate to war.

Table 4.6 displays this pattern in the data, this time using all land grabs since 1918, i.e., not just those during ICB crises. The result is even stronger in these data. Land grabs striking non-gray territories lead to war 64% of the time, whereas land grabs provoke war much less often – 16% of the time – when the territory is a gray area.
Table 4.6: When Do Land Grabs Lead to War?

<table>
<thead>
<tr>
<th></th>
<th>No War</th>
<th>War</th>
<th>% War</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Gray Area</td>
<td>8</td>
<td>14</td>
<td>64%</td>
</tr>
<tr>
<td>Gray Area</td>
<td>37</td>
<td>7</td>
<td>16%</td>
</tr>
<tr>
<td>% Gray Area</td>
<td>82%</td>
<td>33%</td>
<td></td>
</tr>
</tbody>
</table>
This finding supports either of two possible interpretations of the land grabs that strike strong red lines and lead to war. First, these states may have decided to try their luck with getting away with land grabs against strong red lines in the mistaken hope that the adversary would relent to the loss. This requires some underestimation of the target state’s determination to uphold a strong red line against violations. Pakistan’s operation in Kargil offers a possible example. Alternatively, states initiating land grabs against strong red lines may understand that their actions are likely to lead to war but choose to do so nonetheless. The Iran-Iraq War, the Ogaden War, the Cyprus War, and the Arab-Israeli wars all seem to better fit this latter view. If so, these land grabs may have been decisions to start a war, not fait accompli attempts. A fait accompli is, by definition, a strategic attempt to get away with a limited gain without provoking escalation. The results presented here do not enable determining which interpretation is correct; this requires evidence as to intent. Either way, the results sustain the theory’s prediction that states generally cannot and do not “get away” with land grabs that violate strong red lines.

The eight cases from the first cell in Table 4.6 are the most anomalous from the standpoint of the theory. These are land grabs that strike non-gray area territories without ending in full-scale war by the ICB definition. However, four did result in significant hostilities that are sometimes regarded as war. Three are considered interstate wars in the Correlates of War data: Japan’s 1931 seizure of Manchuria, Armenia-Azerbaijan 1992, and Pakistan’s 1999 operation in Kashmir. The 1932 conflict between Peru and Colombia is sometimes referred to as the Letician War, but the remoteness of the region being fought over limited troop deployments and casualties. Two of the other anomalous cases involve gray areas that do not fall in the island or ambiguity categories: India’s seizure of Goa from Portugal in 1961 and Morocco’s “Green March” into the then-Spanish territory of Western Sahara. The final two cases are Romania and Finland’s decisions to exploit the Russian Civil War to make gains. On the whole, the pattern of land grabs violating strong red lines leading to war is quite evident in the data.

But what to make of the lack of a correlation between gray areas and war? Recall from the discussion of the hypotheses (Chapter Two) that the predicted relationship between red lines and war is...
indeterminate. The first hypothesis conjectures that gray areas make war more likely though one mechanism: making *faits accomplis* more likely. The evidence supports this. Because *faits accomplis* make war more likely, in itself this means gray areas should make war more likely (the fourth hypothesis). However, the third hypothesis posits a different mechanism that cuts in the other direction: *faits accomplis* striking gray areas should be less likely to provoke war. Figure 4.4 uses descriptive statistics to illustrate how the two mechanisms counterbalance each other.

In Figure 4.4, the first column shows the rate at which land grabs strike gray area territories vs. other territories: 63% vs. 28% (as shown in Table 4.3). The second column looks within the land grabs which strike each type of territory and displays how often these land grabs lead to war: 64% for gray areas vs. 16% for other territories (as shown in Table 4.6). Gray area territories are more likely to lead to a land grab, but each land grab that strikes a gray area is less likely to lead to war. Those two mechanisms offset. There is no net-positive effect of gray areas on the likelihood of war, and therefore no strong evidence for Hypothesis Four.
Figure 4.4: The Two Offsetting Mechanisms Connecting Gray Areas and War

<table>
<thead>
<tr>
<th>Probability of a Land Grab</th>
<th>Probability Land Grab → War</th>
<th>Probability of War</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray Area</td>
<td>18%</td>
<td></td>
</tr>
<tr>
<td>No Gray Area</td>
<td>19%</td>
<td></td>
</tr>
</tbody>
</table>
Far from having no relationship with the likelihood of war, gray areas have two potent connections to it. But, because the mechanisms push in opposite directions, there is no net correlation between gray areas and war. These two mechanisms embody the main findings of this chapter: gray areas make land grabs more likely to occur, but gray areas also make land grabs less likely to escalate to war.

**Conclusion**

This chapter has presented four major findings, two of which underscore the importance of gray areas and *faits accomplis* in interstate crises and two of which directly test the theory connecting them. First, the *fait accompli*, not coercing concessions, is the primary way in which states gain territory at the expense of uncooperative adversaries – both in crises and in general. Second, fully two-thirds of these land grabs strike gray areas, especially since the 1980. Putting these findings together, the conclusion is that most territorial gains take the form of land grabs seizing gray area territories.

The chapter formally tested three hypotheses drawn from the theory, finding support for two and explaining why the evidence does not conform to the third. Gray areas in territorial deterrent red lines are associated with an increase in the likelihood of a land grab. This probability doubles, increasing from approximately 30% to around 60%. These land grabs that strike gray areas are also less escalatory, resulting in war approximately 30% of the time vs, a 60% rate for land grabs targeting all other territories. Putting these together, states are more likely to seize gray area territories because they can more often get away with doing so without provoking a war.

Neither gray areas in red lines nor *faits accomplis* are unique to territorial conflict. This chapter tested the theory within territorial crises in order to isolate a pool of comparable cases. In this issue area, both gray areas and *faits accomplis* are relatively objective to measure and unusually exogenous to policy decisions. Territorial crises are also of considerable intrinsic importance, representing both the most common and most war-prone type of interstate crisis. The question nonetheless remains as to whether these results generalize beyond this focus on what are usually peripheral territories. Answering this question for all crises and all issue types is beyond any single book, but the next two chapters take up this challenge by presenting case studies of two of the most important crises of the 20th Century: the Berlin Blockade Crisis and the Cuban Missile Crisis. Both chapters look beyond purely territorial and geographic considerations to examine the sources of strength and weakness in the red lines during these pivotal crises. How do states challenge red lines? When and why did *faits accomplis* occur? By what method – signaling vs. the *fait accompli* did the actors pursue victory? How was war avoided?
The Berlin Blockade Crisis of 1948-1949, best known for the Berlin Airlift, marked the escalation of the Cold War to the level of tension that characterized the subsequent decades, shredding whatever remained of the bonds between the Allied Powers and the illusion that a cooperative world order was possible between them. This chapter takes up the challenge of explaining critical aspects of this crisis.

The chapter is structured as a competition between two broad perspectives on crises, the conventional wisdom and an alternative predicated on the relationship between red line characteristics, *faits accomplis*, and outcomes. As discussed in Chapter Three, the conventional wisdom emphasizes first and foremost two classic structural determinants of advantage in crisis: relative power and relative interest in the stakes. It explains crisis strategy primarily in terms of signaling, with each side striving to convince the other of their willingness to fight (i.e., of their high level of interest in the stakes). The theory of red lines and *faits accomplis* developed in Chapter Two, in contrast, explains crisis actions and outcomes in terms of what states can get away with unilaterally taking. Because strong red lines are the key to deterring these *faits accomplis*, four weaknesses in red lines – arbitrariness, imprecision, incompleteness, and unverifiability – function as powerful determinants of the course of crisis. This chapter will focus on the two that shaped this case: arbitrariness and incompleteness.

This chapter adopts the unique observable implications approach to case studies described in Chapter One. Rather than structure the entire case in the form of a chronological narrative, this approach requires identifying the specific facts within the case that can only be plausibly explained by one particular theory, and not by rival theories. Equally important are the certain observable implications. Certain observable implications are facts that must be observed in the case for a particular theory not to be discredited. The intent of the unique (and certain) observable implications approach is to isolate and fully evaluate the most informative pieces of information within the case rather than risk their getting buried within a traditional historical narrative structured chronologically.

Why, for instance, did the Western Powers believe that forcing transport aircraft through to Berlin would succeed, whereas attempting to force through a convoy of trucks and armed vehicles would end in disaster? Both actions would have the same strategic effect of undermining the Soviet blockade. The military balance was the same either way. Yet, one was believed to be likely to succeed and the other was seen by most in Washington as likely to start a war. Some American policymakers even feared that a truck convoy might cross the border successfully only to be trapped in an impossible position when surrounded by Soviet forces and obstacles that it could only penetrate by attacking. The Soviets could destroy bridges both ahead and behind of the convoy, leaving it trapped in a humiliating position. It is difficult to make sense of this seemingly bizarre scenario except through an understanding of crisis that emphasizes unilateral actions designed to push as far as possible without violating strong red lines, in this case the use of force.
Neither side was willing to fire on the other, but neither was eager to back down. The result was that each side worked around the strengths of the other’s red line to the extent possible. The Soviet Union did so first by blocking transit corridors in its own zone rather than assaulting directly. Rather than try to push through these obstructions, the Western Powers responded by bypassing these obstacles through the skies, where the Soviets would have to fire first to stop them. The cut-off convoy scenario contemplated at length in Washington merely takes this dancing around firing first to its logical extreme. The airlift-vs.-convoy decision was the most significant strategy debate on the Western side of the crisis, and as I will explain, it hinged on an appreciation of strong red lines.

One important limitation of the following analysis is the lack of information on the Soviet side; the evidence regarding the Soviets consists largely of detailed accounts of what American policymakers believed to be true about Soviet thought processes. I deal with this problem not by assuming American perceptions were accurate, but rather by limiting the analysis of the Soviets to explaining American perceptions of likely Soviet actions. These perceptions turn out to provide some of the most informative observable implications within the case.

The chapter is arranged in three parts. Part I provides a minimalist historical overview of the Berlin Blockade Crisis. It does not try to summarize all of the details that will be utilized in later sections, but rather merely to provide the basic framework with which a reader who is relatively unfamiliar with this case can understand the remainder of the chapter. Part II portrays the case with two competing and starkly different narratives, one drawing on the theory of red lines and faits accomplis developed in Chapter Two and another rooted in the conventional wisdom as laid out in Chapter Three. This part of the chapter serves to clarify the differences between these two perspectives on crisis and to underscore how much what I will refer to as “red lines theory” as a shorthand can potentially explain about the Berlin Blockade Crisis. Part III, the majority of the chapter, tests both red lines theory and key elements of the conventional wisdom by identifying and evaluating the unique observable implications and certain observable implications of these theories.

Part I: Historical Overview

The defeat of Nazi Germany in the Second World War left Germany divided into four zones by the occupying powers: Britain, France, the Soviet Union, and the United States. Although each power took charge of managing affairs in its own zone, the intent was to govern Germany jointly and cooperatively. Berlin, despite its location at the center of the Soviet Zone in Eastern Germany, was similarly divided into four sectors, one per occupying power. This arrangement was fraught with difficulties from the start. As the Cold War intensified, the situation began to deteriorate.

---

1 I credit the coining of this term to Barry Posen. I use it not to imply that I have created an entirely new theory, but rather merely because of the need for a concise term to refer to the ideas in Chapter Two.
At a conference in London in February, 1948, the three Western Powers agreed to the fusion of their three zones in Western Germany and the gradual formation of a new, pro-Western German state. An important step in this direction was the creation of a new, separate currency for Western Germany. This “West Mark” would solve the problem that a joint currency partially controlled and printed by the Soviets was susceptible to many forms of economic manipulation and uncontrolled inflation. The Soviets, with the damage wrought by German armies still fresh in their minds, vehemently objected to these developments towards a capitalist West German state.

Against this backdrop, Soviet Marshal Sokolovsky walked out of the Allied Control Council governing Germany on March 20th, 1948. The first new Soviet restrictions on train traffic between the Western Sectors of Berlin and the Western Zones of Germany began on April 1st. These restrictions led to the “Little Lift,” a modest initial airlift intended only to supply the Western garrisons in Berlin, approximately 5,000 American forces, 5,000 additional Americans, and 10,000 British and French personnel. Supplying the more than two million German civilians of the Western Sectors was another matter entirely.²

By the end of June, 1948, the Soviets had withdrawn from the Kommandatura council governing Berlin and severed all road, rail, and river (barge) access to Berlin. In response, the United States and the United Kingdom initiated “Operation Vittles” and “Operation Plainfare” (respectively) to supply the entire population of the three Western Sectors. At first the airlift relied mainly on the C-47 Skytrain and fell well short of the quantities of food and coal needed to supply the Western Sectors. Most of the tonnage lifted into Berlin consisted of coal, not food. Over time, improved procedures, a new airfield in Berlin (Tegel), and hundreds of new C-54 Skymasters turned the tide. At its peak day on Easter, 1949, the airlift brought 12,941 tons on 1,398 flights into Berlin, considerably more than double the daily minimum requirement.³

Despite the gravity of the situation, the morale of the citizens of the Western Sectors of Berlin remained high. The continuing symbolism of the airlift contributed, as did the widely-publicized dropping of candy via parachute to children in the Western Sectors. A common joke in Berlin went: “Things could be worse. Imagine what life would be like if the Americans were running the blockade and the Soviets were attempting to supply us.”⁴

⁴ Haydock, City under Siege, p. 249.
The Berlin Blockade Crisis lasted nearly a year, and consisted of more than just the airlift. There were a number of armed standoffs within the city itself. The existing city government disintegrated and was replaced by two separate city governments claiming control over the whole city but in reality controlling only their own side. Efforts to supply Berlin by train failed at the border, and the option to dispatch an armed convoy to do so was discussed at great length but repeatedly rejected. The Western Powers eventually imposed a “counter-blockade” in the form of economic sanctions against trade with the Soviet Zone in September of 1948.\(^5\)

Negotiations at many levels took place throughout the crisis: in Moscow with Stalin directly, in Berlin among the military governors, and even in the United Nations with third-party mediation. None of these negotiations led anywhere. Instead, the Soviet Union eventually accepted that the airlift’s ability to supply the Western Sectors indefinitely meant that the blockade had failed. This led to an offer to end the blockade in exchange for a lifting of the counter-blockade and a conference of foreign ministers, which both sides understood to be negligible concessions by the Western Powers. The Soviet Union terminated the Berlin Blockade on May 12\(^{th}\), 1949.\(^6\)

The Berlin Blockade Crisis ended in a victory for the Western Powers. The Soviet Union failed to achieve its main goals: disrupting the London Conference arrangements for a unified state in Western Germany, acquiring the Western Sectors of Berlin outright, or acquiring de facto control of those sectors in the form of near-total economic authority over them. What led to this outcome? What explains the tactics adopted during the crisis by each side?

**Part II: Two Narratives of the Berlin Crisis**

Part II presents two alternative narratives for the strategies, actions, and outcome of the Berlin Blockade Crisis. The conventional wisdom offers a compelling story that plausibly explains many facets of the crisis in terms of relative power, relative interest, and signaling resolve. In the next section, I lay out an alternative narrative based on the theory of red lines and *faits accomplis* developed in Chapter Two. Through this lens, the crisis can be understood as a competition in locating weaknesses in the other side’s red lines and implementing *faits accomplis* to exploit them. Part II establishes the plausibility of these two frameworks for understanding crises and lays the groundwork for the theory testing in Part III.

**The Conventional Wisdom Narrative**


\(^6\) The airlift would continue to build up a reserve of supplies until it was phased out starting on August 1\(^{st}\), 1949. Miller, *To Save a City*, pp. 181-184.
The story of the Berlin Crisis can be told in the form of the clash of the two sides’ relative power and relative interest, along with their efforts to signal their resolve to the other side. There is no shortage of rhetoric supporting this view. According to Secretary of State Marshall, “The struggle is still in its political phase and anything which tends to reduce the will to resist in the Western democracies is a loss to us and a gain to the Soviets.” Colonel Frank Howley, commander of the American garrison in Berlin, described the prospects for succeeding with forcing an armored convoy through to Berlin as follows, “Our chances will increase as the willingness of the American people to fight increases, and as our rearmament progresses.” At the level of crisis behavior, this story holds that the Western Powers cultivated a perception of resolve through signals such as the forward deployment of nuclear-capable B-29s to England and the airlift. The Soviets, in turn, did so with actions such as buzzing Western transport aircraft, aggressive actions in Berlin, and the blockade itself. Within these parameters, the conventional wisdom offers a plausible narrative for the course and outcome of the crisis.

The process of evaluating the power and interests of both sides never ceased throughout the crisis, and policymakers saw these considerations as highly significant. The American military was keenly sensitive to what it saw as the dangerous imbalance of conventional power in Europe in favor of the Soviet Union, often finding surprisingly little comfort in the U.S. nuclear monopoly. The superiority of Soviet ground forces in Europe was never in doubt. The United States had only a few divisions in continental Europe, and these forces were dispersed in order to carry out the occupation. Still recovering from World War II and without any contribution from Western Germany, the continental European allies were also unready for war. The United States retained only one well-trained air group in theater. It was equipped with antiquated P-47s, although another group of F-80 jet fighters was deployed during the crisis itself.

Secretary of State Marshall deliberately avoided public statements revealing American military weakness in order to bolster the U.S. bargaining position with regard to Berlin. In a memorandum to Secretary of Defense Forrestal, Marshall requested that the Defense Department not make public the current figures on the military balance in Europe and the projected number of weeks it would take the Red Army to reach the Pyrenees Mountains. He was concerned about “disheartening” the Europeans. The Army drew up plans for this withdrawal to the Pyrenees “based on the concept of retaining a foothold in Western Europe.” The U.S. Embassy in Moscow released an annual assessment of Soviet intentions – ironically, on April 1st of each year – and the 1948 assessment made clear the American belief that the Red Army could seize continental Europe and hold it for at least two years.

---

7 “Marshall to Forrestal” March 23, 1948 [FRUS]. Note: FRUS hereafter refers to documents available from the State Department’s Foreign Relations of the United States series.
8 Haydock, City under Siege, p. 255.
10 Shlaim, The United States and the Berlin Blockade, p. 183.
11 “Marshall to Forrestal” March 23, 1948 [FRUS].
13 Joint Intelligence Committee, American Embassy, Moscow, “Soviet Intentions,” April 1, 1948 [FRUS].
Nonetheless, an the 1949 edition of this assessment, written shortly before the end of the crisis, found some reason for optimism,

While the Soviet Army is probably capable of overrunning continental Europe with the exception of Spain and Portugal and of occupying strategic areas in the Near East, the Kremlin is presumably aware of the difficult transport, logistical and other problems which would result from such an attempt and doubtless realizes it would be military folly.\textsuperscript{14}

Despite the expectation that the Red Army would initially sweep across continental Europe, the general view in Washington was that the war would ultimately go badly for the Soviet Union.\textsuperscript{15} This conclusion was usually justified in terms of the superior military potential of the United States and the difficulties of occupying so much hostile territory. Both the United States and the United Kingdom prepared for mobilization in case the situation continued to worsen.\textsuperscript{16} Left generally unsaid was the expectation that the use of atomic weapons against Soviet industrial centers would tilt the longer-term balance further in favor of the Western Powers.

Western policymakers almost universally believed that capitulating to Soviet pressure in Berlin would embolden the Soviets into seizing or demanding further gains, and these fear was often expressed in reputational terms. Analogies to Munich were commonplace.\textsuperscript{17} Clay stated on April 10\textsuperscript{th}, “We have lost Czechoslovakia. Norway is threatened. We retreat from Berlin. When Berlin falls, western Germany will be next. If we mean ... to hold Europe against communism, we must not budge.” \textsuperscript{18} And on another occasion, “We must say, ... ‘this far you may go and no further.’ There is no middle ground which is not appeasement.”\textsuperscript{19} His counterpart Robert Murphy, the top State Department official in Germany, put it as follows,

[The western presence in Berlin] became a symbol of resistance to eastern expansionism. It is unquestionably an index of our prestige in central and eastern Europe. As far as Germany is concerned, it is a test of US ability in Europe. If we docilely withdraw now, Germany and other Europeans would conclude that our retreat from western Germany is just a question of time. US position in Europe would be gravelly weakened, and like a cat on a sloping tin roof.\textsuperscript{20}

\textsuperscript{14} Joint Intelligence Committee, American Embassy, Moscow, “Soviet Intentions,” April 5, 1949 [FRUS].

\textsuperscript{15} “Kohler to Marshall,” September 28, 1948 [FRUS].

\textsuperscript{16} “Douglas to Marshall, July 26, 1948 [FRUS].

\textsuperscript{17} \textit{E.g.}, see the following memorandum which described a draft British communique as “redolent with appeasement.” “Marshall to Douglas,” Washington, July 21, 1948 [FRUS].

\textsuperscript{18} Lucius D. Clay, \textit{Decision in Germany} (Garden City, N.Y.: Doubleday, 1950), p. 361. This quotation is likely a paraphrase, albeit a fairly accurate one, of what he said at the time.

\textsuperscript{19} These sorts of statements fit both the conventional wisdom and red lines theory. As such, they are not unique observable implications. Shlaim, \textit{The United States and the Berlin Blockade}, p. 125.

\textsuperscript{20} “Murphy to Marshall,” June 26, 1948 [FRUS].
The U.S. Ambassador to the United Kingdom, Lewis Douglas, felt much the same in analyzing the option of withdrawing from Berlin under duress:

“We can abandon Berlin. The effect of this course of action, would, I think, be a calamity of the first order. Western European confidence in us, in the light of our repeated statements that we intend to remain in Berlin, would be so shattered that we would, with reasonable expectancy, progressively lose Western Germany, if not Western Europe.”

The concern lay both with Soviet perceptions of Western weakness and European perceptions of American unreliability. These statements were the rule, not the exception.

The Berlin Blockade Crisis has also been viewed as offering several prototypical examples of crisis signaling. Most well-known was the U.S. decision, at Britain’s request, to deploy two squadrons of nuclear capable B-29 bombers to the United Kingdom. This has been regarded as perhaps the most notable action taken in the crisis, and as an unmistakable signal of resolve. Speaking more generally about the situation, the U.S. Ambassador to the United Kingdom (Douglas) said, “Any evidence of softness on our part may, at this particular juncture when there is evidence of Soviet irresolution, react to our serious disadvantage.”

Certain provocative actions taken by the Soviets in the air corridors, such as buzzing C-54s and firing near them, as well as increasingly assertive actions in Berlin itself fit the mold of a strategy to signal resolve via brinkmanship. The blockade and airlift can themselves be cast in these terms. By 1949, American policymakers became convinced that the Soviets were building their foreign policy around a “peace offensive/war scare” propaganda strategy in which they trumpeted their peaceful intentions while ensuring tensions remained high in an effort to place the onus for those tensions on the United States.

The rhetoric and discussions among American policymakers evince a clear concern both for signaling resolve, for brinksmanship, and for probing resolve. President Truman wrote in his memoirs,

Our position in Berlin was precarious. If we wished to remain there, we would have to make a show of strength. But there was always the risk that Russian reaction might lead to war. We had to face the possibility that Russia might deliberately choose to make Berlin the pretext for war, but a more immediate danger was the risk of a trigger-happy Russian pilot or hotheaded Communist tank commander might create an incident that could ignite the powder keg.

Notably, despite the focus on signaling resolve in the academic literature on crisis bargaining, probing the other side’s resolve appears to be comparably important, at least in this case.

The emphasis placed on the structural variables and strategies expected by the conventional wisdom continued throughout the crisis. Even after the crisis subsided, Secretary of State Acheson wrote that “our real protection against [the renewal of] the blockade is our own and Western European strength we all understand that [the NATO alliance is] more important than Russian promises.”\textsuperscript{25} A CIA analysis written for the Council of Foreign Ministers (CFM) meeting following the lifting of the blockade described the situation even more bluntly,

In the context of the global power situation, the real issue before the CFM is not the settlement of Germany, but the long-term control of German power.... None of the parties to the negotiations, including the unrepresented Germans, will overlook the long-term question of who is going to control German potential and thus hold the balance of power in Europe.\textsuperscript{26}

Several of the textbook examples of crisis signaling come from the Berlin Blockade Crisis, the B-29 deployment most notable among them. By standing firm against Soviet pressure, maintaining the airlift, and sending additional signals of resolve where possible, the conventional wisdom offers a \textit{prima facie} plausible explanation for why the Western Powers prevailed in the crisis.

**The Red Lines and Faits Accomplis Narrative**

Red lines theory offers the framework for a starkly different account of the Berlin Blockade Crisis. Facing a red line protecting Berlin from direct seizure by force, the Soviet Union exploited the incompleteness of that red line by interposing its forces and other barriers into the transit corridors. Absent the airlift, this blockade would have left the Western Powers with a stark choice between risking war by transgressing the focal point of assaulting Soviet forces or backing down (more on this in later sections). Until the blockade, this choice had lain with on the Soviet side. However, the new Soviet red line denying Western land access to Berlin was also incomplete; it was susceptible to flanking by air. In the sky, it was again the Soviet Union which had to cross the firing-on-forces focal point first, an action which carried with it an unacceptable and credible threat of escalation to war. In this crisis, each side played out its unilateral options which exploited vulnerabilities in the other side’s red lines, i.e., skirting the firing-on-forces focal point, while demonstrating great reluctance to transgress the main focal points underlying them. Both sides were willing to fight if key focal points (strong red lines) were crossed even as neither was unconditionally willing to fight for Berlin. Neither side ever convinced the other that they were willing to fight to get their way on Berlin, and neither side focused its efforts on trying to do so.

\textsuperscript{25} “Acheson to the Acting Secretary of State,” June 19, 1949 [FRUS].
Strong red lines rest on focal points that encapsulate many units of value into one larger whole that states can more credibly threaten to defend. Like many crises, the two pivotal focal points in the Berlin Blockade Crisis were geographic borders and the use of force. Geographically, Berlin’s isolation reduced the strength of the Western red line in its defense by injecting a degree of arbitrariness. Focal points matter through an “if not here, where?” logic. In this case, one might have hoped to abandon the Berlin exclave and hold the line at the zonal border. In that sense, Berlin was potentially outside the focal point arising from the division between Eastern and Western Germany, and so it was more vulnerable to seizure than any piece of the Western Zones. The Western Powers had reason to worry about their precarious position in Berlin. Chapter Four underscored this point. U.S. Ambassador to the Soviet Union Walter Bedell Smith evinced a clear sensitivity to red line characteristics in saying,

As to the likelihood of war, there is a real possibility of it in the Berlin situation. If we had no exposed salient like that, but instead a firm continuous line around our zone – a line which the Russians could not cross without the onus of direct aggression, there would be relatively less likelihood of war... 27

Perhaps for that reason, the firing-on-forces focal point proved to be more potent than geographic borders in deterring advances by the other side. This focal point, though not without certain attendant ambiguities, was generally understood to be specific to politically directed force used against opposing forces. Acts of violent crime by servicemen or the deliberate killing of Germans did not meet these criteria. Nor did just any Soviet-caused Western deaths. The airlift cost 72 lives, 31 of them American, some in plane crashes attributable to the Soviet blockade. 28 Yet these deaths held far different meaning and far less chance of provoking immediate escalation than fatalities during a politically-sanctioned violent confrontation. Western troops, not geography, underlay the Western red line in Berlin.

The Berlin Blockade Crisis began with a gradual series of fait accompli. The question is which ones. From the Western perspective, it began with the Soviets interposing their forces to sever the rail, road, and river lines of supply to and from Berlin. From the Soviet perspective, the crisis began earlier in 1948 with the gradual implementation of the so-called “London Conference arrangements,” a series of incremental and unilateral moves to merge the three Western Zones into a sovereign state in Western Germany. One early, important step was the creation of a separate currency for the Western Zones, and it was the implementation of this currency reform as a fait accompli that triggered the blockade. 29 Although U.S. policymakers saw a negotiated solution to the currency issue as achievable, they decided to implement the measures as a fait accompli to ensure complete control and progress toward a

27 “Minutes of the 286th Policy Planning Staff Meeting” [State Department], September 28, 1948 [FRUS].
28 Tunner, Over the Hump, p. 218.
29 American analysts concurred with this view: Division of Research on Europe [State Department], “Soviet Intentions in Berlin,” April 27, 1948.
government in Western Germany. As Pravda wrote on April 1st, 1948, “The division of Germany has become an accomplished fact.”

Vulnerable red lines offer a compelling explanation for these faits accomplis. The Soviet Union lacked one standout focal point on which to entrench a deterrent effort against a state in Western Germany. Possibilities included formal tri-zonal merger, the ratification of a constitution, and a formal sovereignty-engendering termination of the occupation. The availability of multiple focal points means no single line truly possesses the if-not-here-where quality of an ideal red line. However, the larger problem for the Soviets was that this focal point did not encapsulate enough value. Geography and firing on forces each touched on an interest that extended far beyond Germany; both sides would be willing to fight than allow one predation after another against their territory or forces. The same could not be said of the Soviet red lines to deter a Western German state.

The problem with the Western red line protecting Berlin from direct seizure was its incompleteness. In the transit corridors on land the Soviets could unilaterally interpose their forces, flip rail switches under their control, block roads, and by other nonviolent means render the continued supply of the Western Sectors of Berlin impossible. In part by interposing their forces, the Soviets were able to deter the Western Powers from driving through to Berlin on land or using force to open those land supply routes. Limiting their actions to barring access through their own territory allowed the Soviets to flank the incomplete red line against direct assault. One slight deviation from the theory (the sort which tends to happen when theory meets reality) is that the blockade, although a unilateral imposition, did not achieve Soviet objectives directly. It relied on eventual shortages and the prospect of starvation to succeed.

Nonetheless, the blockade was a unilateral move which flanked the Western red line protecting their sectors of Berlin which immediately gave the Soviet Union the upper hand in the crisis. U.S. Ambassador to the United Kingdom Lewis Douglas wrote in a telegram to the State Department, “I am sure you will agree that we should, if possible, avoid a situation where we are forced, say, to withdraw from Berlin or use an armed convoy to remain there.” The Soviet strategy was to create such a situation, and for a time they seemed to have done so. Army Chief of Staff Bradley put it as follows, “At present with our passenger trains completely stopped, Russians in effect have won the first round.”

The blockade was implemented using salami tactics. It began with a demand for additional identification checks and inspections on traffic into Berlin. Clay, like other Westerners, saw these actions as “the first of a series of restrictive measures designed to drive us from Berlin.” Army Chief of

---

30 Therefore, the Western Powers underwent currency negotiations with less than good faith. “Wisner to Lovett,” March 10, 1948 [FRUS].
31 Division of Research on Europe [State Department], “Soviet Intentions in Berlin,” April 27, 1948.
33 Teleconference TT-9341 [Clay, Bradley], April 10, 1948, [Clay Papers].
34 Shlaim, The United States and the Berlin Blockade, p. 122.
Staff Omar Bradley asked, “Will not Russian restrictions be added one by one which eventually would make our position untenable unless we ourselves were prepared to threaten or actually start a war to remove these restrictions.” Western policymakers assumed from the first restrictions that a “creeping blockade” was underway, with the alternative belief that the Soviets only intended a marginal change in the transit procedures all but disregarded from the beginning. This suspicion appears to have been warranted.

A CIA intelligence estimate from the beginning of the crisis describes the Soviet strategy as a series of progressive steps to push the West out of Berlin, consolidate a communist state in East Germany, and then woo the rest of Germany to join it. The estimate concluded, “Although each of these successive steps involves the risk of war in the event of miscalculation of Western resistance or of unforeseen circumstances, each move on the program could be implemented without the application of military force if adroitly made.” Equally significant is that none of these moves required convincing the West via signaling that the Soviet Union was willing to fight if its demands went unfulfilled.

In response to the Soviet blockade, the Western Powers relied primarily on a strikingly analogous flanking measure, the airlift. To stop trucks or trains, the Soviets could interpose barriers and forces that left the Western Powers with the decision to violate the firing-on-forces focal point or relent. To stop the airlift, in contrast, the Soviets would have needed to fire on or otherwise attack Western aircraft, thereby crossing the use-of-force focal point. The Soviet red line barring access to the traditional routes to Berlin was therefore incomplete, effective on land but vulnerable to being circumvented in the skies.

The airlift flanked this incomplete red line. Due to the logistical constraints on it, the airlift was not itself a single fait accompli, but rather a continuing series of small ones in the form of individual C-47s and C-54s. The Soviets could issue threats and demands regarding restrictions on the airlift, and they did so. But, in the end this cheap talk was dismissed because the Soviets would have needed to fire first, and the West did not believe they were willing to start a war by crossing that focal point. As Lucius Clay, the U.S. Military Governor of Germany, wrote early in the crisis, “Overflight privileges have been

---

35 Teleconference TT-9341 (Clay, Bradley), April 10, 1948 [Clay Papers].
36 Ibid., pp. 112, 136; “Acheson to the Acting Secretary of State,” June 5, 1949 [FRUS]; Lucius D. Clay, Decision in Germany (Garden City, NY: Doubleday, 1950), pp. 358-362; Miller, To Save a City, p. 27.
39 As discussed in Chapter Two, red line incompleteness is not entirely distinct from the other three red line characteristics (which are entirely distinct from each other). Incompleteness serves to address the problem of multiple red lines set by the same actor at one time. When a red line is incomplete, states can and usually do set an auxiliary red line to cover that gap. The incompleteness problem is therefore related to the degree to which this auxiliary red line is vulnerable, which flows from any of the other three red line characteristics. As a result, the vulnerability in the Soviet red line can be described in two ways that sound different but are the same from the standpoint of red lines theory: 1) The Soviet red line denying land access to Berlin was incomplete in that it neglected the skies: 2) The Soviet red line denying Western resupply of Berlin was arbitrary in the skies where it lacked the support of the firing-on-forces focal point.
constantly under discussion at Soviet insistence but only actual interference possible would be overt attack. Airport is in our sector and flights could be stopped only by Soviet use of force.” Once the airlift proved capable of sustaining Western Berlin indefinitely, the advantage was back to the Western side.

Extensive negotiations between the Soviet Union and the Western Powers took place at many points in the Berlin Crisis, but the resolution of the Berlin Crisis did not occur through negotiation. Instead, each side implemented the *faits accomplis* that they believed they could impose without provoking an intolerably escalatory response from other side, and from there the chips fell where they may. Once the failure of their unilateral moves became clear, the Soviets ultimately approached the Americans and conveyed their willingness to abandon their efforts and end the crisis in exchange for very little. Secretary of State Dean Acheson saw this as a general characteristic of Soviet foreign policy,

The Soviet authorities are not moved to agreement by negotiation – that is, by a series of mutual concessions calculated to move parties desiring agreement closer to an acceptable one. Theirs is a more primitive form of political method. They cling stubbornly to a position, hoping to force an opponent to accept it. When and if action by the opponent demonstrates the Soviet position to be untenable, they hastily abandon it...

The irony of this description of Soviet policy is that the United States behaved no differently; the United States prevailed in the Berlin Blockade Crisis by using a strategy precisely along these lines.

**Part III: Theory-Testing**

Part III evaluates the most important observable implications found within the case and uses them to test both red lines theory and the conventional wisdom. It begins by mapping the interests of the two sides of the crisis against their willingness to fight across a variety of foreseen contingencies, highlighting the surprisingly poor extent to which greater interest successfully predicts greater willingness to fight. This section makes clear why this case is puzzling for the conventional wisdom; red lines theory helps to fill in the blanks. The next two sections proceed from this basis to evaluate the most significant unique observable implications for red lines theory by examining what Western policymakers expected would happen if they used various policy options at their disposal. Why was an airlift seen as so superior to an armed convoy, and a truck convoy in turn to a train? Both the perceived outcomes and the reasoning behind them shed a great deal of light on the considerations which shaped the actions of Western policymakers. Subsequently, the discussion turns to explaining the outcomes of the major incidents which took place in Berlin itself, and the role of red lines therein.

---

40 “Clay to Bradley,” April 2, 1948 [Clay Papers].
The analysis then shifts towards testing the conventional wisdom, with sections exposing serious problems with claims that the American deployment of nuclear-capable bombers to Britain and Soviet harassment measures such as buzzing the airlift were in fact strong signals of resolve. The following section then asks when and why perceptions did change during the crisis, if these traditional signaling strategies failed to make much difference. Finally, the chapter explores two further certain observable implications which appear to be unmet. The conventional wisdom has difficulty explaining why the Soviets failed to wait for their unexpected nuclear test and the leverage it could provide before ending the crisis. Red lines theory struggles to explain why the Soviets failed to use certain passive interference measures at their disposal, such as jamming electronic communications important to operating the airlift in poor weather. The overall weight of the evidence suggests that red lines theory can explain a great deal about the case, including some key details that fly in the face of what the conventional wisdom would expect.

**National Interests and the Perceived Willingness to Fight: A Puzzle**

This section examines the relationship between the national interests at stake for both sides and the contingencies that would lead each side to go to war. This analysis serves three purposes: 1) it tests the explanatory power of interests as a determinant of crisis outcomes in an unusual way, 2) it reveals the empirical puzzle that a theory of red lines can explain, and 3) it lays the groundwork for the discussion of the most significant unique observable implications in the case.

**THE SOVIET UNION**

The United States held clear and consistent perceptions of how much the Soviet Union valued the major issues in play and of how likely the Soviet Union would be to go to war in response to each of the available strategies from which the United States had to choose. What is surprising, however, is just how poorly the degree to which an action was thought to harm Soviet interests predicts the odds that the Soviets were expected to start a war in reaction to it. More than acquiring the Western Sectors of Berlin, the Soviets were believed to desire a halt in the process of forming a sovereign government in Western Germany. Yet, not only were the Soviets unwilling to go to war over this goal, they were unwilling even to threaten to do so. In contrast, the Soviets were seen by most to be willing to fight if the Western Powers attempted to send an armed convoy through to Berlin despite caring less about Berlin than Western Germany. Why were the Soviets perceived as willing to fight a war to prevent the resupply of Berlin by road, but unwilling to fight for an objective that was of considerably more importance to them? Moreover, why were the Soviets seen as willing to fight to prevent resupply by air? The deleterious effect on Soviet interests was the same either way, but the expected odds of war differed greatly. All of this sums to a puzzle. The Soviets were perceived as unwilling to fight for what they wanted most, as willing to fight for their secondary objective if pursued in one manner, and as unwilling to fight for that secondary objective if pursued in a slightly different manner.
There was a general consensus among Western policymakers that the Soviet Union held two immediate objectives in the Berlin Crisis: 1) terminating, or at least slowing, the formation of government in Western Germany and 2) acquiring the Western Sectors of Berlin or, failing that, an economic stranglehold over it that amounted to control. Over time, the prevailing view in Washington became that the Soviet Union prioritized events in Western Germany over control of the Western Sectors of Berlin. For instance, a December 1948 memorandum from CIA Director R. H. Hillenkoetter to President Truman concluded that the Soviets had given up on leveraging Berlin to influence events in Western Germany and had shifted to the secondary objective of forcing the West out of Berlin or at least gaining major concessions on the status of the Western Sectors of Berlin.

Given the events of the two World Wars, a high level of insecurity and even paranoia in the Soviet Union with regard to renewed German aggression may have been inevitable. Early in the occupation, Soviet authorities protested an American move to teach baseball to German children on the grounds that it was quasi-military training. The potential for a reinvigorated Germany to swing the balance of power in the Cold War also did not escape notice in the Kremlin. The Soviet interest in the formation of the government in Western Germany should be understood in this light.

The Moscow negotiations in August, 1948 helped consolidate Western perceptions of Soviet interests. These negotiations consisted of a series of direct meetings between the Western ambassadors in Moscow and Stalin or Molotov. They produced what seemed to be a deal on Berlin exchanging the lifting of the blockade for the use of the Soviet Zone currency in the Western Sectors of Berlin. Stalin and Molotov had initially sought the explicit inclusion of this concession in the Moscow talks, but were rebuffed. To the surprise of the West, Stalin ultimately relented to its exclusion, resulting in the Moscow Directive. However, this agreement fell apart when it was referred to Berlin to iron out the details, and may never truly have reflected a mutually-acceptable bargain. In Berlin, the Soviet determination to link the end of the blockade to the establishment of a state in Western Germany continued to resurface despite an agreement in principle (the Moscow Directive) that excluded it. Ambassador Smith (Moscow) warned that any agreement could be temporary and the blockade re-imposed in response to movement towards a government for Western Germany. The first

---

44 Haydock, City under Siege, p. 56.
45 Some in the West shared these views to some extent, especially in France. According to George Kennan, “The Germans, from all accounts, are confused, embittered, self-pitying and unregenerate. Western concepts of democracy have only a slender foundation among them. There is a very good prospect that they will move toward a strongly nationalistic and authoritarian form of government.” “Kennan to Marshall and Lovett” [“Policy Questions Concerning a Possible German Settlement”], August 12, 1948 [FRUS]. Also see Department of State Policy Paper, “Security Against Germany,” February, 1948 [FRUS].
explicit Soviet statement linking the continuation of the blockade to the London Conference arrangements came only on July 29th, 1948, courtesy of Clay’s counterpart, Marshal Sokolovsky. Over time, it became clear to Western negotiators that no bargain was obtainable without a concession on the London Conference arrangements.

There was a broad consensus among Western policymakers that the Soviet Union did not want a war. This conclusion was never taken for granted, and analysis of Soviet intentions with regard to the prospects for war continued relentlessly throughout the crisis. According to Clay, “I am still convinced that the Soviets do not want war. However, they know that the Allies also do not want war and they will continue their pressure to the point at which they believe hostilities might occur.” Intelligence assessments concurred with this view throughout the crisis. Responding to a cable from Ambassador Smith in Moscow, Secretary of State Marshall wrote on April 29th, 1948,

We did not have in mind the probability of some Russian counter move in Europe proper since we agree with you that the present indications are that, with the exception of a possible miscalculation in Berlin or Vienna, the Kremlin does not intend to mount any action in Europe proper which would carry the risk of actual hostilities.

These informed observers generally saw two potential paths to war, neither of which entailed a sudden Soviet decision to attack. The first involved a Soviet miscalculation of the West’s willingness to fight, whereas the second involved the potential consequences of trying to force an armed convoy through to Berlin. As I explore further in subsequent sections, the intriguing aspects of these perceptions are the disconnects between what American policymakers thought the Soviets cared about most and what they thought would cause the Soviets to use force. The Soviets would not use force for what they cared about most, halting state formation in Western Germany. The Soviets would use force

---

48 Miller, To Save a City, p. 83.
49 For instance, the following memorandum from the Secretary of the Navy considers three possible scenarios for Soviet Pearl Harbor-like attacks are, including long-range air raids on the Seattle area, a large-scale sabotage campaign, and a set of coordinated submarine strikes on key ports. “Sullivan [Secretary of the Navy] Marshall,” Washington, December 6, 1948 [FRUS].
50 “Clay to the Department of the Army,” July 10, 1948 [FRUS].
52 “Marshall to Smith,” April 29, 1948 [FRUS]. Marshall also stated, “We do not feel ... that the Soviet Government has committed itself so irretrievably to maintain the blockade to preclude the possibility of some face-saving retreat on their part.” Shlaim, The United States and the Berlin Blockade, p. 252. For Smith’s telegram drawing similar conclusions: “Smith to Marshall,” April 26, 1948 [FRUS].
to stop the resupply of Berlin by land, but not by air (more on this below). These perceptions suggest that red lines can trump interests in driving crisis outcomes, at least to a limited extent within the confines of this case.

THE WESTERN POWERS

The Western side of the equation is less puzzling, but still informative. Why was the United States willing to fight a war if the Soviets seized the Western Sectors of Berlin by force, but potentially willing to withdraw under duress were the airlift to fail? The cost of losing Berlin would exist either way, but the intended policy responses differed greatly. And why would a deliberate Soviet attack on the airlift so likely lead to war, whereas successful Soviet passive interference (e.g., jamming or blinding searchlights) that killed as many American airmen would not? Much of the rest of the chapter analyzes these issues, and the remainder of this section lays the foundation for that discussion.

The West’s goals entering the crisis mirrored those of the Soviets. Keeping Western Germany in the Western camp took precedence, followed by the restoration of the economy and self-governance in the three Western Zones as part of the larger effort to stabilize Europe.53 From the European perspective, this goal was intrinsically valuable, while to the United States it had value in large part because it would create a bulwark against Communism. Soviet control of all of Germany was seen as a grave threat.54 Control of the Western Sectors of Berlin was seen as important, but less so than progress in Western Germany.

Available evidence suggests that most Western policymakers expected to use force in response to Soviet uses of force, including limited uses of force in Berlin or the air corridors. Partly for that reason, these policymakers feared the escalation of the crisis into a war throughout the crisis. Reacting to the failure of the Moscow talks and Soviet threats to harass the airlift, Truman wrote in his diary on September 9th, “I have a terrible feeling that we are very close to war. I hope not.”55 An October 6th National Security Council memorandum asks the Joint Chiefs of Staff to submit recommendations on how to respond to aggressive Soviet actions such as an attack on a U.S. transport aircraft. It asks that two types of measures be prepared:

1) those essential for the defense of the personnel and military equipment of U.S. forces and

53 “Jessup to Acheson,” April 19, 1949 [FRUS]; Shlaim, The United States and the Berlin Blockade, pp. 27, 105. One important caveat was the reticence by many in the French government to allow the creation of a centralized German state even just in the Western Zones due to the threat is might again pose. “Clay to the Department of the Army,” November 22, 1948 [Clay Papers].
54 “Department of State Policy Statement,” August 26, 1948 [FRUS].
55 Haydock, City under Siege, pp. 212-213.
2) offensive and retaliatory measures against USSR forces which may be interfering with the airlift to the extent of causing loss of U.S. lives or planes; measures which in the present tense situation might well result in the outbreak of hostilities leading to war.\textsuperscript{56}

Although nonviolent responses to limited Soviet first uses of force (especially against an armed convoy) were discussed, there is no indication that these reflected national policy.\textsuperscript{57} The British and the United States came to agreement on a general formula for the use of force, namely that they would fight if attacked but that they would not unconditionally commit to choosing war over withdrawing from Berlin.\textsuperscript{58}

One of the most intriguing unanswered questions about the Berlin Blockade Crisis is what would have happened had the airlift been unable to sustainably supply the Western Sectors of Berlin. Red lines theory predicts that the Western Powers would ultimately have backed down in the crisis, lacking any remaining viable method of circumventing Soviet red lines and believing that Berlin was not intrinsically worth a world war. It is impossible to know with any certainty what would have happened, but there is suggestive evidence in this direction.

President Truman’s account of the crisis would seem to strongly disagree. Truman announced to his advisors a clear decision to hold firm in Berlin early in the crisis and never wavered from it. On July 19\textsuperscript{th}, Truman wrote in his diary, “We’ll stay in Berlin-come what may” and “I’d made the decision ten days ago to stay in Berlin” (emphasis in original).\textsuperscript{59} A State Department telegram to the London Embassy reports on a meeting with Truman in which he made the decision to stay in Berlin and “to use any means that may be necessary.” The first draft then added “whatever the consequences,” but this line was removed before transmission.\textsuperscript{60} As clear as these facts may seem, the evidence suggests that they indicate a decision to avoid any major concessions to the Soviets at that time rather than a decision to go to war if it was necessary to break the blockade. The Truman Administration never declared (even internally) that it was willing to attack first over Berlin.\textsuperscript{61} Instructions to American forces in Germany were clear: do not initiate the use of force.\textsuperscript{62} Ambassador Smith (Moscow) highlighted his concerns about this absence of an explicit decision over whether to fight or make concessions if the blockade could not be circumvented by nonviolent means:

I have also pointed out that our governments must decide whether we are prepared to deal indefinitely with the situation now existing in Berlin in the event of breakdown of present

\textsuperscript{56} National Security Council Memorandum, “Re: Possible Soviet Interruption of Airlift,” October 6, 1948.
\textsuperscript{57} E.g., Teleconference TT-9286 [Clay, Royall, Bradley, Collins, Wedemeyer], March 31, 1948 [Clay Papers].
\textsuperscript{60} “Marshall to Douglas,” July 20, 1948 [FRUS].
\textsuperscript{61} Shlaim, \textit{The United States and the Berlin Blockade}, p. 363.
\textsuperscript{62} “Marshall and Lovett to Douglas,” April 30, 1948 [FRUS].
conversations, or if our ability to cope indefinitely with the Berlin situation is doubtful, what concessions we would be willing to make to relieve the situation. I have indicated that these concessions might have to be substantial. In none of the replies that I have received has there been an indication that this basic strategic question has been considered and a definite line of action, beyond immediate reference to UNO [United Nations], decided upon. I do not ask to know what this decision is, but it would certainly help my digestion if I knew that it had been taken.63

Many officials in the United States, Britain, and France harbored thinly veiled – and sometimes explicit – desires to extricate themselves from the problematic Berlin situation. Lucius Clay, among the strongest advocates of holding firm in Berlin, wrote on May 15th, 1948 that he understood that the airlift was not a permanent solution, but that it should continue until Western Europe stabilized. At this point, “the political consequences of the abandonment of Berlin would be minimized.”64 Robert Murphy, another of the strongest advocates of staying in Berlin, wrote to Washington of his fears that the “strength of determination” of the United States to stay would wane over time.65 Lewis Douglas, U.S. Ambassador to the United Kingdom, offered suggestions for minimizing the damage of abandoning Berlin to America’s reputation, including formal alliance commitments, a regional pact, increased troop deployments, and evacuating pro-Western Berliners.66 A CIA Report also spoke of countermeasures to mitigate the loss of prestige from backing down.67 British Commandant Edwin Herbert initially predicted Western Powers would be driven from Berlin by October and that “by that time, the people of Berlin will be so fed up with starvation rations that they’ll start rioting.”68 The French, reliant on the Americans to support their sector and even their garrison in Berlin, were widely believed to favor withdrawal throughout the crisis by both the Americans and the British.69

Contrary to some preconceptions about the aggressive military mindset, support in Washington for a withdrawal was strongest within the Department of Defense.70 On July 28th, 1948, Secretary of Defense Forrestal wrote, “It may not be altogether out of the question to consider ... the possibility that some justification might be found for withdrawal of our occupation forces from Berlin without undue loss of prestige.”71 Army Chief of Staff Omar Bradley debated the issue with Clay early in the crisis,

64 “Clay to Gruenther [Secretary of the Joint Staff],” May 15, 1949 [Clay Papers].
65 “Murphy to Marshall,” April 13, 1948 [FRUS].
68 Haydock, City under Siege, p. 175.
69 Dean Acheson, “Memorandum of Conversation,” March 31, 1949 [FRUS]; “Clay to the Department of the Army,” January 15, 1949 [FRUS].
70 Haydock, City under Siege, p. 229; Shlaim, The United States and the Berlin Blockade, pp. 108, 136-137, 176, 224, 267.
saying, “Here we doubt whether our people are prepared to start a war in order to maintain or position in Berlin and Vienna.” Later in the conversation, Bradley asked, “Should we now be planning how ... we might ourselves announce withdrawal and minimize loss of prestige rather than being forced out by threat.”

Bradley also ordered plans to implement a withdrawal (if so ordered) to be drawn up in Germany, and at least one plan for the “Orderly Evacuation of Berlin by Air” did exist. Secretary of the Army Royall called for a “definite decision” on staying in Berlin late in March 1949, shortly before the Soviets backed down, and advocated that this decision be to abandon the city. This skepticism about Berlin within the military leadership remained true even after the successful conclusion of the crisis, with a June (1949) Defense Department report to the National Security Council calling for serious consideration for evacuating Berlin if the blockade were re-imposed.

The counterfactual of what would have happened if the airlift had proved inadequate to sustainably supply the Western Sectors of Berlin can never be known with certainty. Nor can the counterfactual in which the Western Powers had to respond to the Soviet use of force against the airlift or in Berlin itself. Nonetheless, the evidence suggests that the red line against firing on their forces was a stronger determinant of the Western willingness to use force than their interest in retaining the Western Sectors of Berlin. In other words, the Western Powers may have been willing to lose Berlin to non-violent coercive pressure from the blockade but were not willing to abide the use of force by the Soviet Union without responding in kind. There was never any decision or widespread agreement that force should be used if needed to stay in Berlin, but the decision to use force if the Soviets did so was often taken for granted. In this crisis, the red lines seem to have mattered more than the stakes.

Airlift vs. Armed Convoy

Why did most American policymakers believe that the Soviets would ultimately relent to the resupply of Berlin by air but not by road? Why was resupplying Berlin by road so different that it was generally expected to lead to war, a humiliating retreat, or both? Relative power cannot explain this difference, and Soviets interests would be equally harmed by a shipment of food or coal reaching Berlin regardless of how it arrives. Similarly, without a theory of red lines, it is unclear why aerial resupply would send a different signal than road resupply. The airlift-convoy discrepancy provides one of the most telling unique observable implications in this case, because the conventional wisdom has so much trouble explaining it.

72 Teleconference TT-9341 [Clay, Bradley], April 10, 1948, [Clay Papers].
73 Teleconference TT-9287 [Clay, Royall, Bradley], March 31, 1948 [Clay Papers]. Part of this planning involved the number of Germans to be evacuated for fear of Soviet reprisals. “Clay to Wedemeyer,” July 28, 1948 [Clay Papers]; “HQ Dept of the Army from Dir Plans and Operations to EUCOM,” September 20, 1948.
75 The Acting Secretary of Defense, “Report to the National Security Council: Possible U.S. Courses of Action in the Event the USSR Reimposes the Berlin Blockade” June 1, 1949 [FRUS].
The Western Powers saw themselves as having four basic policy options in confronting the blockade: withdrawal, negotiation, airlift, and armed convoy (also referred to as “armored convoy” and “guarded convoy”). Aside from withdrawal, the options were not mutually exclusive. Negotiations continued off-and-on throughout the year-long crisis but played little role in the outcome. With withdrawal as a last resort, the key strategy decision, then, was whether to supplement the airlift with the armed convoy option. This option came in several forms. A May 1948 version of the plan envisioned assembling three divisions, one from each of the three Western Powers occupying Germany, at Helmstedt and informing the Soviets that these forces would proceed to Berlin at a certain time. A June, 1948 version consisted of a constabulary regiment, an engineer battalion, a bridge train, road repair equipment, two hundred trucks, a British infantry battalion, and French tank destroyers. The orders would be to advance, not to fire unless fired upon, and to clear obstacles even under threat of force. Later versions of the plan envisioned a purely American effort due to opposition by the British and French.

The two top American officials in Germany, Military Governor Lucius Clay and top political representative Robert Murphy, both supported the use of an armed convoy, but few others in Washington, London, or Paris shared that view. Around the peak of the crisis, Clay wrote,

I am still convinced that a determined movement of convoys with troop protection would reach Berlin and that such a showing might well prevent rather than build up Soviet pressures which could lead to war. Nevertheless I realize fully the inherent dangers in this proposal since once committed we could not withdraw.

Clay justified this position in part by arguing, “If the USSR does intend war, it is because of a fixed plan. Hostilities will not result because of action on our part to relieve the blockade unless there is such a fixed plan.” Robert Murphy echoed this thinking almost verbatim. Murphy added, “If an unfortunate incident should occur, there would be no good reason to regard it as more than local and not a casus belli.” All of this reasoning is strongly at odds with both red lines theory and the conventional wisdom’s understanding of brinksmanship, both of which allow more scope for aggressive crisis actions to trigger a war that neither side wanted. After the West prevailed in the crisis without resorting to an armed convoy, Clay moved away from his controversial stance, “While at one time I believed movement by surface routes under armed guard would be feasible, I am sure that if there is a re-imposition of the blockade it is to be expected that an attempt to move into Berlin under armed convoy would be met by force.” He added than an airlift was the “only acceptable option” in that circumstance. However,

---

77 Teleconference TT-9768 [Mayo, Pritchard], July 13, 1948, [Clay Papers].
78 “Clay to the Department of the Army,” June 25, 1948 [FRUS].
79 “Clay to the Department of the Army,” July 10, 1948 [FRUS].
81 Robert D. Murphy, “Comments on JCS Analysis,” June 1, 1949 (FRUS).
82 “Clay to Gruenther [Secretary of the Joint Staff],” May 15, 1949 [Clay Papers].
writing years later he returned to his initial view, “I shall always believe that the convoy would have reached Berlin.”

The dominant view in the State Department, the Defense Department, the White House, London, and Paris, however, was that a convoy was unlikely to reach Berlin and quite likely to lead to war. According to Clay, President Truman informed him that he had been open to the convoy option until confronted with the strong unified opposition of the Joint Chiefs of Staff. Truman later described the convoy option as overly risky.

Opposition to an armed convoy was particularly strong in the American military. Secretary of Defense Forrestal opposed an armed convoy because its chances of success were “remote” and because the “distinctly probable consequence of war” would be disastrous given the “inadequacy of United States preparation for global conflict.” For this reason, both the U.S. Department of Defense and the British Government called for military mobilization prior to an armed convoy effort. A report to the National Security Council as the crisis came to a close offers perhaps the clearest verdict on the armed convoy option,

Armed motor convoy ... was thoroughly considered by the Joint Chiefs of Staff shortly after the Berlin blockade was first imposed and the conclusion then reached is still considered sound, that such an attempt would be fraught with the gravest military implications, including the risk war, and would probably prove ineffective even if faced with only passive interference.

The British military informed the U.S. Army that they considered an armed convoy attempt to be “militarily unsound and politically undesirable.” Even Clay’s top subordinate in Berlin, Colonel Frank Howley, responded to a question months into the crisis about how an initial armed convoy attempt would have gone by saying simply, “We would have gotten our asses shot off.” When interviewed decades later, Soviet junior officers who had manned the border posts expressed the belief that they would have fired on what would, to them, have looked like an invasion force.

83 Clay, Decision in Germany, p. 374.
84 Haydock, City under Siege, pp. 144-145.
85 Shlaim, The United States and the Berlin Blockade, p. 264.
86 Ibid., p. 262.
89 Even the phrasing is similar to the conclusions reached a year early. The Acting Secretary of Defense, “Report to the National Security Council: Possible U.S. Courses of Action in the Event the USSR Reimposes the Berlin Blockade,” June 1, 1949 [FRUS].
90 “HQ Dept of the Army from Dir Plans and Operations to EUCOM,” September 10, 1948.
91 Haydock, City under Siege, p. 255.
Other Western policymakers envisioned a different scenario, one in which an armed convoy entered Eastern Germany only to find itself trapped in an impossible position after the Soviets destroyed several of their own zone’s bridges and erected other obstacles. A July, 1948 report to the National Security Council warned that “Soviet passive interference, such as road and bridge obstruction or destruction, could make an armed convoy method abortive.” Potential Soviet responses included destroyed bridges, roadblocks that included tanks and trucks, and other obstacles. Clay’s British counterpart, General Robertson, feared that the Soviets would simply block the road with tanks. This would leave the convoy with no choices beyond violating the use-of-force red line and halting its progress. According to Army Chief of Staff Bradley, the Soviets “could stop you in so many ways short of armed resistance.... A bridge could go out just ahead of you and then another bridge behind you, and you’d be in a hell of a fix.” In response to these concerns, Clay and Murphy increasingly emphasized bridging equipment as part of their armed convoy plans, but there was little confidence in Washington or London that this would solve the problem, e.g., of Soviet tanks blocking the road.

Many Western policymakers took seriously this rather bizarre scenario that leaves a Western convoy stuck in Eastern Germany by the Soviet’s destruction of their own bridges. This predicted outcome for an armed convoy attempt makes sense only if the tactics of both sides are deriving from a desire to maneuver around the firing-on-forces red line without violating it. Taking that tactic, used by both sides, to one plausible logical extreme produces a convoy stuck in the middle of Eastern Germany. It therefore offers a compelling unique observable implication for red lines theory.

The decisive advantage of the airlift over the armed convoy was that it could be accomplished as a fait accompli without needing to cross the red line against attacking Soviet forces deployed to block access to Berlin. Even General Clay, the principal advocate of an armed convoy, saw this advantage keenly in the period in which he had stepped back from his ardent support of the convoy, “It must be remembered that the surface convoy places the responsibility for the first use of force on our hands, whereas the airlift can be interfered with only by aggressive action on the part of the Soviet Government.” At the start of the crisis, however, Clay was far more sanguine than others about the possibility of circumventing this problem with an armed convoy, “I believe if we advised Soviet authorities of our intent to move supplies into Berlin with armed escort and gave 48 hours’ notice, convoy would get through.” This view is at odds with red lines theory, which expects that the Soviets could disregard merely verbal warnings, interpose their forces, and hold fast behind the red line against

94 Teleconference TT-9768 [Mayo, Pritchard], July 13, 1948, [Clay Papers].
95 Shlaim, The United States and the Berlin Blockade, p. 131.
98 Miller, To Save a City, pp. 25-26.
99 “Clay to Gruenther [Secretary of the Joint Staff],” May 15, 1949 [Clay Papers]. This advantage was widely recognized, e.g., here: Teleconference TT-9768 [Mayo, Pritchard], July 13, 1948, [Clay Papers].
a direct attack on their forces. Clay added to his recommendation, “Am sure neither British nor French would join us.” Washington shared their concerns. Clay’s proposals were rejected.

The Soviet Union faced the same problem in confronting the airlift. They could issue declarations that no flights would be permitted, or that they would somehow halt a flight, but in the end the burden was on them to fire or relent. In one incident on May 4th, 1948, a Soviet officer phoned the U.S. Berlin Air Safety Center to inform them that a particular scheduled flight would not be permitted. The flight landed in Berlin that night. The Soviets respected the red line against firing on Western forces. Even though the value of one plane and its crew paled in comparison to the costs of a war, shooting down a plane deliberately could have crossed a focal point and led to that outcome. A State Department Policy Planning Staff report from mid-way through the crisis dealt with the question of how to respond to such an incident. This report recommended that the United States should react to the downing of a U.S. aircraft by demanding an explanation and implementing “defensive measures” (fighter escorts, presumably). Unless the Soviet Union denied responsibility and avoided repeating the aggression, the United States should “assume that there exists a grave risk of imminent war.”

There is no compelling explanation for the sharp difference between the anticipated results of an airlift vs. an armed convoy that does not make use of a theory of red lines. Normally, any action taken in a crisis can be interpreted as a calibrated signal (see Chapter Three for more on this point). Despite signaling’s impressive malleability, however, it is hard to explain the observed difference between the expected outcomes of these two actions with a logic predicated on signaling. Why is one mode of transport so different from another? If anything, the airlift alone would seem to represent the inferior signal of resolve. Convoy advocate Robert Murphy saw it that way, stating that the airlift “carries with it also a confession of inability or unwillingness to enforce a well-earned right of surface passage.”

As the Berlin Blockade Crisis came to a successful conclusion and with the benefit of hindsight, American policymakers embarked on an extensive analysis of what to do if the blockade were to be reimposed. The consensus was to adopt the same strategy predicated on an airlift and eschewing an armed convoy. A National Security Council report on this question regarded an armed convoy probe as “impractical” and a renewed airlift as “the only practicable step short of great risk of hostilities or decision to leave Berlin.” Undersecretary of State James Webb described President Truman’s views as follows, “He agrees that the reinstitution of the air lift is probably the only answer, he would like to

---

100 Teleconference TT-2066 [Clay, Haislip], March 30, 1949 [Clay Papers]. Murphy continued to believe that this approach of announcing a convoy in advance and then dispatching it would succeed throughout the crisis. Robert D. Murphy, “Comments on JCS Analysis,” June 1, 1949 [FRUS].
102 “Report by the [State Department] Policy Planning Staff,” October 1, 1948 [FRUS].
104 The Acting Secretary of Defense, “Report to the National Security Council: Possible U.S. Courses of Action in the Event the USSR Reimposes the Berlin Blockade,” June 1, 1949 [FRUS].
come up with a better answer, and would not be averse to reconsidering the possibility of breaking the blockade if some means of surface transportation showing reasonable possibilities of success could be found."\textsuperscript{105} Truman ordered that, if the blockade was re-imposed, American traffic should only stop when the Soviets emplace a physical barrier or an armed guard. Purely verbal warnings were to be disregarded. However, Truman also ordered that no “show of force” such as an armed convoy then be made if a physical barrier was in place.\textsuperscript{106} This balancing act mirrors the strengths and weaknesses in Soviet red lines, with the United States willing to act unilaterally where it felt it could get away with it but unwilling to directly engage or challenge Soviet forces.

The conventional wisdom expects crisis strategy to revolve around signaling resolve to the adversary while trying to gauge their resolve to ascertain the best possible bargain that can be reached. Western policymakers consistently considered this approach to crisis strategy, and consistently went another way. After replacing Marshall as Secretary of State in January, 1949, Dean Acheson initially floated the proposal for a one or more limited “probes” of Soviet restrictions, making clear that he was sympathetic to the option as a means for “testing out Soviet intentions.”\textsuperscript{107} He was promptly rebuked by his own subordinates, with Undersecretary of State Webb writing, “The Dept supports the above views of the JCS [Joint Chiefs of Staff] on the isolated issue of probing tactics. Our general view is that half-way measures in a matter of this type with the Soviets are likely to be unsuccessful and are, in some respects, even more dangerous than more deliberate courses of action.”\textsuperscript{108} Acheson seems to have been persuaded, saying, “Neither side wishes to be driven by miscalculation to general hostilities or humiliation. Therefore initial moves should not, if it is possible to avoid it, be equivocal – as a small ground probe would be – or reckless – as a massive one would be.”\textsuperscript{109}

The Futility of Attempting to Supply Berlin by Rail

Trains were the first tool used to test the Soviet announcement of new, intrusive inspections on trains to and from Berlin, a measure seen as unacceptable by the Western Powers and as a prelude to an inevitable full blockade.\textsuperscript{110} In response, on April 1\textsuperscript{st} the United States and United Kingdom dispatched five trains with orders to proceed to Berlin. All were repulsed in a humiliating failure, save for one American train which submitted to inspections in contravention of its orders. The commandant of that train was immediately court-marshaled, but was later acquitted on the grounds that his orders to reach

\textsuperscript{105} The Acting Secretary of State, “Possible Courses of Action in Event Berlin Blockade is Renewed,” May 31, 1949 [FRUS].
\textsuperscript{106} “Director of Plans and Operations, US Army to CINCEUR,” June 17, 1949.
\textsuperscript{107} “The United States Delegation at the Council of Foreign Ministers to the Acting Secretary of State,” May 22, 1949 [FRUS]. Also see “Acheson to the Acting Secretary of State,” June 5, 1949 [FRUS].
\textsuperscript{108} “The Acting Secretary of State to Acheson,” May 26, 1949 [FRUS].
\textsuperscript{109} Shlaim, The United States and the Berlin Blockade, p. 301.
\textsuperscript{110} For Clay’s explanation of why the inspections were intolerable: Teleconference TT-9286 [Clay, Royall, Bradley, Collins, Wedemeyer], March 31, 1948 [Clay Papers].
Berlin and to refuse inspections were contradictory.\(^{111}\) After these failures, the option to send further trains largely fell into disfavor, with advocates of a land supply approach consistently preferring a truck convoy. Why the change? Why were trucks any different? It is difficult to imagine why one would send a stronger signal than the other, and from a commitment standpoint the train offers an advantage: trains cannot stop on a dime, so moving a train towards the border at high speed would have the tactical effect of “throwing the steering wheel out the window” in a game of chicken. That is, the Soviet border guards would be the only ones in a position to choose to fire or not to fire, placing the burden on them. So, why were trucks seen as superior?

The pivotal weakness of using trains to assail the blockade was that the Soviets could more easily stop train movements with passive measures that did not require crossing the red lines against attacking Western forces. These Soviet measures to block access and force the burden of firing first on the Western Powers took two forms, neither of which even required directly interposing military forces on the tracks, let alone firing on trains. First, the Soviets announced “technical difficulties” with the main lines of track, requiring repairs. The fact that Soviet trains continued to use one available line of track up to within a few miles of the border puts this statement in the appropriate light, as does the deliberate tearing up of at least one hundred-yard stretch of track. On a critical highway bridge over the Elbe, the Soviets erected barriers while also removing planking and sawing off cross-beams.\(^{112}\) Shooting was not necessary.

Second, the Soviets had control over switches which determined the line of track a train would take, enabling them to shunt an intruding train along a dead-end line of track without firing a shot. The Soviets made full use of this advantage to divert uncooperative American trains, and it was the inability to solve this problem which best explains the shift in Western emphasis from trains to the armed convoy option.\(^{113}\) Soviet control of canal locks enabled them to stop barge traffic in the same way, an action that has received little attention but was of no less importance from a logistical standpoint.\(^{114}\)

American policymakers sought an option that would resupply Berlin while forcing the burden of firing first onto the Soviets. Lucius Clay and others repeatedly proposed announcing to the Soviets that a train would proceed to Berlin at a certain time, placing the burden on the Soviets to stop it, then following through.\(^{115}\) The problem, however, is that a verbal commitment was not the same as an automatic action beyond Western control. The train could still be stopped by the Americans. Soviet soldiers could still have emplaced barriers to block the train’s route, and the train could still have been

---


diverted to a siding near the border. The consideration of this option is strongly suggestive of a search for an equivalent of the airlift on the ground, but the Soviet red line was simply not vulnerable to flanking on the ground like it was in the skies. Although reluctantly acceding to orders that train guards would not fire unless fired upon, Clay ordered a train movement to Berlin along these lines on June 21st, 1948. Predictably, the Soviets prevented the train from proceeding by diverting it down a siding. Eventually, the Soviets attached their own locomotive and sent it back to Western Germany.116

The debate over how to conduct the initial train probe also revealed the keen sensitivity to the use-of-force red line. General Clay, a striking exception to this general approach of working around strong red lines, initially sought to challenge the new Soviet restrictions by sending a train to Berlin defended by armed guards with orders to fire upon any Soviets boarding the train. Clay’s superiors in the Defense Department immediately ordered the train to be delayed due to their grave reservations about this course of action. President Truman, per their advice, instead ordered the train to proceed with the guards ordered to fire only if fired upon.117 Clay was also ordered not to increase the number of guards or their armament, contrary to his intent.118 After repeatedly overruling Clay, the commander on scene, Secretary of the Army Kenneth Royall offered him an apology, “We are sorry that so much chaperonage was necessary but the war danger element made it necessary.”119

The ability of the Soviets to shunt the train to a siding and leave it there until it became more cooperative may have rendered some of this debate about rules of engagement irrelevant. More generally, the events surrounding the failed attempt to force trains through to Berlin demonstrate the power of the use-of-force red line. The United States anticipated and endured a significant tactical setback by failing in its efforts to uphold a train line of communications with Berlin in order to respect this red line. Clay’s suggestion that train guards might fire first under certain circumstances was met with nearly universal hostility in Washington. Perhaps most informative is the importance of something as small as which side controlled the switches able to divert trains in determining whether or not trains would reach Berlin. At a minimum, this seemingly trivial consideration sufficed to shift the debate away from trains and toward the two core options: airlift and armed convoy.

The Crisis in Berlin

Although the border between the Eastern and Western Sectors of Berlin did not prove inviolable, the pattern of events shows that it provided clear advantages to the force aiming to hold

118 Shlaim, The United States and the Berlin Blockade, p. 127.
119 Teleconference TT-9287 [Clay, Royall, Bradley, Wedemeyer], March 31, 1948 [Clay Papers].
ground on its side of the line. A series of incidents in Berlin demonstrate both the power of borders and the degree to which each side maneuvered around and the red line of deliberately using force against the other side. This section 1) reviews several of the most prominent border incidents in Berlin during the crisis; 2) considers the process through which the occupying powers disestablished the Berlin city government and infrastructure, replacing it with two rival city administrations; and 3) discusses the pattern of smaller incidents of arrests and harassment in Berlin throughout the crisis. Even at what could be considered the tactical level of the crisis, strong red lines appear to have provided powerful advantages.

Prior to the crisis and despite its division into four sectors, Berlin was intended to function as a single unified city with one city government and relatively free movement of people and goods across the sectors. This flexible status quo ante, while preferable for the daily life of Berliners, left Berlin without a border focal point as stark as would exist with the construction of the Berlin Wall in 1961. The resultant incidents seem to follow a pattern in which two key advantages determine the outcomes: 1) the other side must use force first and 2) the incident was on one’s own side of the sector border. Although the sector border did not completely prevent incursions, the home side consistently out-escalated the other in terms of the number of troops deployed and the (perceived) willingness to use them – this despite the undisputed ability of the Soviets to ultimately out-escalate the Western Powers in Berlin and in Germany.

When the Soviets began imposing restrictions on rail traffic starting April 1st, 1948, the Central Railway Administration in Berlin naturally became an object of dispute. Up to this point, it had been controlled by the Soviets despite being located in the U.S. Sector. Perhaps fearing trouble, the Soviets stationed armed guards outside the building. Like the Soviets confronting the airlift, the United States opted not to take the building by force. Instead, American forces surrounded the building to deny entry. The Soviets dispatched reinforcements, but the American MPs held firm. Eventually, the Soviets withdrew, first the reinforcements and then the guards inside the building on April 4th. In a similar if smaller incident on April 1st, a squad of Soviet soldiers advanced 100 yards into the British Sector and established a roadblock. The British surrounded them on three sides with hundreds of troops, and the Soviet squad took the open line of retreat back to their own sector.

On August 14th, the Soviets raided a large black market involved in smuggling supplies into the Western Sectors of Berlin on Potsdamer Platz, located at the intersection of the American, British, and Soviet sectors. The Soviet raid crossed the sector border into Western territory. German civilians at the scene responded with stones, leading Soviet-sector police to fire into the crowd, wounding six Germans. Both sides reacted with reinforcements, and the ensuing standoff eventually ended with the Soviets withdrawing to their sector. The next day, the Soviets again sent hundreds of police and soldiers to the market. The Western Powers were prepared this time, directly interposing their forces along the

120 Clay, *Decision in Germany*, p. 373; Haydock, *City under Siege*, p. 127.
121 Haydock, *City under Siege*, p. 127.
border. The United States deployed a full battalion of MPs, and the British did the same plus barricades. The Soviets stood down.122

On December 16th, 1948, French forces surrounded two radio towers in their sector used by a Soviet-controlled station, sealed them off, evacuated the staff, and dynamited them. This was done primarily to clear flight lanes for the newly-built Tegel Airfield. Although this action infuriated the Soviets, reportedly sending Soviet NKVD Colonel Sergei Tulpanov into a rage strong enough to cause a gallstone attack, the Soviets did not directly retaliate.123 No Soviet personnel were killed, and no Soviet border was violated. Perhaps perceiving the crisis more in terms of reciprocity than red lines, a State Department telegram had warned in October, “We strongly advise against any direct action now interfering with Berlin radio station since this would be sure to provoke violent Soviet action.”124 This prediction turned out to be unduly pessimistic, as red lines theory would expect.

Although most of the major incidents in Berlin took the form of Soviet incursions into the Western Sectors, perhaps the single gravest incident occurred in the Soviet Sector, where the Berlin City Government was located. Freely elected city governments would favor the Western Powers despite rigged electoral victories for the communists in the Soviet Sector. As the Moscow negotiations came to naught, the Soviets moved to consolidate an entirely communist city government presiding over their sector. Rather than move directly, the Soviets orchestrated the storming of the building by a communist mob. Western officials barricaded themselves in their offices. Soviet agents broke into the U.S. offices, beating some Americans there, including reporters, but killing no one. Pro-Western Germans found there were arrested. Eventually, the French struck a deal with the Soviets to evacuate by truck. After three blocks, the Soviets ordered the trucks to pull over on the Schloss Bridge, surrounded the trucks, and proceeded to arrest all the pro-Western Germans. The trucks were then allowed to proceed. In the aftermath, French commandant Major General Jean Ganeval wrote to Soviet General Kotikov, “You and I reached an agreement. I did not doubt your word for a moment and issued my orders accordingly. I still cannot believe that an agreement personally guaranteed by you could have been violated in so flagrant a form.”125 The incident was a clear tactical victory for the Soviets.

In a revealing exchange, Clay declared on September 6th to a concerned Under-Secretary of Defense William Draper that he would send U.S. military police into the Soviet Sector to prevent exactly this sort of mob takeover of the city government. This ambitious approach would seem to contradict the expected deterrent advantage that the border offers to the Soviets per red lines theory. However, in a September 8th follow-up discussion that included Army Chief of Staff Bradley and Secretary Royall, Clay was urged to avoid that action if at all possible. Clay then backed away from his earlier comments,

122 Ibid., pp. 191-192.
123 Clay, Decision in Germany, p. 384; Miller, To Save a City, pp. 113-114.
124 It is not certain that this telegram envisioned disrupting Soviet-controlled radio in the same way that the French went about it, but it is hard to imagine that it referred to a more aggressive approach. “The Acting Secretary of State to the Embassy in France,” October 2, 1948 [FRUS].
125 Haydock, City under Siege, pp. 206-208; “Murphy to Marshall,” September 7, 1948 [FRUS].
saying, “Sometimes I let off steam to Draper which I can’t let off here and tell him what I would like to do but am not doing.”

Most of the actions taken to divide Berlin into two cities were conducted unilaterally and elicited little more than pro forma protests. Each of the occupying powers was able to restructure the utilities and governance of its sector. This involved sealing off, re-routing, and often rebuilding key parts of the electrical, communication, transportation, and sanitation infrastructure, all with the goal of providing no aid to the other side. For the Soviets, this was part of the blockade; for the Western Powers, retaliation. In one incident, U.S. forces cut a gas main to Soviet Marshal Sokolovsky’s house, leading him to move rather than deal with them about it. In the course of that move, a truck carrying some of his furniture attempted to pass through the American Sector, where it was seized. Police forces, media, labor unions, currencies and the city government all divided in two.

Arguably, the Soviet faits accomplis that led to the administrative and economic division of Berlin backfired for the Soviet Union. On May 2nd, Clay argued that the use of the Western German currency or a separate third currency in Berlin would be “most difficult and probably untenable in the long run” due to the importance of the links between the Western Sectors of Berlin and the surrounding areas. Over time, however, the costs of restructuring the Western Sectors around the loss of those linkages became an accepted reality. Once the Western Sectors held what amounted to a separate election late in 1948, re-integrating the city would amount to betraying the pro-Western political parties who were elected, the voters who turned out en masse to elect them, and perhaps even the principle of democracy. This principle would necessarily be disregarded in any plausible compromise with the Soviets, due to Soviet unpopularity. By December, 1948 Clay described a return to joint governance as “virtually impossible.” By 1949, the United States had abandoned its objective of restored joint control of Berlin, with one official writing, “We have now found by bitter experience that direct quadripartite economic operations of this complexity are impossible.” The Soviet faits accomplis succeeded in a tactical sense, but may have resulted in an inadvertent strategic setback to their ambitions for influence in the Western Sectors of Berlin.

---

126 Teleconference TT-1162 [Clay, Draper], September 6, 1948 [Clay Papers]; Teleconference TT-1182 [Clay, Royall, Draper, Bradley, Wedemeyer, Byroade], September 8, 1948 [Clay Papers].
127 Clay, Decision in Germany, p. 365; “Murphy to Marshal,” November 24, 1948 [FRUS].
128 Miller, To Save a City, p. 54.
129 Haydock, City under Siege, pp. 181, 191.
130 For much of the crisis, the CIA shared this Soviet view that these measures would strengthen the Soviet bargaining position. “Hillenkoetter to Truman,” December 10, 1948.
131 “Clay to the Department of the Army, May 2, 1948 [FRUS].
132 “Clay to Draper,” August 19, 1948 [Clay Papers].
133 “Murphy to Marshal,” November 26, 1948 [FRUS].
134 “Clay to the Department of the Army,” December 4, 1948 [Clay Papers].
135 “Adams [Policy Planning Staff] to Jessup,” April 15, 1949 [FRUS].
Although strong red lines appear to have played a powerful part in determining the outcomes of many types of local conflicts within Berlin, some issue areas defied this form of crisis management. Where strong red lines were unavailable, the two sides generally seemed to fall back on tit-for-tat reciprocity.

Soviet forces arrested U.S. servicemen who crossed into the Soviet sector of Berlin almost routinely during the crisis. For instance, at least 24 were arrested between July 15th and September 15th, 1948. Far from being political acts, however, many of these incidents appear to result from the greater availability of alcohol in the Soviet sector, with drunkenness reported in many of the cases even in U.S. sources. Although incidents of long detentions, poor confinement conditions, and beatings did occur on occasion, the norm was for the enlisted soldier in question to be held for about a day in reasonable conditions and released. In addition to detaining Western servicemen, the Soviets arrested several thousand pro-Western Germans, many of them police officers. There were frequent allegations that these arrests took place covertly in the Western Sectors, exploiting this limited element of unverifiability in the Western red lines.

In retaliation for Soviet actions like these detentions, Clay ordered the enforcement of traffic laws on Soviet vehicles long accustomed to speeding through the Western Sectors in total disregard of them. In one resultant incident, Soviet Marshal Sokolovsky’s vehicle was pulled over, leading to a confrontation with his bodyguards that ended when a gun was shoved into Sokolovsky’s chest. In another incident, a Soviet jeep refused to stop for an American MP on a motorcycle, and instead ran him off the road, breaking his arm. Other American soldiers fired on the jeep and pursued it into the Soviet sector before giving up and turning back. These incidents elicited protests but not major escalation.

The latter incident poses the question of whether the results would have been more severe if the firing had killed any Soviet soldiers and/or if the firing had been sanctioned directly by Washington.

The Western powers adopted a general policy of something like tit-for-tat retaliation to minor Soviet acts of harassment that did not cross any strong red lines. Lucius Clay was a strong believer in this approach to crisis strategy,

I know that these measures have the appearance of opéra bouffe and that it does not seem possible that they could take place between the representatives of great nations. However, surrounded in Berlin and subjected to continued and deliberate annoyances, there was no other

---

136 “Hays to Huebner,” July 21, 1949; Haydock, City under Siege, p. 128.
138 Haydock, City under Siege, p. 192.
139 Clay, Decision in Germany, p. 372; Haydock, City under Siege, pp. 146, 200-201.
The countermeasures were effective, and Soviet-created incidents were always reduced when we retaliated.  

Clay reported that in one exchange the Soviets initially refused to budge in response to protests regarding the detainment and treatment of American officers and Western-Berlin policemen in their sector, but that the Soviets became more cooperative when explicitly threatened with added restrictions on access to and transit through the Western Sectors.

A State Department report of April 26th, 1948 listed both a series of Soviet measures short of force to harass the Western position in Berlin and a series of possible retaliatory countermoves. These included barring Soviet publications, evicting Soviet personnel, denying entry, and intensified propaganda. The Western Powers also moved against communists in their zones, arresting some leaders and banning the SED, the Soviet-arranged communist party. However, the aforementioned report makes clear that the power of the border as a deterrent loomed large:

Within Berlin, there are few major means available to the US for exerting practical countermoves against the Soviet forces. In any campaign of mutual retaliation the USSR would have the advantage. American retaliatory action would of necessity be confined to specific Soviet installations located in the western sectors, such as the Soviet Zone Railroad Administration Headquarters in the American sector and the Soviet-controlled Berlin Radio studio in the British sector.

Despite these disadvantages, the Western Powers prevailed in the crisis writ large and most of the major incidents in Berlin during the crisis period. They won the day in most of the Berlin incidents not because of greater strength or better signaling in Berlin — they had neither — but because they were more careful than the Soviets to avoid even small-scale fait accompli attempts on the other side of the border in Berlin.

Deploying B-29s to Britain: A Pivotal Signal?

No event taking place within the Berlin Blockade Crisis has received more academic interest than the deployment of two groups (60 planes in total) of nuclear-capable B-29 heavy bombers to

---

140 Clay, *Decision in Germany*, p. 373.
141 “Clay to Draper and Bradley,” September 8, 1948 [Clay Papers].
143 Division of Research on Europe [State Department], “Possible Soviet Measures to Harass the Western Forces in Berlin,” April 26, 1948. For a similar line of reasoning: Teleconference TT-1756, December 29, 1948 [Clay Papers].
This action seems to offer a prototypical signal of resolve. That interpretation has some truth behind it. The intent was to signal, and the deployment was understood not to entail any great military significance in itself. In advocating this action, Clay described the deployment as “essential” and British Foreign Minister Bevin as “highly important.” However, there are several reasons to doubt the significance of the B-29 deployment.

First, the deployment generated no visible change in Soviet policy. The enthusiastic British request for the deployment came from Foreign Minister Bevin on late June 27th, shortly after the full imposition of the blockade, and was approved by President Truman the next day. There was no obvious change in Soviet in the period following the decision to deploy, nor the deployment itself. After many ebbs and flows in the level of tension and a series of (largely fruitless) negotiations undertaken in several different venues, the crisis ended ten months later.

Second, the British were insistent that both they and the Americans lie about the purpose of the deployment, recommending the phraseology of “routine training flights” and publicly denying any link to events in Berlin. This disingenuousness hardly seems consistent a strong signal of resolve. Nor was there any public announcement of the deployment of nuclear weapons (none were deployed), though one could have hoped the Soviets would draw this conclusion nonetheless. The bomber squadrons selected were not those used for the nuclear mission. Some B-29s even seem to have been moved back from Germany to England to reduce vulnerability in the event of Soviet attack.

Third, the timing of the deployment was delayed nearly a month in order to fit nicely alongside the current status of the ever-evolving exchange of diplomatic notes. The bombers did not reach Britain until late July. The delay had to do with a desire to avoid having the deployment appear to be a response to specific Soviet diplomatic notes, stalling implementation until an appropriate lull in the diplomatic correspondence. If this signal was seen as a potentially decisive tool rather than a minor act more on par with yet another diplomatic note, this delay becomes more difficult to explain.

Finally, despite some strongly-worded support for the deployment, nowhere in the documentary record do American policymakers suggest that they expected the B-29 deployment to lead

---

146 “Clay to Bradley,” June 28, 1948; “Message from the British Secretary of State for Foreign Affairs [Bevin],” July 14, 1948 [FRUS].
149 “Marshall to Riddleberger,” July 12, 1948; “Message from the British Secretary of State for Foreign Affairs [Bevin],” July 14, 1948 [FRUS].
to a change in Soviet policy.\textsuperscript{150} In the aftermath of the B-29 deployment, it receives virtually no further discussion. There were no expressions of surprise or disappointment at the lack of impact. The B-29s faded from high-level attention virtually upon arrival.

The limited significance of the B-29 deployment suggests that the pivotal events and strategies during the crisis were less a matter of signaling than the conventional wisdom would expect. The problem with this tactic was as follows: for the B-29 deployment to be a credible signal of resolve, it would need to be true that the Western Powers would have been less likely to send this signal if they were irresolute. However, forward deploying bombers to Britain in no way prevented an eventual decision to abandon Berlin rather than fight for it. Even if the Western Powers had already decided to do just that if necessary, they could still have hoped to bluff the Soviets into backing down first. Therefore, the incentive for the Western Powers was to send the B-29 signal regardless of their true level of resolve, implying in turn that the Soviets could learn little about Western resolve from the deployment. American and British policymakers seem to have recognized this, widely regarding the deployment as a desirable measure with scant prospects of changing Soviet perceptions, eliciting Soviet concessions, or risking war. The Western Powers could have found ways to send stronger signals had they wished to do so, but, in the words of Harry Truman, “This is no time to be juggling an atomic bomb around.”\textsuperscript{151}

**Soviet Harassment of the Airlift: Brinksmanship to Signal Resolve?**

Brinksmanship offers perhaps the most important way in which the conventional wisdom believes that states signal resolve in order to prevail in crisis. By incurring a significant risk of war, a state reveals something about its willingness to fight over that issue. This section explores the extent to which this strategy was in fact used during the crisis. In particular, once the airlift put it in a disadvantageous position, did the Soviet Union use this strategy to attempt to halt or supersede the airlift? Soviet fighter aircraft could, after all, have literally played the game of chicken with the airlift. I show first that U.S. policymakers consistently worried about accidental escalation and various readily-available Soviet brinksmanship options. Most of these options involved harassment of the airlift, such as close buzzing of transport aircraft. However, despite some accounts which blow the Soviet usage of these tactics out of proportion,\textsuperscript{152} the Soviets consistently decided against using assertive brinksmanship measures that would truly have brought the crisis to a head. This signaling strategy was available, but not used.

Western policymakers continually feared that the Soviets would resort to brinksmanship tactics to combat the ongoing success of the airlift and in so doing trigger inadvertent escalation. According to

\textsuperscript{150} It is difficult to prove a negative, but for the same conclusion: Shlaim, *The United States and the Berlin Blockade*, p. 239.

\textsuperscript{151} Shlaim, *The United States and the Berlin Blockade*, p. 255.

Ambassador Walter Bedell Smith (Moscow), “there is of course the possibility of a miscalculation on their part of the probable consequences of harassing actions that they might take.” He went on to predict that the Russians “will harass the airlift; we may occasionally lose a pilot or a plane.” The State Department’s Policy Planning Staff put the problem in more general terms,

Where forces of mutually antagonistic great powers are operating in such close proximity ... there is always a danger of incidents which, although not so intended, would lead directly to military complications.

Intelligence assessments warned of miscalculation on the Soviets’ part leading to overly aggressive actions. General William Tunner, the commander of the airlift, believed that what harassment did take place happened after (and because) the airlift began to demonstrate its sustainable endurance. He described the harassment as follows,

The Russians resorted to many silly and childish stunts in their efforts to harass us. Their first action was to announce that on the morrow they would be flying in formation over Berlin and East Germany, including the corridors ... I was convinced all along that the Russians were bluffing ... I put out orders to all pilots to continue boring ahead and not to pay attention to the Russians if they did show up. The threatened formation never developed.

What harassing measures did the Soviets take? Were they merely “silly and childish,” or did the Soviets incur a meaningful risk of accidental escalation so as to signal resolve. Many accounts of the crisis emphasize the dangerous harassing measures employed by the Soviets, including buzzing transport aircraft with fighters, aerial and AAA target practice near transport aircraft, and the use of barrage balloons. On further inspection, however, Tunner’s characterization was apt; the Soviets never used the type of intentional brinksmanship that Western policymakers feared, at least not on a large scale.

The most notable single incident cited as an instance of brinksmanship-type harassment was the April 5th, 1948 collision of a Soviet Yak-3 fighter with the British Vickers Viking transport aircraft that it was buzzing in the vicinity of Gatow Airfield on the edge of the British Sector of Berlin. This was the sole collision of this sort during the crisis, and seems to offer a textbook example of brinksmanship. However, this incident did not arise out of a pattern of close encounters in which one was a bit too close. There were no other buzzing or alleged buzzing incidents at that stage of the crisis, including the

153 “Minutes of the 286th Policy Planning Staff Meeting,” September 28, 1948 [FRUS]. Also see “Director of Plans and Operations, Department of the Army to Clay,” Cable WX-83789 [date unknown]
155 E.g., Division of Research on Europe [State Department], “Soviet Intentions in Berlin,” April 27, 1948.
156 Tunner, Over the Hump, p. 185.
157 Haydock, City under Siege, pp. 177, 212, 224.
preceding year and the next month. All accounts suggest that Marshal Sokolovsky’s initially appeared shocked and quite uncharacteristically apologetic, although eventually the Soviets would blame the British pilot for the incident.\textsuperscript{158} Given the lack of any pattern of harassment at this time, both the British and the American governments concluded the incident was not ordered by the Soviet Government.\textsuperscript{159}

Intelligence reports and studies by the State Department’s Policy Planning Staff continued to warn of a potential escalation in Soviet buzzing that would create a real risk of collisions and therefore the escalation of the crisis as a whole.\textsuperscript{160} These dangerous buzzing incidents never came. After the collapse of the Moscow Negotiations, the Soviets announced that they would begin large-scale air maneuvers in the Berlin flight corridors in early September, 1948. This announcement inspired a certain amount of alarm that the long-feared buzzing-type brinksmanship campaign was set to start, but it quickly became apparent that the Soviet announcement was a bluff.\textsuperscript{161} Although incidents of buzzing did take place and heavy Soviet fighter activity around the corridors was common,\textsuperscript{162} in the end the Soviets never adopted a policy of playing chicken with Western transport aircraft. One historian’s rare look inside Soviet archives from this period also found no evidence of orders to buzz Western aircraft.\textsuperscript{163}

The Soviets had options beyond buzzing for engaging in brinksmanship with the airlift, including conducting target practice in the vicinity of Western aircraft. There were instances of dry firing passes by Soviet fighters, Soviet fighters firing live ammunition at dummies trailing behind other aircraft, and AAA fire near the corridors.\textsuperscript{164} However, few of these incidents occurred in dangerously close proximity to airlift aircraft. In his memoirs, Clay wrote, “Frequent Soviet warnings of aerial gunnery practice and formation flying in the air corridors did not materialize in threatening form.”\textsuperscript{165} General Tunner regarded the actions that did occur along these lines as “never more than a morale threat.”\textsuperscript{166}

\section*{When and Why Perceptions Changed}

\begin{flushright}
\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{158}Haydock, \textit{City under Siege}, p. 127; Shlaim, \textit{The United States and the Berlin Blockade}, pp. 133-134; Miller, \textit{To Save a City}, p. 25.
\item \textsuperscript{159}“Murphy to Marshall,” April 6, 1948 [FRUS].
\item \textsuperscript{161}“Murphy to Marshall,” September 4, 1948 [FRUS]; Shlaim, \textit{The United States and the Berlin Blockade}, p. 391; Miller, \textit{To Save a City}, pp. 84-85.
\item \textsuperscript{162}“OMGUS to HQ Dept of the Army for Director Intelligence,” July 20, 1948; Haydock, \textit{City under Siege}, pp. 167, 224.
\item \textsuperscript{163}Gobarev, “Soviet Military Plans and Actions during the First Berlin Crisis,” p. 20.
\item \textsuperscript{164}“USAFE to CINCEUR,” April 22, 1948; “OMGUS SGD Hays to HQ Dept of the Army for CSGID,” July 23, 1948; Haydock, \textit{City under Siege}, p. 224; Riddleberger to Saltzman and Hickerson, Telegram #1799 [date unknown].
\item \textsuperscript{165}Clay, \textit{Decision in Germany}, pp. 373-374.
\item \textsuperscript{166}Tunner, \textit{Over the Hump}, p. 185.
\end{itemize}
\end{footnotesize}
\end{flushright}
The lack of compelling evidence that the most notable signals of resolve – the B-29 deployment and buzzing incidents – affected perceptions raises several questions. Why did these signals fail to affect perceptions? Did perceptions change at all during the crisis? Is a blanket dismissal of signaling to cultivate a perception of resolve appropriate, at least with regard to this case? This section explores these questions, starting with the instances in which perceptions did change. By comparing these instances to the seeming insignificance of more traditional signals, it is possible to draw conclusions about when and why actions taken in crisis affect the adversary’s perceptions.

The most important change in perceptions during the crisis regarded the effectiveness of the airlift. Soviet leaders, like most Western policymakers, entered the crisis with the prior belief that an airlift could not sustainably supply the Western Sectors of Berlin. A January 1948 report by the U.S. Army General Staff clearly stated that an airlift to supply the population of the Western Sectors of Berlin was infeasible. According to the U.S. Ambassador to Moscow (Smith), the Soviets also did not think the airlift could prevent starvation in Berlin. A CIA Report of June 30th, 1948, shortly after the full restrictions came into effect, reported that the Soviets had given orders to Eastern Berlin judicial authorities to begin treating the Western Sectors as part of the Soviet Zone because the Western Powers would be gone within three weeks. Secretary of State Marshall wrote on June 27th that “our general estimate is that the current supply situation in Berlin means that the zero hour there will not be reached for two to three weeks.” At that time, British Foreign Minister Ernest Bevin also did not believe the airlift would be enough to supply the city, but that it would at least be “a symbol of our determination” that would buy time for negotiations.

Over time, these views changed, and the airlift became seen able to continue for an indefinite period, albeit at a cost for the United States, Britain, and the economy of Berlin. In a cable from Clay to Army Chief of Staff Bradley on April 1st, 1948, Clay conveys the decision to begin the airlift, saying “I believe [the airlift] will meet our needs for some days.” Although Clay was initially skeptical about the airlift as more than a stopgap measure, but members of his staff began to persuade him that it could transport more material than he had assumed. By April 2nd, Clay’s views had already begun to shift, with a follow-up cable to Bradley speaking of a sustained airlift for a “much longer period.” Clay had initially estimated a ceiling of 700 tons for the airlift, a figure which paled in comparison to the 4,500 estimated requirement for Berlin. Berlin had imported 15,500 tons daily prior to the blockade. By July, the US-UK effort reached 2,250 tons. By Autumn, 1948, the airlift began to exceed the daily

168 “Smith to Marshall,” July 24, 1948 [FRUS].
169 “Hillenkoetter to Truman” [“Russian Directive Indicating that Soviets Intend to Incorporate Berlin into Soviet Zone”], June 30, 1948.
170 “Marshall to Douglas,” June 27, 1948 [FRUS].
172 “Clay to Bradley,” April 1, 1948 [Clay Papers].
173 Haydock, *City under Siege*, p. 167.
174 “Clay to Bradley,” April 2, 1948 [Clay Papers].
requirements, at times surpassing 10,000 tons. Not surprisingly, the success of the airlift produced a change in perceived capability among Westerners and, one can assume, Soviets. Notably, it was perceptions of a very specific logistical capability which enabled an ongoing series of *faits accomplis* that changed, not perceptions of aggregate relative power or willingness to fight.

Perceptions of the logistical potential of an airlift changed for the straightforward reason that the airlift proved itself. The most direct way to signal the strength of a capability is to visibly use that capability to full effect, and the airlift did just that. The Soviets could count the volume of traffic and calculate its implications for themselves. Eventually it became clear that the airlift could continue for years, and the United States began to plan for that contingency. Why, then, did the crisis take so long? To fully convince the Soviets of its effectiveness, the airlift needed to supply Berlin through the winter. Flying conditions were at their worst in the winter in Central Europe, as the Battle of the Bulge illustrated. Winter weather resulted in (by one U.S. estimate) a 40% reduction in airlift operations. Coal consumption also peaked in wintertime, and most of the tonnage flown into Berlin consisted of coal. The Soviets hoped that winter would break the airlift. Making it through the winter successfully was necessary to demonstrate that the airlift could continue indefinitely. The commander of the airlift, General William Tunner, described Soviet perceptions along these lines:

> The Russians had never had an airlift themselves, and they didn’t take ours seriously until it was too late. I have another personal opinion on this. The Russians did not understand instrument flying themselves and therefore did not believe that we could maintain the Airlift during the long European winter.

The Soviet decision to end the blockade came not long after the end of the Winter of 1949.

American perceptions of the likelihood of war generally declined during the crisis as well – why? First, the conventional wisdom’s emphasis on military mobilization finds support in this case. The Western Powers carefully monitored Soviet troop movements and readiness levels for signs of preparations for war, in line with the conventional appreciation of mobilization as an effective signal. Western analysts interpreted the lack of additional Soviet forces in Eastern Germany as a significant indicator that no attack was imminent. The lack of mobilization, however, was a constant that cannot explain reductions in the perceived probability of war.

---

177 “Cable Ref# UAX-1736,” December 16, 1948.
178 Miller, *To Save a City*, p. 86.
180 Tunner, *Over the Hump*, p. 184.
181 Miller, *To Save a City*, pp. 47-48; “USMA Warsaw SGD Betts to HQ Dept of Army for Director of Intelligence,” August 20, 1948.
182 “Summary of the Daily Meeting with the Secretary” [State Department], March 22, 1949 [FRUS].
The most compelling explanation for the reduction in the perceived probability of war over the course of the crisis is that the absence of a Soviet attack led Western policymakers to become increasingly confident that the crisis was not a smokescreen for premeditated aggression.\(^{183}\) This observation may apply to crises quite generally. Lebow (1981) distinguishes traditional coercive crises, which he refers to as “brinksmanship crises” from “justification-of-hostility” crises in which the appearance of coercion is merely a charade used to obscure and legitimize what would otherwise be naked aggression.\(^{184}\) Once crises pass the point at which an adversary merely seeking a contrived justification for war would have initiated hostilities, the other side can become more confident that this adversary is not dead-set on war. An American official at the Warsaw Embassy during the crisis described the change in perceptions over time, “Local rumors of impending war now reach peaks about monthly instead of bimonthly as here-to-fore.”\(^{185}\) The importance of screening out justification-of-hostility crises reappears in the mid-1949 discussion of a potential Soviet re-imposition of the blockade. Several senior officials in Washington explicitly warned that such a move might presage a Soviet decision to start a war while placing as much blame as possible on the West.\(^{186}\) These fears, in turn, would likely have waned as that future crisis continued without the onset of hostilities.

Although useful in partially defusing crisis tensions, this reduction in the perceived likelihood of war does not extend to long-term trust in the adversary’s intentions. George Kennan aptly summarized this problem of conveying cooperative intentions to the Soviets,

Nothing short of complete disarmament, delivery of our air and naval forces to Russia, and resigning of powers of government to American Communists would dent this problem; and even then, ... Moscow would smell a trap and would continue to harbor most baleful misgivings.\(^{187}\)

Reversing the perspectives, Soviet Foreign Minister Molotov cited the recent lenient treatment of Finland as evidence of benign Soviet intentions. The U.S. Ambassador in Moscow (Smith) recorded his comments as follows,

Molotov replied that surely after two years in the Soviet Union I must realize that the Soviet Government does not pursue any aggressive aims in its foreign policy but that its first and foremost desire is to rehabilitate and reconstruct its internal economy.\(^{188}\)

Nothing of the sort would happen for the next forty years.


\(^{185}\) “USMA Warsaw SGD Betts to HQ Dept of Army for Director of Intelligence,” August 20, 1948.

\(^{186}\) The Acting Secretary of Defense, “Report to the National Security Council: Possible U.S. Courses of Action in the Event the USSR Reimposes the Berlin Blockade,” June 1, 1949 [FRUS].


\(^{188}\) “Smith to Marshall,” May 4, 1948 [FRUS].
Why Didn’t the Soviets Wait for a Nuclear Arsenal?

The blockade began in part during March, 1948 and in full by June, 1948. It ended in May, 1949. The Soviet Union tested its first nuclear weapon in August, 1949. If the Soviets believed that Western perceptions of Soviet power were a potentially decisive factor, why didn’t the Soviets continue the crisis until they could exploit this sharp increase in their perceived power? There was no immediate, urgent need to lift the blockade. The situation had stabilized into a stalemate of sorts that, although costly to each side in certain ways, was far from catastrophic. The Soviets were also very likely aware that the Western Powers were underestimating the progress of the Soviet nuclear program, and therefore that a sudden, favorable change in Western perceptions of Soviet power loomed on the horizon. A nuclear test, after all, is a signal of increased power whose credibility is not in doubt; only a nuclear power would be able to send it. So, why not take advantage of it?

If perceived relative power is a decisive advantage in crises, it is difficult to explain why the Soviets chose not to wait and at least see if this sudden increase in perceptions of their military capabilities could bring them victory. More formally, the relative power hypothesis in its stronger forms has a certain observable implication that the Soviets should have waited and exploited this advantage. They did not do so. More information on Soviet strategic thinking would help to elucidate this point, but it is not easy to imagine how it could alter this conclusion. Relative power matters, of course, but this case offers two reasons to view its explanatory power as more circumscribed than the conventional wisdom tends to assume. The first is the decision by the Soviet Union not to wait to lift the blockade until after its nuclear test. The second is the range of possible outcomes – from Soviet victory to Western victory to war – that Western policymakers considered possible despite relative power remaining virtually constant during the critical period of the crisis in mid-1948. Relative power and relative interest alone cannot explain the outcome of the Berlin Crisis.

The Surprising Rarity of Soviet Passive Interference to the Airlift

Among the most surprising aspects of the Berlin case – and the most problematic for red lines theory – is the lack of non-violent Soviet technical measures to disrupt the airlift. These actions avoid the pertinent focal points, direct attack on military forces and border-crossing, and so should have held great appeal for the Soviets. This section asks why the Soviets refrained from jamming western radio communications and radar. One possibility is that the Soviets lacked an understanding of the electronic warfare options available to them. Other explanations such as human error and organizational dysfunction remain possible pending further information on the Soviet side of the crisis. Nonetheless, the rarity of Soviet passive interference with the airlift remains a piece of evidence that potentially calls red lines theory into question to some extent.

---

189 Because the Soviet nuclear test occurred unexpectedly and after the crisis ended, there was no discussion of a possible Soviet nuclear test in the documents on the Berlin Crisis that I have seen.
The Soviet Union did not engage in a sustained effort to jam Western communications in order to disrupt the airlift. Most or all of the jamming activity mentioned in various historical accounts seems to have consisted only of low-level jamming of police radio bands used by German officers in the Western Sectors of Berlin. Little effort was made to jam the various electronic frequencies used to coordinate so many aircraft in the air at once and to conduct instrumental landings in difficult weather conditions. The Soviets did remove Western navigation beacons in Eastern Germany on April 9th, 1948, but could have gone much farther. One historical account suggests that this Soviet restraint may have been motivated by a desire to signal the limits of their intentions and thereby aid in avoiding war, but there is no evidence available to support that interpretation. General Tunner believed that the Soviets may simply have lacked an understanding of instrumental flying, a new technology, and that in any case the airlift would have been able to persevere despite jamming. Red lines theory, however, predicts that they could have gotten away with this action without undue risk because the red line against attacking Western aircraft did not extend to passive electronic interference. It remains somewhat puzzling for the theory that the Soviets did not at least try.

The Soviets had at least two other, higher-risk options for (somewhat) passively interfering with the airlift: high-powered searchlights and barrage balloons. Searchlights were used at times to interfere with landing at Gatow at the outer edge of the British Sector of Berlin. Although occasionally forcing pilots to put up newspapers on the cockpit windows to avoid temporary blinding, this tactic never caused a crash. There is something odd about the fact that taking down an American aircraft by shooting searchlights at it has such different implications as taking a plane down by shooting bullets or shells, but this likely has to do with searchlights falling in a gray area within the in-this-sense-imprecise red line against using force on Western armed forces.

Barrage balloons are the other option that was available to the Soviets. Used during World War II to defend against attack by low-flying enemy aircraft, barrage balloons are stationary balloons tethered to the ground by a cable. The threat to aircraft is not so much the balloon as the cable, which damages aircraft that collide with it. Although there were reports of barrage balloons during the airlift, the Soviets never attempted to use them on a scale that could interfere with the airlift. For instance, on July 2nd, 1948 the Soviets warned the Western Powers that a barrage balloon would go up in the corridor. They gave the time, location, and altitude along with a promise not to place the balloon in  

190 Miller, To Save a City, pp. 26, 126; Shlaim, The United States and the Berlin Blockade, p. 288.
193 I find this explanation unconvincing for two reasons: 1) as U.S. sources suggest, these actions would have been perceived as little more than a logical extension of the blockade that was already underway and 2) these actions would have been less provocative than many other actions taken by the Soviet Union, especially the incidents in Berlin itself. Shlaim, The United States and the Berlin Blockade, p. 288.
194 Tunner, Over the Hump, p. 224.
195 Tunner, Over the Hump, p. 185; Michael D. Haydock, City under Siege, p. 224.
clouds. The balloon was never sighted by a Western aircraft. A retrospective assessment of reports of barrage balloons in the air corridors around June 30th 1948 concluded that only one balloon near Magdeburg was present and that it was “considered of no importance.” These two incidents reflect the minimal extent of the Soviet usage of barrage balloons, a tactic that could be viewed as passive interference or as brinksmanship. Barrage balloons fell into a gray area for the firing-on-forces red line, because they were equipment but not living soldiers. Western policymakers anticipated a more aggressive use of barrage balloons and wrestled with the problem of how to respond. Eventually they gave orders to refrain from shooting down balloons without receiving orders from Washington or London to do so, despite the ambiguity of whether firing on balloons violated the focal point of firing on Soviet forces. It is unclear whether the Western powers would have exploited this imprecise aspect of the Soviet red line to try to get away with shooting down Soviet balloons as a fait accompli if the Soviets had made more widespread use of this tactic.

The absence of these passive interference measures, especially jamming, is puzzling not just for the theory being tested, but also surprised U.S. policymakers. Even in October, months into the airlift, a report by the State Department Policy Planning Staff mentioned continuing fears of “indirect mechanical interference to navigation through such means as jamming of radio control systems, smoke smudges, etcetera.” As expected, this report also recommends that such measures should be met with “immediate vigorous protest” in the form of statements, but the report refrains from endorsing any stronger responses. These fears in Washington and London, at least, meet the expectations of red lines theory. Future research exploiting Russian sources might search for evidence against red lines theory by exploring why the Soviets refrained from widespread passive interference measures that sought to disrupt the airlift without clearly violating the red line against firing on it.

Conclusion

The course of the Berlin Blockade Crisis is best understood as the Soviet Union flanking the incomplete red line protecting Berlin by interposing forces and obstacles to cut the land supply routes. The Western Powers then mirrored this tactic by using the airlift to flank this new – and also incomplete – Soviet red line, taking to the sky to avoid having to dislodge Soviet forces in their path. The auxiliary red lines that might have filled these gaps suffered from arbitrariness. Non-arbitrary red lines use a focal point to encapsulate many individually small valued units into a larger whole that states can more credibly threaten to defend. Red lines against overflying occupied Germany in the pre-arranged flight corridors and against sealing roads into Berlin fall far short of this standard. Importantly, it is not just the focal point itself, but a focal point that encapsulates enough value to be worth defending. In this sense, red lines theory is not entirely divorced from state interests.
remove them from the land routes to Berlin did not. Even the Western Sectors of Berlin writ large fell short of the ideal on this score, being as they were an isolated picked in Eastern Germany. Imprecision and unverifiability featured less prominently in this case, but the next chapter will explain how both played major roles in the Cuban Missile Crisis. Both in the actions they took and in their assessments of what would transpire in contingencies that never came to pass, Western policymakers consistently respected the deterrent power of strong red lines.

Over the course of the chapter, Lucius Clay emerged as the standout exception, a policymaker with little if any concern for strong red lines. His views differed in pronounced ways from those of his superiors in Washington, resulting in his being repeatedly overruled on questions relating to the conduct of U.S. policy during the crisis. Clay was not an outsider or thought to have extreme views in general; for instance, his views on U.S. interests and goals were quite mainstream. His disregard for red lines seems to be a reliable predictor of his disagreements with Washington, most of which concerned questions of crisis strategy. This disregard also led to some erroneous predictions; on April 2nd, 1948, Clay declared to his superiors, “I am sure our position has stopped for time being further interference with air and highway movements which would require force to implement.”

Clay left the Berlin Blockade Crisis with a reputation for strong leadership in a difficult situation. For this reason, he was brought back from retirement during the 1961 Berlin Crisis and charged with command of the U.S. forces in West Berlin (although his authority was vaguely defined). He promptly began a series of provocative actions, including aggressive patrols and sending American officials into the Soviet Zone without submitting to inspection by East Berlin police. His intent was to enforce the right of such access. Clay was the main instigator of the famed tank standoff at Checkpoint Charlie. At one point, Clay ordered a mock Berlin Wall created so American soldiers could practice tearing it down. He also inquired as to Washington’s view of a quick military raid into East Berlin to destroy obstructions and segments of the Wall. These positions were at odds with the majority views in Washington and NATO, leading to Clay’s marginalization, his eventual resignation, and a more conciliatory policy. Whereas most policymakers thought in terms of working around strong red lines, Lucius Clay’s recommendations consistently showed how radically different crisis policy would be if run by those with little concern for them.

It is easy to take the impact of strong red lines for granted. A counterfactual illustrates: what if Western Berlin had been geographically contiguous with Western Germany? Tactically, no blockade would have been possible. The Soviets could have seized the Western Sectors by violating the use-of-force red line, but this they were unwilling to do even with Berlin as an exclave encased by their territory

---

201 Teleconference TT-9300 [Clay, Royall, Bradley], April 2, 1948 [Clay Papers]. He was wrong in part because blocking highway movements did not require force.


203 American and British planners in fact sought a settlement along these lines, but Berlin is too far east for the resultant Soviet Zone to have been of adequate size. Daniel J. Nelson, Wartime Origins of the Berlin Dilemma (University: University of Alabama Press, 1978), pp. 141-142.
and forces. Moreover, if red lines theory is correct, the strategic stakes would have been further tilted in favor of the West. After giving up on Berlin, there would have been no fallback focal point on which to declare, “Here and no further” short of the French and Dutch borders. In contrast, allowing the Soviets to tear off a swathe of territory from a single zonal border would leave no compelling focal point against further Soviet advances. As a result, the incentive for the West to hold would have been stronger, and their deterrent correspondingly enhanced. Walter Bedell Smith, U.S. Ambassador in Moscow for the critical phases of the crisis, clearly recognized the importance of strong red lines in regulating crises:

Although a clear-cut zonal line separating us from them would minimize the danger of armed conflict now, the confused relationships existing in Berlin and the corridors leading to it have present potentialities for incidents that could readily lead to war.  

The arrangements made for Berlin during the Second World War have rightly come under a great deal of criticism, and red lines theory underscores the pitfalls of creating this sort of isolated geographic pocket. However, red lines theory also implies that these wartime planners may deserve a great deal of credit for the decision to effectively pre-partition Germany into occupation zones. This unheralded move may have averted a more severe crisis, and perhaps a war. If the occupying powers had attempted to govern all of Germany in a truly joint fashion, there would have been no focal points to serve as the basis for a partition of Germany that gave each side enough to reduce their fears of the full might of a re-militarized Germany allied to their Cold War adversary. Instead, each side would have jockeyed for total control of the political orientation of the German Government in a situation of even greater uncertainty and consequence. With such high stakes and no strong red lines to step in and shape outcomes when diplomatic negotiations deadlocked, the situation could easily have taken a dire turn.

The Berlin Blockade Crisis galvanized the creation of NATO and accelerated the Cold War. Despite the distinct legal and historical status of defeated, occupied, demilitarized Germany, American policymakers clearly saw a need to incorporate it into one broader whole (NATO) that they could more credibly threaten to protect. By making the Berlin exclave a part of this, the Western Powers increased the costs to themselves of a Soviet move on the Western Sectors. Consequently, they were able to deter Soviet moves against it. The 1961 Berlin Crisis ended with the Soviet Union’s unilateral construction of the Berlin Wall. The West excoriated that action, which fell just outside the red lines protecting West Berlin, for imprisoning thousands of would-be refugees behind the Iron Curtain. But, it also effectively marked the end of the Berlin crises. The Berlin Wall even contributed to further strengthening the focal point protecting Berlin from a more aggressive Soviet fait accompli. From this

---

204 Minutes of the 286th Policy Planning Staff Meeting,” September 28, 1948.

205 This decision was not without its critics. Some State Department officials opposed occupation zones because they were a step towards dismemberment as advocated by those who sought to indefinitely suppress German power. Nelson, Wartime Origins of the Berlin Dilemma, p. 28.

206 Jacob D. Beam, “Set of Principles for Treatment of Western Germany in Event It Is Impossible to Repair the Split of Germany (Revised),” February 29, 1949 [FRUS]; “Lovett to Smith,” April 24, 1948 [FRUS].
point on, the geography of Cold War crises shifted away from the Central European front, most notably to the Caribbean Sea.
Ch. 6 The Cuban Missile Crisis of 1962

The Cuban Missile Crisis is widely regarded as the closest that the world has come to nuclear war. Both sides sought a strategy that would let them get their way without an unacceptable risk of nuclear war. The interplay of these strategies ultimately led to the Soviet withdrawal of missiles from Cuba and the avoidance of war. But what were these strategies, which theoretical framework do they best fit, and why did they produce this outcome?

The prevailing narrative of the Cuban Missile Crisis envisions the case as a clash of wills in which the Soviet Union blinked first. The United States deployed the gamut of traditional crisis signaling tactics – military mobilization, firm statements of resolve, brinksmanship, etc. – to decisive effect. Fearful of war, the Soviet Union turned back its ships and withdrew the missiles. The Cuban Missile Crisis is more than just a source of evidence in favor of the significance of signaling. It is the “home turf” case from which the conventional wisdom on draws so heavily to inspire its emphasis on signaling to coerce concessions.

Yet the story of the Cuban Missile Crisis is not so simple, and there is much in the case that red lines theory best explains. The crisis arose because Soviet leaders saw an opportunity to exploit an unverifiable American red line against deploying missiles to Cuba, which they did in a covert fait accompli. Throughout the case, both sides maneuvered to make unilateral gains without crossing the strong deterrent red line (and focal point) against overtly attacking the other. This finding mirrors the Berlin case, where the airlift proved to be decisively superior to a truck convoy because the Soviets had to fire first to stop planes, but the United States had to push through Soviet forces on land. With respect to Cuba, American policymakers chose the blockade option over airstrike or invasion due to fears that an attack on Cuba would spill Soviet blood and therefore make war too likely. The strategic logic was to work around a line that could not be safely crossed.

The result was a new red line – the blockade line. The United States deterred the Soviet Union from running its blockade line not because it convinced the Soviet Union that violating the line would lead to a U.S. attack, but rather because the United States could disable and/or board a Soviet freighter without having to overtly attack it. The decisive factor was the credible threat to disable a blockade-running ship in a fait accompli that walked right up to the edge of the line of an outright attack without unambiguously violating it.

Red lines theory also aids in understanding the brinksmanship which established the credibility of American coercive threats, in particular the underappreciated importance of intensive American reconnaissance overflights. Although traditional policies intended to signal resolve – mobilization, forward deployment, firm public statements, etc. – were necessary for the success of signaling, American brinksmanship also relied on the potential for reconnaissance overflights to generate shooting incidents. Because Khrushchev distrusted his ability to restrain Cuban and even Soviet anti-aircraft
units, these overflights created the uncontrollable risk of a shoot-down. The loss of one pilot pales in comparison to the casualties from a nuclear war, but that one death would cross the red line of an overt attack. Khrushchev’s fears surrounding these incidents and potential retaliation for them galvanized his decision to withdraw the missiles. This brinksmanship functioned in large part at the intersection of the two theoretical frameworks: signaling resolve and the significance of violations of strong red lines.

Whereas the Berlin case divides neatly into two competing narratives and the evidence favors the narrative from red lines theory, the Cuba case paints a more balanced picture. Certain aspects of the case fit one; some the other. The Soviet decision to withdraw the missiles does provide an example of successfully signaling resolve; over the course of only a few days, the United States persuaded Soviet leaders that it was likely enough to attack Cuba – and potentially start a major war in doing so – that Khrushchev ordered the immediate withdrawal of the missiles. However, red lines theory explains the Soviet decision to deploy the missiles, the American decision for blockade, the way in which that blockade line succeeded in turning back Soviet ships, and – in conjunction with established ideas about brinksmanship – the significance of the American reconnaissance overflights for the eventual outcome.

This chapter employs the same approach to case studies as the Berlin chapter. The analytical leverage comes from identifying unique observable implications and certain observable implications. These are observations – sometimes fine-grained details of the case – that are uniquely explained by one theory or that must be observed for the theory to plausible explain the case.

This chapter makes use of American, Soviet, Cuban, and – less frequently – other nations’ documents.¹ Most of these sources have been translated into English.² Although the records on the Soviet side are less extensive than for the United States, these sources allow for a more balanced and complete investigation of the crisis than is currently possible with the Berlin case.

The first section provides a minimalist historical overview of the Cuban Missile Crisis. The second delves into more detail on the interests and objectives of the United States and the Soviet Union, setting up the rest of the chapter. The third considers the Soviet decision to covertly deploy ballistic missiles to Cuba in a fait accompli, focusing on the significance of covertness to the Soviet strategy. The fourth section turns to the American decision about how to respond: airstrike, invasion, or blockade. It analyzes the decision for blockade from the standpoint of the success of the Soviet deterrent red line protecting Cuba from a fait accompli attack in the form of either an airstrike or an invasion. The fifth section focuses in on how and why the United States succeeded in deterring the

¹ Many of these documents come from the same collections. I mark documents [FRUS #] from the Foreign Relations of the United States collection on the Cuban Missile Crisis, along with the document number. I mark as [WC] documents from the Wilson Center Digital Archive. I mark as [GWU NSA] documents from the National Security Archive at George Washington University.
² My thanks to those involved in the translation of these documents. Because the supply of Russian documents is limited and the interest in the case among historians has led to impressive efforts at translation, my inability to speak Russian turned out to be less of an obstacle than I had feared. I was able to read Cuban documents in the original Spanish, but found that the most important of them were also available in English translations.
Soviet Union from running the blockade. The final section examines how the United States signaled resolve with sufficient effectiveness to induce the Soviet Union to immediately withdraw the missiles.

**Historical Overview**

In 1959, Cuban revolutionaries led by Fidel Castro defeated the forces of dictator Fulgencio Batista and formed a new government. Before long, relations between Cuba and the United States began to deteriorate. The Cuban government increasingly embraced communism and began to forge strong ties with the Soviet Union. The United States responded by embargoing trade with Cuba in 1960 and sponsoring the disastrous Bay of Pigs invasion by Cuban exiles in 1961, then declining to salvage the operation by intervening directly. By May 1962, Khrushchev had made the decision to deploy medium and intermediate range ballistic missiles with nuclear warheads to Cuba.\(^3\)

The Cuban Missile Crisis came as the culmination of perhaps the tensest period of the Cold War. The 1958-1961 period saw two major crises over Berlin – or, from another view, one prolonged crisis over Berlin – that peaked with the construction of the Berlin Wall in 1961. Throughout 1962, both sides feared and anticipated the next round over Berlin, which never came.

The overall balance of power had changed in several respects from the late 1940s. Soviet conventional forces remained larger than their NATO counterparts in Europe, but by a smaller margin than in 1948. The American nuclear monopoly was long gone, but the nuclear parity of the 1970s was only visible on a distant horizon. The size of the U.S. nuclear arsenal had expanded rapidly, and the United States held a fleeting but considerable advantage in that it could to deliver far greater numbers of nuclear weapons to targets in the Soviet Union than the Soviet Union could against targets in North America. More specifically, by the end of 1962 the United States fielded more than 200 ICBMs, a significant IRBM force located in Europe and Turkey, more than 600 B-52 bombers, and more than 800 B-47s. The Soviet Union, in comparison, had less than 50 ICBMs and a force of around 100 long-range bombers at the time of the crisis.\(^4\) Soviet leaders recognized that it would take the better part of a decade to construct the ICBM force needed to reach nuclear parity.\(^5\)

The distribution of military power in the Caribbean was more imbalanced than the overall strategic balance. In Khrushchev’s words to a senior Cuban official, “I asked [Soviet Defense Minister] Malinovsky, who knows all that you have in Cuba, how much time it would take him with forces like

---


\(^5\) Fursenko and Naftali, *One Hell of a Gamble,* p. 171.
those of the United States to occupy Cuba, and he answered three days. I am completely in agreement with Malinovsky’s judgment.”

The United States detected the installation of the Soviet missiles on October 16th, sparking intense deliberations about how to respond. Confronted with Soviet missiles in Cuba nearing operational readiness, American policymakers framed a choice among three plausible options: blockade, airstrike, and invasion. Other options received less attention, including inaction and negotiation from a position of weakness. After a great deal of internal debate, President Kennedy decided on the blockade option while continuing diplomacy and reserving the option to follow the blockade with either of the more forceful options.

Because international law considers blockades to be acts of war, the United States labelled its blockade a “quarantine” in order to maintain a fig leaf of legality. President Kennedy announced the quarantine on October 22nd and demanded the withdrawal of the missiles, “Our unswerving objective, therefore, must be to prevent the use of these missiles against this or any other country, and to secure their withdrawal or elimination from the Western Hemisphere.” This speech marked the onset of the Cuban Missile Crisis from the Soviet standpoint.

Although the moment in which Soviet ships cut their engines and, in some cases, turned around to return to Europe continues to be celebrated as a decisive American victory, the crisis was far from over at this juncture. The quarantine did nothing to directly remove the missiles and warheads already in Cuba.

Diplomatic and propaganda efforts took off at a frantic pace. In what many Americans regard as a seminal moment of the crisis, U.S. Ambassador to the United Nations Adlai Stevenson responded to repeated stonewalling by Valerian Zorin, the Soviet Ambassador to the United Nations, on the question of whether Soviet missiles were in Cuba. Stevenson famously declared, “I am prepared to wait for my answer until hell freezes over.” He then brought out reconnaissance photos of the Soviet missiles, showing them to the world. Zorin reported the encounter to Moscow a bit differently: “The attempts of the USA representative to turn the Council into a tribune for base propaganda met no support from other members of the Council.”

At this stage, the burden fell on the United States to make credible the threat to remove the missiles by force. The United States combined an array of signals to this end, including military

---

6 Note however that, in context, Khrushchev had an incentive to exaggerate. “Record of Conversations between Rodriguez and Khrushchev,” December 11, 1962 [WC].
mobilization, the forward deployment of forces to the Caribbean theater, firm public statements in which Kennedy put his reputation on the line, and the quarantine itself. The United States succeeded in changing Soviet perceptions, convincing the Kremlin that an imminent attack was likely enough that the immediate withdrawal of the missiles was necessary.

Both sides sought to avoid nuclear war without having to back down. This led to negotiations to end the blockade and the crisis with a compromise deal. The United States focused on one demand: the withdrawal of Soviet missiles, although some consideration was given to adding further demands concerning Soviet conventional forces in Cuba, Cuban support for rebels in Latin America, or even changes in the Cuban regime. In return for removing the missiles, the Soviets would receive the lifting of the blockade along with other concessions. Possible added incentives included an American non-invasion pledge toward Cuba, the withdrawal of American nuclear missiles from Turkey, the withdrawal of similar missiles elsewhere in Europe (e.g., Italy), and withdrawal from the Guantanamo Bay Naval Base. In the end, the Soviet Union obtained a public non-invasion pledge and a secret promise to withdraw the missiles from Turkey after a decent interval. Due to strong Turkish, NATO, and domestic opposition to that concession, it was omitted from official correspondence and acknowledged only in private conversations between Robert Kennedy and Soviet Ambassador Dobrynin.

The American threat to attack the missiles in Cuba was so credible that on October 28th the Kremlin decided to immediately announce the withdrawal of the missiles on Radio Moscow without waiting for the Soviet Embassy to decrypt the message with the decision and without even informing the Cuban government. Ambassador Dobrynin received orders to “quickly get in touch with R. Kennedy” to ensure that he was informed of the withdrawal.

Cuban leaders were unanimously furious with Khrushchev, both for withdrawing the missiles and for doing so without consulting or even notifying them beforehand. Soviet Ambassador (Havana) Alekseev described the atmosphere after the Soviet radio broadcast, “Confusion and bewilderment are reigning inside the Cuban leadership.” The Soviet leadership sent First Deputy Chairmen Anastas Mikoyan to Cuba to mend fences and solicit Cuban cooperation with the missile withdrawal and with fulfilling the Soviet promise to permit international inspections — something to which the Cubans never relented. Khrushchev did not react well to Castro’s bitter and uncooperative attitude, speaking in private of Cuban “stubbornness” and “arrogance” and contemplating severing ties with their “unreliable allies.” Mikoyan eventually had to fabricate an obscure Soviet law that prevented them from leaving

---

10 “State Department to the Embassy in the Soviet Union,” October 27, 1962 [FRUS #95].
11 “Minutes of Conversation between the Delegations of the CPCZ and the CPSU,” October 30, 1962 [WC].
12 “Gromyko to Dobrynin,” October 28, 1962 [WC].
15 “Central Committee of the Communist Party of the Soviet Union Presidium Protocol 66 [Instructions to Comrade A. I. Mikoyan],” November 16, 1962 [WC]; “Mikoyan to the Central Committee of the Communist Party of the
behind tactical nuclear weapons in Cuban hands – weapons that the United States did not know were in Cuba.  

Who won the Cuban Missile Crisis? The outcome is generally regarded as an American victory. The Soviets retreated once when they declined to run the blockade and again when they decided to withdraw the missiles. Khrushchev’s decision to remove the missiles was met with elation in the White House, with McGeorge Bundy famously declaring that “everyone knew who were hawks and who were doves, but that today was the doves’ day.” The Americans and the British discussed the importance of not “crowing” about the “greatly strengthened position” of the West after the crisis.

Anticipating the perception that the Soviet Union backed down, Khrushchev wrote to Kennedy saying,

To our mutual satisfaction we maybe even sacrificed self-esteem. Apparently, there will be such scribblers who will engage in hair-splitting over our agreement, will be digging as to who made greater concessions to whom. As for me, I would say that we both made a concession to reason and found a reasonable solution which enabled us to ensure peace for all including those who will be trying to dig up something.

Khrushchev was right about the scribblers. He came under intense public criticism from China and withering private criticism from Cuba. The secrecy of the American concession in Turkey hurt him in this regard. The perception that Cuba ended in a Soviet defeat contributed to his eventual loss of power in 1964. In defending his actions, Khrushchev would emphasize that he avoided nuclear war and safeguarded Cuba against invasion.

A strict comparison from before the Soviet missile deployment to the aftermath of the crisis supports Khrushchev’s claims. The Soviet Union returned to the status quo ante, i.e., no missiles in Cuba, in return for which it received two new concessions from the United States: the non-invasion pledge and the withdrawal of the missiles in Turkey. A cable from the U.S. Mission to the United Nations unwittingly made exactly this point: “We pointed out [to the UN Secretary-General] that effect of Khrushchev’s Turkey proposal was that, as the result of his own clandestine nuclear intrusion into the

---

17 “Record of the Tenth Meeting of the Executive Committee of the National Security Council,” October 28, 1962, [FRUS #106].
18 “Khrushchev to Kennedy,” October 30, 1962 [FRUS 120].
19 Chinese Embassy in Moscow, “Khrushchev’s Reconciliation with the United States on the Question of the Cuban Missile Crisis,” October 31, 1962 [WC].
20 Dmitry Polyanski,” Report on Khrushchev’s Mistakes in Foreign Policy,” October 24th, 1964 [WC].
Western Hemisphere, he gets a guarantee of Cuban integrity and the removal of the Turkey base, whereas all we get is removal of intrusion which he should not have made anyway.”

Yet few have ever been inclined to see the outcome this way, i.e., as a Soviet victory. Regardless, history does side with Khrushchev in remembering the crisis above all as a shared victory for both sides over the devastation of nuclear war.

Setting the Stage: The National Interests in Play

Were the United States and the Soviet Union willing to fight in order to achieve their preferred outcome with respect to the presence or absence of nuclear missiles in Cuba? What specific adversary actions would have provoked one side or the other to use force? This section explores these questions.

COMMON INTEREST IN AVOIDING WAR

As in the Berlin case, both sides sought above all to avoid a major war. Soviet leaders initially believed that the missile deployment would not provoke war, but came to harbor grave fears that the United States would be willing to use force to remove the missiles. In the United States, policymakers debated whether a fait accompli destroying the missiles could take place without provoking war. Both sides sought to make gains, but only insofar as they could do so while avoiding war. In Khrushchev’s words during a Central Committee meeting, “The point is that we do not want to unleash a war, we wanted to intimidate and restrain the USA vis-à-vis Cuba.”

Both sides also understood that their adversary sought to avoid nuclear war. According to Secretary of State Rusk, “We could be just utterly wrong, but, uh, we’ve never really believed that, that Khrushchev would take on a general nuclear war over Cuba.” Khrushchev emphatically rejected arguments that peaceful coexistence was impossible. Looking to control escalation even in the event of an attack on Cuba, the Kremlin sent the following order to Soviet forces in Cuba, “We categorically confirm that you are prohibited from using nuclear weapons from missiles, FKR [cruise missiles], ‘Luna’ [rockets] and aircraft without orders from Moscow.” At the height of the crisis, Khrushchev sent a cable to President Kennedy expressing a sentiment that both shared:

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

---

24 “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS 21].
26 “Malinovsky to Pliev,” October 27, 1962 [WC].
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.\textsuperscript{27}

However, whether this shared interest in avoiding war extended to Fidel Castro and the Cuban government remains open to question. By October 26\textsuperscript{th}, Castro was writing to Khrushchev to warn him that “aggression in the next 24 to 72 hours is almost inevitable.”\textsuperscript{28}

**U.S. INTERESTS**

Soviet missiles in Cuba mattered to the United States for two reasons: the change in the strategic balance and the message sent about the inability of the United States to contain the Soviet Union even in its own backyard.\textsuperscript{29} In explaining U.S. policy to British Prime Minister Macmillan, President Kennedy emphasized the strategic considerations: “The build-up in Cuba, if completed, would double the number of missiles the Soviets could bring to bear on the U.S. They would also overcome our warning system which does not face south. Furthermore, the short distance involving short times of flight would tempt them to make a first strike.”\textsuperscript{30}

However, fears about what Soviet missiles in Cuba would mean for American prestige may have been more important than the purely military considerations. During a White House meeting, McGeorge Bundy posed the question, “What is the strategic impact on the position of the United States of MRBMs in Cuba? How gravely does this change the strategic balance?” Secretary of Defense McNamara replied, “Mac, I asked the Chiefs that this afternoon, in effect. And they said, substantially. My own personal view is, not at all.” Among attendees, opinions on this point were also divided.\textsuperscript{31} President Kennedy was open to this line of reasoning, saying later in the conversation, “You may say it doesn’t make any difference if you get blown up by an ICBM flying from the Soviet Union or one that was ninety miles away. Geography doesn’t mean that much.”\textsuperscript{32}

Confronted by General Earle Wheeler’s (Army Chief of Staff) assertion that the missiles represented a “sort of quantum jump in [the Soviet] capability to inflict damage on the United States,” President Kennedy disagreed, “Well, it is a fact that the number of missiles there, I would say that no matter what they put in there, we could live today under. If they don’t have enough ICBMs today,\textsuperscript{33}

\textsuperscript{27} “The Embassy in the Soviet Union to the Department of State,” October 26, 1962 [FRUS #84].
\textsuperscript{28} “Castro to Khrushchev,” October 26, 1962 [WC].
\textsuperscript{29} For a clear statement from Secretary of State Rusk of why both considerations necessitated a strong response, see the following: John McCone, “Memorandum for the File,” Washington, October 19, 1962 [FRUS #28]. For the claim that the Soviets acted based on precisely these two goals: “Carroll to McNamara,” October 29, 1962 [GWU NSA].
\textsuperscript{30} Memorandum of Telephone Conversation between Kennedy and Macmillan, October 22, 1962 [FRUS #45].
\textsuperscript{31} John McCone, “Memorandum for the File,” October 17, 1962 [FRUS #23].
\textsuperscript{32} In context, this quotation should be interpreted as Kennedy giving serious consideration to this proposition, not necessarily as his endorsement of it. “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS #21].
they’re going to have them in a year.”\textsuperscript{33} A Special National Intelligence Estimate of October 19\textsuperscript{th} concluded that the missile deployment would not be able to undermine the American second strike capability.\textsuperscript{34} Despite the temptation to ascribe their determination to remove the missiles to the military threat they posed, American policymakers found the argument difficult to make. President Kennedy encountered this problem in a meeting on October 21\textsuperscript{st},

JFK: It’s just as if we suddenly began to put a major number of MRBM{s} in Turkey. Now that’d be goddam dangerous, I would think.
Bundy: Well, we did, Mr. President.
U.A. Johnson: We did it. We ...
JFK: Yeah, but that was five years ago.\textsuperscript{35}

Although the missiles in Cuba carried with them a larger shift in the short-term nuclear balance than the deployment of American missiles to Turkey, broader doubts remained about their military significance in a world with ICBMs increasingly rendering distance less significant.

Senior American policymakers did not completely agree on American goals during the crisis. The “doves”, i.e. blockade advocates and airstrike skeptics, prioritized removing the missiles. This camp came to include both Kennedy brothers and largely set American policy. Some “hawks,” i.e. advocates of airstrike and/or invasion, shared this goal but differed on how best to achieve it. However, other hawks prioritized a second American interest that could only be achieved by invasion: the removal of Castro’s regime. CIA Director McCone repeatedly emphasized that “we must all bear in mind that we have two objectives, one, disposing of the missile sites, and the other, getting rid of Castro’s communism in the Western Hemisphere.”\textsuperscript{36} He feared a domino effect in which the Cuban Revolution would spread across Latin America.\textsuperscript{37} The Joint Chiefs of Staff consistently included removing Castro as an overall objective during the crisis, with one memo regarding this as an argument for “all out invasion.”\textsuperscript{38} Weeks after the Soviets announced they would withdraw the missiles, a Joint Staff memo on Cuba listed three goals, “1) Attain removal of remaining offensive weapons, 2) Obtain removal of USSR military presence, 3) Obtain removal of Castro and/or reorientation of the Communist regime.” It added, “A minimum goal should be the accomplishment of 1 and 2 above and 3 as a long term objective.”\textsuperscript{39} This difference in basic objectives influenced the strategy debate over blockade vs. airstrike vs. invasion.

\textsuperscript{35} “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS #21].
\textsuperscript{37} Fursenko and Naftali, \textit{One Hell of a Gamble}, p. 204.
\textsuperscript{39} Joint Staff, “Highlights: Relationship of Cuban Objectives to Present Decisions,” November 16, 1962 [GWU NSA].
SOVIET INTERESTS

Historians continue to debate the motives behind the Soviet missile deployment. Was the objective to improve the strategic balance? To better deter an invasion of Cuba? To acquire leverage or a bargaining chip with respect to Berlin? In an October 17th assessment, CIA Director John McCone interpreted the missile deployment as Khrushchev’s attempt to establish a “hall mark” achievement, to bolster the defense of Cuba, and to improve the Soviet capability to strike the United States. Others drew a variety of connections to Berlin, with notions ranging from the missiles as a bargaining chip for a trade to the missiles as a way to provoke an incident which the Soviets would use to justify seizing West Berlin.

Washington was quick to interpret the missile deployment as symptomatic of Khrushchev’s “high degree of anxiety concerning the nuclear balance.” During the initial round of White House discussions on October 16th, the conversation touched on the American nuclear advantage and the Soviet desire to redress it. It was thought that the presence of American Jupiter missiles in Turkey may have prompted Khrushchev to look for a way to return the favor of emplacing missiles in the U.S. backyard.

Soviet sources tend to emphasize the loyal defense of the Cuban ally as the primary motive. Speaking on October 30th, Khrushchev explained,

We agreed to dismantle the missiles also because their presence in Cuba is essentially of little military importance to us. The missiles were meant to protect Cuba from attack; they helped us to wrench out of the imperialists the statement that they would not attack Cuba, and they thus served their main purpose. Otherwise we can hit the USA from elsewhere, and we do not need missiles in Cuba for that.

However, Khrushchev had every reason to make these claims regardless of his true motives. Admitting that changing the strategic balance was a major goal would have meant acknowledging failure. Nonetheless, many American analyses also emphasized this goal, including the initial CIA memo interpreting the newly-discovered missiles in Cuba, “The Soviet leaders’ decision to deploy ballistic missiles to Cuba testifies to their determination to deter any active US intervention to weaken or overthrow the Castro regime, which they apparently regard as likely and imminent.”

---

41 Sorensen, Kennedy, pp. 676-678.
43 “White House Meeting Transcript,” October 16, 1962 [FRUS #18].
44 Fursenko and Naftali, One Hell of a Gamble, pp. 68, 73, 182-183.
45 “Minutes of Conversation between the Delegations of the CPCZ and the CPSU,” October 30, 1962 [WC].
By this point in the Cold War, Berlin so transfixed both sides that key figures in Washington and Moscow often interpreted the other side’s policy as a stratagem with Berlin in mind. Secretary of State Rusk’s comments are representative of how some Americans thought Berlin motivated the Soviet missile deployment:

I think also that, uh, Berlin is, uh, very much involved in this. Um, for the first time, I’m beginning really to wonder whether maybe Mr. Khrushchev is entirely rational about Berlin. We’ve [hardly?] talked about his obsession with it. ... They may be thinking that they can either bargain Berlin and Cuba against each other, or that they could provoke us into a kind of action in Cuba which would give an umbrella for them to take action with respect to Berlin.47

Ironically, Soviet analysts tasked with explaining why the United States chose to initiate the Cuban Missile Crisis by reacting with such unexpected hostility to the deployment also drew links to Berlin. A Soviet intelligence retrospective written in November concluded that Kennedy took such a strong line in Cuba to “demonstrate to the Soviet Union the US readiness to resort to force in defense of its positions not only in Cuba, but in Western Berlin as well.”48 Soviet Ambassador Dobrynin (Washington) speculated that President Kennedy might have chosen a crisis over Cuba, where the United States held the conventional military advantage, as a way of preempting the next Berlin crisis in Europe where the Red Army was strong.49

Deploying Missiles to Cuba

With characteristic eloquence, Ted Sorensen described the Soviet deployment of nuclear missiles to Cuba, “Undertaken in secrecy, accompanied by duplicity, the whole effort was based on confronting Kennedy and the world in November with a threatening fait accompli.”50 The Soviet Union understood from the start that the United States would react with outrage to the deployment, but believed that a covert fait accompli would succeed nonetheless. The opportunity to covertly violate the red line shaped Soviet policy. In contrast, President Kennedy’s clear public statement of American resolve not to permit missiles in Cuba had little effect on Soviet perceptions or behavior. I begin with this latter point, before turning to the role of coverts in Soviet strategy.

On September 4th 1962 – six weeks before the detection of the missiles – President Kennedy drew an explicit, public red line against the introduction of nuclear missiles to Cuba:

There is no evidence of any organized combat force in Cuba from any Soviet bloc country; of military bases provided to Russia; of a violation of the 1934 treaty relating to Guantanamo; of

47 “White House Meeting Transcript,” October 16, 1962 [FRUS #18].
50 Sorensen, Kennedy, p. 678.
the presence of offensive ground-to-ground missiles; or of other significant offensive capability either in Cuban hands or under Soviet direction and guidance. Were it to be otherwise, the gravest issues would arise.

The result of prior deliberations, the statement was intended to aid in deterring the Soviet Union from taking these actions. By the time of this statement, the missile deployment was well underway. Kennedy’s statement increased Soviet fears that the United States had detected signs of the missiles more than it affected Soviet perceptions of the American response to the deployment. The Soviet Union already assumed that the deployment was inimical to American interests, so the statement in itself did little to change Soviet calculations. Kennedy’s statement contained not one but five prohibitions. The Soviet Union violated four – all save Guantanamo, which would have required attacking American forces. In the subsequent crisis, American policymakers quickly decided not to insist on the removal of Soviet conventional forces from Cuba – in effect conceding the failure of deterrence on that issue – in order to prioritize Soviet nuclear forces capable of striking targets in the United States.

The Soviet disregard for his September 4th statement surprised, embarrassed, and angered Kennedy. In his words, “My press statement was so clear about how we wouldn’t do anything under these conditions and under the conditions that we would. He must know that we're going to find out...” Briefing his subordinates around the world after the crisis subsided, French Foreign Minister Murville similarly wrote, “One cannot understand why the very clear warnings of President Kennedy in September were ignored.” Lemay, among others, reminded Kennedy about this statement in pushing for a more aggressive policy. Kennedy’s decision to make the already-existing red line explicit not only failed to bolster deterrence, but alarmed the Soviet Union that the covertness of their fait accompli was fleeting. The results of the September 4th statement: increased Soviet fears, the acceleration of the missile deployment, and the deployment of tactical nuclear weapons to safeguard the ballistic missiles.

Soviet leaders expected the missile deployment to succeed, i.e. that the United States would ultimately abide the presence of the missiles. In Khrushchev’s words to a senior Cuban leader (from a Cuban source), “The Yankees will scream, but they will have to put up with the missiles.” Defense Minister Malinovsky similarly told the Cubans, “You don’t have to worry … there will be no big reaction from the U.S. side.”

American policymakers immediately grasped the logic of the Soviet fait accompli strategy. In an October 19th meeting with the Joint Chiefs, President Kennedy described the missile deployment as “a

51 John McCon, “Memorandum of Meeting with the President,” August 23, 1962.
53 “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS #21].
54 “Murville to French Diplomatic Posts,” November 10, 1962 [WC].
55 “Meeting between Kennedy and Joint Chiefs of Staff,” October 19, 1962 in Widmer and Kennedy, Listening In.
56 Fursenko and Naftali, One Hell of a Gamble, p. 206.
57 “Record of Conversations between Rodriguez and Khrushchev,” December 11, 1962 [WC]. The quotation is from a prior trip to Russia that occurred before the crisis.
58 Fursenko and Naftali, One Hell of a Gamble, p. 196.
rather dangerous but rather useful play of theirs,” one which created “a very satisfactory position from their point of view.” Admiral Anderson agreed, “The Communists have got in this case a master situation, for their point of view, where every course of action posed to us is characterized by unpleasantries and disadvantages.”

A COVERT DEPLOYMENT

The Cuban Missile Crisis erupted out of a Soviet fait accompli specifically designed to exploit an unverifiable American red line. Covertness was central to the strategic calculations that led Khrushchev to deploy missiles to Cuba. Many Soviet officers discovered their destination was Cuba only when they opened sealed orders while at sea. Great care was taken to unload the missiles at night and to transport them to the launch sites via cleared routes. Soviet leaders contemplated making public their treaty with Cuba and the missile deployment, but decided that the strategic advantages of operating covertly trumped the downsides.

Even amidst the crisis, the Soviet Union sought to maintain and maximize what secrecy it could. Khrushchev was under considerable stress throughout the deployment, continually monitoring the progress of ships and searching for any sign that the Americans had detected the missiles. An October 27th telegram from Soviet Defense Minister Malinovsky to General Pliev in Cuba ordered, “Stop all work on deployment of R-12 and R-14 [missiles]—you are aggravating the United Nations. Camouflage everything carefully, work only at night.” The Soviet strategy relied on this secrecy.

Many in Washington recognized the vulnerability in the American red line against Soviet missiles (unverifiability) and warned of a potential Soviet fait accompli very much like the one which eventually took place. In 1961, Robert Kennedy told his brother, “If we don’t want Russia to set up missile bases in Cuba, we had better decide now what we are willing to do to stop it.” He foresaw the threat before current accounts suggest that Khrushchev had the idea. In the period immediately preceding the deployment, CIA Director John McCon argued that a Soviet missile

---

59 “Meeting between Kennedy and Joint Chiefs of Staff,” October 19, 1962 in Widmer and Kennedy, Listening In.
61 Nikolai Beloborodov, “The War Was Averted,” Early 1990s [GWU NSA].
65 “Malinovsky to Pliev,” October 27, 1962 [WC].
66 Fursenko and Naftali, One Hell of a Gamble, p. 97.
deployment to Cuba was likely in progress, repeatedly butting heads on this point with the majority of his subordinates. 68

If covertness was so central to Soviet strategy, why did they fail to effectively conceal the ballistic missile sites in Cuba? Is this a failed certain observable implication for red lines theory? This question resembles the concern about the puzzling absence of nonviolent countermeasures to the Berlin airlift such as jamming raised in the previous chapter. Fortunately, more evidence is available about Soviet policy during the Cuban Missile Crisis.

Soviet documents reveal that lack of adequate camouflage resulted from sloppy implementation of the intended policy. The Soviet intent was to covertly emplace missiles in Cuba without ever revealing their exact location. Their general presence would only have been revealed after assembly and if circumstances warranted. 69 In Moscow, the blame ultimately fell on the advance team that signed off the feasibility of the plan. 70 Soviet forces arrived expecting to conceal the missiles in palm groves that, they found on arriving, were too sparse to serve that role. 71 Ironically, some in Washington fell prey to the same misconception, with General Taylor skeptical of an ultimatum before a airstrike because, “They can pull in under trees and forest and disappear almost at once, as I visualize.” 72 Soviet plans similarly failed to prepare in any way for the problem of hurricanes and tropical storms. 73

During one of many acrimonious meetings between the Cuban leadership and Soviet envoy (Mikoyan) after the unilateral Soviet decision to withdraw, Fidel Castro pointed out that the missiles could have been hidden beneath large tents used at tobacco plantations or inside fake buildings. 74 The Soviets, having failed to anticipate the problem, considered neither option until it was too late. In a November 21st speech in Havana intended primarily to rally the morale of departing Soviet forces and persuade them that withdrawal was a victory rather than a defeat, Mikoyan diverged from these themes to excoriate the lack of attention to camouflage and the harm it did to Soviet plans. 75

While scrambling to justify their retreat to Cuban leaders, Mikoyan repeatedly emphasized that the discovery of the missiles undercut their deterrent value, presumably because it rendered them

---

68 He based this suspicion not just on the opportunity for it, but also on the fact that the newly-installed Soviet air defense system in Cuba was too extensive to be justified by any other purpose. “Memorandum of Conversation between McConie and Bundy,” October 5, 1962 [FRUS #9]; “Carter to McConie,” September 19, 1962 in Central Intelligence Agency, CIA Documents on the Cuban Missile Crisis, 1962 (Central Intelligence Agency, Washington, D.C., 1992).

69 “Memorandum of Conversation between Mikoyan, Castro, Dorticos, Others,” November 22, 1962 [GWU NSA].

70 Ibid.


72 “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS #21].


74 “Memorandum of Conversation between Mikoyan, Castro, Dorticos, Others,” November 22, 1962 [GWU NSA].

vulnerable to a first strike. Covertness was very much the Soviet intent, the execution merely fell short.

**Blockade, Airstrike, or Invasion?**

President Kennedy and his advisors faced a pivotal choice between invasion, airstrike, blockade, some combination thereof, or inaction. Various figures raised the last option – usually framed as diplomatic negotiations – but it was always ruled out quickly as too feeble a response. An airstrike could immediately eliminate the threat. An invasion could do that plus remove Castro and the potential re-introduction of Soviet missiles in the future. Both options represented a possible *fait accompli* that the United States could impose to its benefit and leave the Soviet Union in the unenviable position of having to react while still wanting to avoid war. A blockade, in contrast, provided no direct solution to the missiles already in Cuba. So why the choice for a blockade, at least as the opening move?

The most important part of the answer is simple but critical: too many American policymakers, including the President, did not believe they could get away with a violent *fait accompli* in Cuba without provoking a Soviet response with grave risks for the United States. Soviet deterrence of an attack on Cuba was, at least initially, successful. The Soviet red line was credible. Despite their strong preference for a the firmest possible response to the missile deployment, American policymakers reluctantly decided to get away with as much as they could – the blockade – without violating the Soviet red line against overt attack. The blockade was a means of applying pressure that worked around that red line without crossing it.

This conclusion emerges from the contours of the debate won by the blockade advocates. Initially, however, the majority favored using force, including key figures who would eventually steer policy away from the use of force. On October 16th, the first day, Defense Secretary McNamara said, “If we saw a warhead on the site and we knew that that launcher was capable of launching that warhead, I would … Frankly, I would strongly urge against the air attack, to be quite frank about it.” Later in the same meeting, President Kennedy seemed to decide in favor of force,

We’re certainly going to do number one; we’re going to take out these, uh, missiles. Uh, the questions will be whether, which, what I would describe as number two, which would be a

---

76 These claims, which he repeated on multiple occasions, may have been sincere or may have been a desperate stab at finding a justification that would mollify the Cubans. Mikoyan never explained how missiles unknown to the United States could deter the United States. I infer that he meant only for their exact locations to be secret, not their presence somewhere in Cuba. “Mikoyan to the Central Committee of the Communist Party of the Soviet Union,” November 6, 1962 [WC]; “Record of Conversation between Mikoyan and Castro,” November 13, 1962 [WC]; “Dorticos to Castro,” November 25, 1962 [WC].


78 “White House Meeting Transcript,” October 16, 1962 [FRUS #18].
general air strike. That we're not ready to say, but we should be in preparation for it. The third is the, is the, uh, the general invasion. At least we're going to do number one, so it seems to me that we don't have to wait very long. We, we ought to be making those preparations.²⁹

National Security Advisor McGeorge Bundy recommended “decisive action [airstrike] with its advantages of surprises and confronting the world with a fait accompli.”³⁰ On October 27th, the day before the Soviet Union announced the withdrawal of the missiles, the Joint Chiefs of Staff recommended the implementation of OP Plan 312 (airstrike) to be followed one week later with OP Plan 316 (invasion).³¹ The Pentagon consistently pushed for a forceful fait accompli. Some allies were willing to support such a policy. Despite his nation’s exposed position at the edge of the Iron Curtain, West German Chancellor Konrad Adenauer made clear that he would support an airstrike and an invasion if Washington deemed them necessary.³²

American policymakers were cognizant of another incentive to strike: the rapidly closing window of opportunity to strike before the weapons became operational.³³ Because an airstrike would not prevent the re-introduction of strategic missiles, advocates of using force tended to support either a long-term blockade or an invasion as a follow-on to the initial airstrike as a way of ensuring the continuing absence of missiles from Cuba.³⁴ All – even the blockade advocates – acknowledged that a blockade in itself did nothing to remove the missiles.³⁵

During a contentious meeting between President Kennedy and the Joint Chiefs, Curtis Lemay insisted that “we don’t have any choice except direct military action,” referring to a blockade as “political action.” Taking a perspective radically at odds with Kennedy, Lemay disagreed with the notion that an attack on Cuba would provoke a response in Berlin, “I don’t think they’re going to make any reply if we tell them the Berlin situation is just like it’s always been.” In contrast, he claimed a blockade “will lead right into war. This is almost as bad as the appeasement at Munich.” This not-so-thinly veiled criticism culminated in the exchange:

---

²⁹ Bundy then replied, “You want to be clear, Mr. President, whether we have definitely decided against a political track?” “White House Meeting Transcript,” October 16, 1962. CIA Director McCone left the meeting with this impression. John McCone, “Memorandum for the File,” October 17, 1962 [FRUS #23].

³⁰ “Record of Meeting, Washington, 11 a.m.,” October 19, 1962 [FRUS #31].

³¹ “Taylor to Kennedy,” October 27, 1962 [GWU NSA]; “Record of the Eighth Meeting of the Executive Committee of the National Security Council,” Washington, 4 p.m., October 27, 1962, [FRUS #94].

³² Adenauer did not explain his reasoning, but he did suggest that an ultimatum with a 24 hour deadline precede an attack. “Memorandum of Conversation between Konrad Adenauer and Walter Dowling,” October 28, 1962 [WC].


³⁴ “White House Meeting Transcript,” October 16, 1962 [FRUS #18].

³⁵ “Minutes of the 505th Meeting of the National Security Council,” Washington, 2:30-5:10 p.m., October 20, 1962, [FRUS #34]; “Summary Record of the Sixth Meeting of the Executive Committee of the National Security Council,” Washington, 10 a.m., October 26, 1962, [FRUS #79].
Lemay: I think that a blockade, and political talk, would be considered by a lot of our friends and neutrals as being a pretty weak response to this. And I’m sure a lot of our own citizens would feel that way, too. In other words, you’re in a pretty bad fix, Mr. President.

JFK: What did you say?
Lemay: You’re in a pretty bad fix.
JFK: You’re in there with me.86

Despite the initial appeal of the airstrike option and the continuing efforts of hawks like Lemay, opinion quickly began to shift toward a blockade. Robert Kennedy clearly favored the blockade by October 19th, as the minutes of a White House meeting record:

He thought it would be very, very difficult indeed for the President if the decision were to be for an air strike, with all the memory of Pearl Harbor and with all the implications this would have for us in whatever world there would be afterward. For 175 years we had not been that kind of country. A sneak attack was not in our traditions. Thousands of Cubans would be killed without warning, and a lot of Russians too. He favored action, to make known unmistakably the seriousness of United States determination to get the missiles out of Cuba, but he thought the action should allow the Soviets some room for maneuver to pull back from their over-extended position in Cuba.87

Robert Kennedy believed that an attack kill Soviet soldiers and therefore “would lead to unpredictable military responses by the Soviet Union which could be so serious as to lead to general nuclear war.” Consequently, the best course was to impose the quarantine and “play for the breaks.”88 He remained, however, open to an airstrike if a blockade and intense pressure failed to coerce the Soviet Union into withdrawing the missiles.89

President Kennedy was openly skeptical of an airstrike by October 22nd, the day of the speech announcing the blockade.90 Despite his initial reaction in favor of a strike, McNamara’s views had already begun to shift by the evening of the first day (the 16th): “Now this alternative [blockade] doesn’t seem to be a very acceptable one, but wait until you work on the others.”91

WHY THE SOVIET RED LINE WAS CREDIBLE

86 “Meeting between Kennedy and Joint Chiefs of Staff,” October 19, 1962 in Widmer and Kennedy, Listening In.
87 “Record of Meeting, Washington, 11 a.m.,” October 19, 1962 [FRUS #31].
88 Robert McNamara, “Notes on Meeting with President Kennedy,” October 21, 1962 [FRUS #36].
89 Note individuals’ positions during these discussions were not always consistent. “Minutes of the 505th Meeting of the National Security Council,” October 20, 1962 [FRUS #34].
90 “Minutes of the 507th Meeting of the National Security Council,” October 22, 1962, [FRUS #41].
91 “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS 21].
The strength of the Soviet red line – as found in the statements and reasoning of American policymakers seeking to overcome it – came primarily from the presence of Soviet troops. The United States was willing to attack Cuba, but not those troops. On October 17\textsuperscript{92}, CIA Director McCone wrote:

Consequences of action by the United States will be the inevitable “spilling of blood” of Soviet military personnel. This will increase tension everywhere and undoubtedly bring retaliation against U.S. foreign military installations, where substantial U.S. casualties would result, i.e., Tule, Spanish bases, Moroccan bases, and possibly SAC bases in Britain or Okinawa. Jupiter installations in Southern Italy, Turkey ... do not provide enough “American blood.”\textsuperscript{92}

By October 19\textsuperscript{93}, McNamara was arguing against the airstrike option, emphasizing Soviet casualties, the costs of a Soviet response, and the difficulties of controlling subsequent events.\textsuperscript{93} While explaining the decision for blockade to French President De Gaulle, Dean Acheson cited “the high number of Soviet technicians that would have been killed” as the main reason President Kennedy rejected an airstrike.\textsuperscript{94} On the Soviet side, Mikoyan thought much the same, “After all, we have a whole army here [Cuba]. If an invasion on the part of the Americans began, it would have led to a global confrontation.”\textsuperscript{95} An exchange between President Kennedy, Secretary McNamara and Senator Fulbright underscores the weight Kennedy placed on the presence of Soviet troops:

[Fulbright] stated in his opinion it would be far better to launch an attack and to take out the bases from Cuba. McNamara stated that this would involve the spilling of Russian blood since there were so many thousand Russians manning these bases. Fulbright responded that this made no difference because they were there in Cuba to help on Cuban bases. These were not Soviet bases. There was no mutual defense pact between the USSR and Cuba. Cuba was not a member of the Warsaw Pact. Therefore he felt the Soviets would not react if some Russians got killed in Cuba. The Russians in the final analysis placed little value on human life. The President took issue with Fulbright, stating that he felt that an attack on these bases, which we knew were manned by Soviet personnel, would involve large numbers of Soviet casualties and this would be more provocative than a confrontation with a Soviet ship.\textsuperscript{96}

The Soviet Union lacked a publicly-acknowledged alliance with Cuba and denied the presence of its forces in Cuba, except in the roles of technicians and advisors. As the literature often approaches the credibility of extended deterrent commitments, this commitment should be quite weak.\textsuperscript{97} Some

\textsuperscript{92} John McCone, “Memorandum for Discussion [The Cuban Discussion],” October 17, 1962 [FRUS #26]. The redaction could relate to the U.S. base at Guantanamo Bay.

\textsuperscript{93} John McCone, “Memorandum for the File,” October 19, 1962 [FRUS #28].

\textsuperscript{94} “Record of a Meeting between Charles De Gaulle and Dean Acheson, October 22, 1962 [WC].

\textsuperscript{95} In context, he could have said this merely to placate the Cubans. “Memorandum of Conversation between Mikoyan, Castro, and Dorticos,” November 22, 1962 [GWU NSA].

\textsuperscript{96} “National Security Action Memorandum 196,” October 22, 1962 [FRUS #42].

\textsuperscript{97} The Soviet Union had issued threatening statements to the effect that invading Cuba could lead to war, but refrained from taking the traditional actions to publicly stake a commitment to Cuba’s defense. CIA Director McCon described the most important of these statements as follows, “Moscow has once again used vague and
American policymakers contemplated attempting to regard all missiles and forces in Cuba as Cuban, and so get around the focal point of firing on Soviet forces. Khrushchev did contemplate this possibility, but only as one among many possible contingencies. Fulbright made the case for a policy predicated on exploiting this (above), as did a few others. A Special National Intelligence Estimate written on October 18th provides the clearest articulation. Because the Soviet Union lacked an explicit obligation to defend Cuba, the estimate predicted a Soviet counterstroke of manageable proportions that would not lead to war.

The Soviets have no public treaty with Cuba and have not acknowledged that Soviet bases are on the island. This situation provides them with a pretext for treating US military action against Cuba as an affair which does not directly involve them, and thereby avoiding the risks of a strong response. We do not believe that the USSR would attack the US, either from Soviet bases or with its missiles in Cuba, even if the latter were operational and not put out of action before they could be readied for firing. ... Since the USSR would not dare to resort to general war and could not hope to prevail locally, the Soviets would almost certainly consider retaliatory actions outside Cuba.

However, most of the major players involved at the White House never accepted this logic or made this argument. Instead, the belief that shaped U.S. policy was that the Soviet Union would feel compelled to respond more forcefully to an American attack on their forces in Cuba despite the lack of any explicit public commitment to do so.

Advocates of force tended to be more sanguine about the Soviet response, but the virtually everyone expected that the Soviets would feel compelled to do respond somehow. According to Dean Acheson, “We should proceed at once with the necessary military actions and should do no talking. The Soviets will react some place. We must expect this; take the consequences and manage the situations as they evolve.” In McNamara’s words, “I, I don't know quite what kind of a world we live in after we've struck Cuba, and we, we've started it.”

WHAT HAPPENS AFTER AN ATTACK ON CUBA?

---

101 Note that this is implicitness rather than imprecision. Hawk or dove, every American official understood that killing Soviet forces and attacking Cuba would have violated a Soviet red line; they differed only on whether the United States could get away with it.
103 “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS 21].
Across the hawk-dove divide, American policymakers largely agreed that the Soviets would likely respond to a violation of their red line in Cuba and made a concerted effort to identify and prepare for counterattacks around the globe.\textsuperscript{104} Specifics varied, but most American officials envisioned an initially proportional response. Berlin consistently ranked at the top of the list of likely targets. The President held this view.\textsuperscript{105} Many American policymakers also anticipated a symmetrical response to the blockade in the form of a blockade of Berlin.\textsuperscript{106} Perhaps due to the brevity of the crisis, the Soviets never did so. Others thought the Soviet Union would find a place to retaliate that did not require an attack on a NATO member.\textsuperscript{107} A strike on the Jupiter Missiles in Turkey also potentially fit the bill for a proportional response.\textsuperscript{108} General Taylor, Chairman of the Joint Chiefs of Staff, regarded a Cuban attack on the base at Guantanamo Bay as an inevitable consequence of using force, although he left open the possibility that this attack might consist only of artillery fire rather than an attempt to storm the base.\textsuperscript{109} Scenarios in Iran, Korea, Taiwan, and elsewhere received consideration as well.\textsuperscript{110}

Do these perceptions support red lines theory or the conventional wisdom? A Joint Staff memo of October 17\textsuperscript{th} examines how the Soviet Union would respond to an attack on Cuba helps clarify the distinction. It asserts that two “basic principles” are at issue: relative power and the relative involvement of vital interests in Cuba. Building on that premise, the report saw the situation as advantageous for the United States because “US vital interests are clearly involved in Cuba; the Soviet vital interests do not yet appear vitally involved” and “The United States has the advantage over the USSR in a general war.” The report therefore concludes that Soviet policymakers would retaliate in some form for an attack on Cuba, but only with measures that they could be confident would avoid war.\textsuperscript{111} The logic is compelling, corresponding neatly to the conventional wisdom’s view of the two foremost structural determinants of crisis outcomes: relative power and relative interest.

However, President Kennedy and most of the other senior policymakers on the American side of the crisis did not share this view. Instead, they believed that violating the firing-on-forces focal point

\textsuperscript{104} McNamara and Roswell Giplatric emphasized the need for this study of possible countermoves from the outset of the crisis. “Off the Record Meeting on Cuba,” Washington, 6:30-7:55 p.m., October 16, 1962 [FRUS 21].
\textsuperscript{106} “Minutes of Meeting of the Executive Committee of the National Security Council,” October 23, 1962 [FRUS #47].
\textsuperscript{107} “Holtoner, Gray, and Wylie to Taylor [“Prospective Soviet Responses to US Action”],” October 17, 1962 [GWU NSA]; “Kent to McCon [“Reactions to US Air Strike against Strategic Missile Bases MIG-21s and IL-28s”],” October 27, 1962 [GWU NSA].
\textsuperscript{108} “White House Meeting Transcript,” October 16, 1962 [FRUS #18].
\textsuperscript{109} Off the Record Meeting on Cuba,” October 16, 1962 [FRUS #21]. For a similar view: “Kent to McCon [“Reactions to US Air Strike against Strategic Missile Bases MIG-21s and IL-28s”],” October 27, 1962 [GWU NSA].
\textsuperscript{110} “Holtoner, Gray, and Wylie to Taylor [“Prospective Soviet Responses to US Action”],” October 17, 1962 [GWU NSA].
\textsuperscript{111} “Holtoner, Gray, and Wylie to Taylor [“Prospective Soviet Responses to US Action”],” October 17, 1962 [GWU NSA].
would risk unacceptable levels of escalation. This report from the Joint Staff reflects the views of some of the most hawkish elements in Washington, not the views that ultimately shaped American policy.

Soviet leaders were thinking along similar lines as President Kennedy. In Khrushchev’s words during a Central Committee meeting on October 23rd, “The tragic thing — they can attack, and we will respond. This could escalate into a large-scale war.” Mikoyan explained both the logic of an American 

fait accompli

and why it would backfire: “[an airstrike] could have been done in several minutes, and we could have faced this fact and a choice – to initiate World War III, swallow the defeat, or retaliate with a localized strike on points that were sensitive for the West, such as American missiles in Turkey.” Mikoyan continued,

If our group of forces, totaling 42,000 people, was destroyed, it would have meant the strike on the missiles in Turkey would not have been a sufficient action; most likely, in such a case, West Berlin could have been involved … it would have been taken by the allied forces under a formal control of [East Germany] after a blitzkrieg. Certainly, we would have tried to limit our actions only with the Berlin operation and the use of conventional weapons in hope that both theaters of war ‘would be localized conflict zones.’ Despite all the efforts of both the USSR’s and the United States’ governments, there could have been the danger of a nuclear war – in other words, a total catastrophe.

When John Scali asked KGB Station Chief Aleksandr Feklisov how the Soviet Union would respond to an airstrike, he replied that “West Berlin would be at the very least occupied.” Soviet Foreign Minister Gromyko stated, “There is no consensus as to how and where this riposte would come, but it would happen – about this there is no doubt.”

There was no confidence on either side that escalation could be contained once the firing-on-forces focal point was crossed. This fear, at least as much as the fears of an initial proportional Soviet response (in Berlin or elsewhere) deterred the United States from using force. In President Kennedy’s words, “It isn’t the first step that concerns me, but both sides escalating to the fourth and fifth step – and we don’t go to the sixth because there is no one around to do so. We must remind ourselves that we are embarking on a very hazardous course.” Kennedy also worried that escalation would not require even that many additional steps. Noting that attacking Cuba could provoke a Soviet seizure of Berlin, “which leaves only one alternative, which is to fire nuclear weapons – which is a hell of an alternative – and begin a nuclear exchange.” In his account of the crisis, Sorensen emphasizes the

---

113 Mikoyan, The Soviet Cuban Missile Crisis, p. 114.
115 Fursenko and Naftali, One Hell of a Gamble, p. 232.
116 Polmar et al., Defcon-2: Standing on the Brink of Nuclear War During the Cuban Missile Crisis, p. 190.
117 “Meeting between Kennedy and Joint Chiefs of Staff,” October 19, 1962 in Widmer and Kennedy, Listening In.
importance of these fears and the doubts about whether cooler heads would truly prevail after initial hostilities.\(^{118}\)

Although the Pentagon expressed optimism that the United States could use force against Cuba without provoking the use of nuclear weapons, the White House and Kremlin were more skeptical. A Joint Staff memo predicted that Soviet leaders would not order a nuclear missile launch from Cuba even if the United States attacked, but that there was a “remote possibility” of a Soviet officer in Cuba doing.\(^{119}\) Soviet forces in Cuba did have the orders, “All the forces initially should not use atomic weaponry,” but Soviet plans called for the use of tactical nuclear weapons in the defense of Cuba.\(^{120}\) The Pentagon was late to planning for the possibility of tactical nuclear weapons – due mainly to having no knowledge of their presence – but a document from November 2\(^{\text{nd}}\) belatedly addressed the issue, concluding that the Soviet Union would be reluctant to use these weapons “to repel invasion.” It added, “However, if the Cuban leaders took this foolhardy step, we could respond at once in overwhelming nuclear force against military targets.”\(^{121}\) The potential that escalation past the firing-on-forces threshold would lead to nuclear war was real.

President Kennedy explained the overall strategy to British Prime Minister Macmillan, “We are attempting to begin the escalation in a way to prevent WW III. Maybe this will result anyway, but we cannot accept his [Khrushchev’s] actions.”\(^{122}\) Secretary of State Rusk repeatedly emphasized the importance of moving forward to a planned extent and then creating a “pause” in which the Soviet Union could agree to a settlement that would walk the two sides back from the brink.\(^{123}\) I examine this strategy of compellence and signaling in more detail in a later section.

The blockade was not in itself a complete strategy, because it left the missiles in place. The central reason for the choosing the blockade option despite this obvious deficiency was straightforward: the United States did not believe it could get away with a violent fait accompli in Cuba without provoking a Soviet response too likely to lead to war. The problem: the need to attack Soviet forces to achieve the objective. At the strategic level, the United States took the strongest unilateral action it could without having to overtly attack the other side. The same logic governed the crisis at the tactical level with respect to the blockade.

The Blockade Quarantine Line at the Tactical Level

---

\(^{118}\) Sorensen, *Kennedy*, p. 685.


\(^{120}\) “Central Committee of the Communist Party of the Soviet Union Presidium Protocol 60,” October 23, 1962 [WC].

\(^{121}\) “Taylor to Kennedy,” November 2, 1962 [GWU NSA].

\(^{122}\) “Memorandum of Telephone Conversation between Kennedy and Macmillan,” October 22, 1962 [FRUS #45].

“We’re eyeball to eyeball, and I think the other fellow just blinked.” Secretary of State Dean Rusk’s reaction to reports that Soviet ships halted short of the blockade line remains a signature phrase in the public memory of the Cuban Missile Crisis. Khrushchev gave that order on October 23rd, and the White House received the good news with jubilation on the morning of October 24th, the day the blockade went into effect. This section examines the blockade at the tactical level. It asks: why did the United States succeed in stopping those ships?

The blockade red line itself did not hold. The United States was not able to uphold as a general rule that no Soviet ship moving toward Cuba could cross the line. Soviet leaders never believed that unilaterally maneuvering a ship through the blockade line would lead directly to the risk of unacceptable punishment, or even necessarily to the U.S. Navy sinking the ship. Instead, the effectiveness of the blockade hinged on both sides’ expectations of how an interception on the high seas would go. Because the United States Navy could likely disable or board Soviet freighters without sinking them or firing (to kill) first, the advantage lay with the Americans. These actions, albeit aggressive, fell in a gray area in the Soviet red line against the overt use of force. It was ambiguous whether these were intolerable attacks or something different and lesser. Red lines theory, therefore, implies that the United States should be able to get away with these actions, and that the Soviet Union would react less forcefully to these ambiguous attacks than to unmistakably overt attacks. The Soviet decision to “blink” was not a strategic-level retreat due to the credibility of American threats to go to war. On the contrast, the determinants of the withdrawal were quite tactical in nature. Soviet ships turned back due to the upper hand that the U.S. Navy would have in an intercept encounter in which both sides refrained from overt attacks on the other.

The American blockade line was, in the terminology of red lines theory, entirely arbitrary and more than a little imprecise. President Kennedy inadvertently summed up the arbitrariness problem nicely, “When we talk about a barrier, we talk about 500 miles, we can always take it at 300 miles if you have to.” Unlike Berlin, the blockade line did not fall on an extant land boundary. Even after the blockade commenced, the Defense Department considered moving it both closer to and further from Cuba. The timing of this red line was equally arbitrary. The State Department made clear to the Soviet Union that the quarantine “will go into effect at 1400 hours Greenwich time October twenty-four,” but why that time rather than a few hours earlier or later? Why the next ship slated to arrive after that time, but not those already near or past it?

---

125 “Memorandum of Telephone Conversation between Kennedy and Ball,” October 24, 1962 [FRUS #64].
126 Both options were ostensibly discussed as signals of resolve, but the tenor of the document suggests its authors held little faith that either would make much difference. “Blockade Measures,” n.d. [25 October 1962] [GWU NSA].
127 “The Department of State to the Embassy in the Soviet Union,” October 23, 1962 [FRUS #52].
American policymakers initially wanted to deter any violation of the blockade line, but the arbitrary line made this difficult. The following conversation between UN Ambassador Adlai Stevenson and Under-Secretary of State George Ball regarding a Soviet tanker approaching the blockade line captures this dynamic:

Ball—There is a ship which was going to be challenged at 2 o'clock tomorrow morning, which is just about 2-1/2 hours from now. We have got that held off.
Stevenson—Would we stop it anyway?
Ball—Yes, we were going to stop it.
Stevenson—Although it was not carrying [weapons]—
Ball—that is on the theory that we stop everything and challenge it and find out what's on board.

... Ball—You know, there is a little flexibility in this because we are challenging these ships 500 miles out; we could outrun most of them and we could challenge them 200 miles out if necessary.128

Although specifying an exact line reduced imprecision, it did not eliminate it. However the line might appear on a map, its exact position in the middle of the ocean from the vantage point of the ships in theatre was another matter. It was marked only by the presence of 16 U.S. Navy destroyers, three cruisers, one carrier, and six additional ships.129 These ships did not all hold constant positions. Some imprecision at the level of an intercept was unavoidable, and there would be no all-or-nothing moment at which the line was clearly crossed.

In keeping with the weak quality of the blockade red line, the Soviet Union did not unconditionally respect it. Even as most ships halted or turned back, Soviet tankers continued toward Cuba.130 The first tanker, the Bucharest, declared that its cargo was petroleum products and was allowed to proceed without physical inspection or boarding.131 After allowing the Bucharest through the line without an inspection, American policymakers debated how to address the next Soviet ship approaching the line, the Grozny. Key figures fell on both sides, but the ship ultimately came to a halt before the final decision had to be made.132 American policymakers also decided to allow the East German passenger ship Voelker Fruendschaft through the line, citing its lack of weaponry as the public excuse for what was in actuality motivated by the desire to avoid a first confrontation with a ship full of

128 “Memorandum of Conversation between Ball and Stevenson,” October 24, 1962 [FRUS #66].
129 Sorensen, Kennedy, p. 708.
130 This was perhaps due to the fact that they were so unlikely to carry weapons. It may also have been due to the value of their cargo for Cuba. “Memorandum of Telephone Conversation between Ball and Bundy,” October 24, 1962 [FRUS #58].
civilians. In contrast, a Lebanese-flagged vessel carrying dry cargo would have been inspected, in part because no resistance was expected from it. Overall, it was not the line itself that held, but rather a more specific strategic calculation for each ship as an isolated case.

Soviet submarines also successfully violated the blockade line, with the U.S. Navy consistently failing to limit their movements. Although an unlikely way to move cargo, American officials did worry that submarines could eventually be used to deliver sensitive materials such as nuclear warheads. The submarines mattered more, however, for the threat they posed to American warships conducting the blockade and to the general principle of enforcing the line. Although one might imagine that the obstacle to stopping Soviet submarines was locating them, the U.S. Navy had a great deal of success at discovering and harassing the noisy Soviet submarines operating in their backyard. The difference between the subs and the freighters was not finding them, but rather the fact that the U.S. Navy could not safely disable or board a Soviet submarine without overtly attacking it, with the attendant possibility of sinking it and the escalation that could ensue. In this sense, submarines are more like aircraft than surface ships.

Encounters between American naval vessels and Soviet submarines produced some of the tensest moments of the Cuban Missile Crisis, often unbeknownst to the White House. American sources record American warships stalking and harassing Soviet submarines, up to and including dropping hand grenades and other explosive signaling devices to pressure the submarine to surface. In this way, the American Navy walked right up to the edge of the firing-on-forces focal point. From the Soviet standpoint, it was difficult to distinguish these non-lethal explosions from potentially-lethal depth charges. Soviet sources record submarines being forced to the surface by extremely high temperatures of up to 140 degrees Fahrenheit. The submarines then found themselves confronted with swarms of ships and aircrafts, blinding searchlights, numerous weapons aimed at them, and warning shots. In some cases, Soviet forces out of immediate contact with Moscow feared that they had been fired on or that war had begun. One of these incidents led to the arming of a nuclear torpedo amid an intense debate between Soviet officers that ultimately saw cooler heads prevail. The extent of these

133 “Summary Record of the Fifth Meeting of the Executive Committee of the National Security Council,” October 25, 1962 [FRUS #73]; “Summary Record of the Sixth Meeting of the Executive Committee of the National Security Council,” October 26, 1962 [FRUS #79].
134 “Summary Record of the Sixth Meeting of the Executive Committee of the National Security Council,” October 26, 1962 [FRUS #79].
135 According to President Kennedy, “Of course, I don’t see how we could prevent further ones from coming in by submarine.” “White House Meeting Transcript,” October 16, 1962 [FRUS #18]. “Smith to McCone [Soviet Challenge to the Quarantine],” October 23, 1962 [FRUS #54].
138 Vadim Orlov, “Recollections,” 2002 [GWU NSA]; Polmar et al., Defcon-2: Standing on the Brink of Nuclear War During the Cuban Missile Crisis, pp. 160-162.
confrontations and the potential for non-violent harassment to come across as violent attack may well have exceeded the intent of the White House. In the end, however, both sides respected the firing-on-forces focal point and forewent attacking each other in these incidents.

The American blockade never did stop Soviet submarines, its success in making their lives miserable notwithstanding. In the same Central Committee meeting where the decision was made to turn back the surface ships, the orders immediately following those were, “Keep the submarines on their approaches.”

**WHY THE BLOCKADE WORKED**

If the United States was focused on signaling resolve, why did it allow vessels through the blockade line without the inspections on which it was insisting? The answer: the strength of the U.S. quarantine came not from the inviolability of the line, but rather from the advantages the U.S. would have an intercept encounter with a Soviet ship. Letting harmless ships through the line did nothing to change this advantage and therefore did not undermine the blockade line.

The decisive determinant of the blockade’s tactical success was the ability of the U.S. Navy to disable or board and halt Soviet freighters without needing to overtly destroy or sink the ships. These actions fell in a gray area in the Soviet red line against overt attack. Both sides anticipated what a climactic intercept might look like, and both sides saw it ending in either an American success or an escalatory incident neither wanted. Because neither contingency suited the Soviets, they backed down.

On October 24th, the pivotal day for the blockade, the recommendation on the President’s desk was that Soviet vessels continuing toward Cuba “be stopped and searched in accordance with the Proclamation.” The memo continued, “Should a vessel refuse to comply with orders to stop, we should proceed to execute the terms of the Proclamation with special attention to the minimum use of force.” American officials at the highest levels talked through how these incidents could play out. Officials at the highest levels paid careful attention to the Rules of Engagement for an interception. According to Secretary of State Rusk,

The principal aim of quarantine operations, to keep further nuclear weapons out of the island, should be attained with a minimum of force. Therefore, in case of hostile conflict, there would

---

139 I have not seen evidence that the White House sanctioned the full extent of this harassment.
141 “Rostow to President Kennedy,” October 24, 1962 [FRUS #57].
142 John McCone, “Memorandum for the Files [Executive Committee Meeting on October],” October, 23 1962 [FRUS #51].
be only one “wounding shot” after the usual “warning shot” to affect the ships’ maneuverability but not to sink them if possible.\textsuperscript{143}

Although McNamara at one point threatened Soviet Ambassador Dobrynin, saying the United States “would not stop short of sinking Soviet ships which are bringing offensive weapons to Cuba,” this seems likely to have been bluster.\textsuperscript{144} Regardless of his view, the intent of the Rules of the Engagement was to stop the ships with minimum force, “disabling” only if required to search the ship, and “not sinking ships unless absolute necessary.”\textsuperscript{145} The strategic logic was predicated on finding an effective way to stop the Soviet ship while sidestepping the Soviet red line against an attack destroying their vessels, thereby minimizing the risks of war. Even as he remained open to an air strike, the Attorney General explicitly feared that if the United States navy fired on even one Soviet tanker, “The balloon would go up.”\textsuperscript{146} The fact that disabling falls into a gray area on the edge of the attacking or destroying is critical here, because it does not unmistakably fall into those latter categories. American policymakers feared the consequences of a disabling shot, but not nearly so much as the consequences of deliberately sinking a Soviet ship. In more detail, the Rules of Engagement ordered the following:

- Visit and search procedure
- Non-compliance will result in stopping, boarding, and seizing of the ship for diversion into selected port.
- The blockade would be carried out with minimum use of force … However, enforcement of the blockade may require force.
- [After a signal to stop] Failing to stop … would result in a shot being fired across the bow.
- If forced to engage, an attempt would be made to stop the ship by damage in non-vital areas, such as the rudder, and with minimum loss of life.
- If necessary to destroy the ship, warning of intentions will be given, time will be provided for debarkation of passengers and crew, and maximum help, in accordance with operational conditions, would be rendered.\textsuperscript{147}

Although the last point (above) leaves open the possibility of firing to destroy the ship, it never specifies circumstances in which this would occur for an unarmed merchant ship. Given that omission and the clear order to disable first, the intent to avoid overt attack is clear. The rules of engagement sanctioned the use of force only under these circumstances: “hostile interference” with the inspection or “actions which can reasonably be considered as threatening a U.S. ship.” Neither description seems

\textsuperscript{143} This quotation is from a West German source because it came as Rusk briefed European allies. “Knappstein to the German Foreign Ministry,” October 24, 1962 [WC].
\textsuperscript{144} “Dobrynin to the Soviet Foreign Affairs Ministry,” October 23, 1962 [WC].
\textsuperscript{145} “Houser to Taylor,” October 19, 1962 [GWU NSA].
\textsuperscript{146} “Summary Record of the Seventh Meeting of the Executive Committee of the National Security Council,” October 27, 1962 [FRUS #90].
\textsuperscript{147} These are direct quotations, but intervening text has been removed and the bullets were added. “Riley to the Deputy Secretary of Defense [“Rules of Engagement”],” October 22, 1962. (Attachments) [GWU NSA]. Also see “Transcript of Call between Kennedy and Gilpatric,” October 23, 1962 in Widmer and Kennedy, Listening In.
to apply to actions by unarmed merchant shipping.\textsuperscript{148} And in fact, the orders added, “Armed merchant ships suspected of carrying prohibited material will be treated in the same manner as an unarmed merchantman.”\textsuperscript{149} Overall, the orders were to try to board and to disable if necessary, but only to fire on if attacked.

Other documents suggest a different and perhaps more hawkish view of a quarantine intercept. Some suggested that these incidents could be used to harass the Soviet Union, sending a firmer signal of resolve. One Defense Department memo explored options for upping the pressure on the Soviets as a form of leverage, proposing that an intercept “gives us the opportunity of seizing one or more ships at will on basis of ‘misinterpretation’ of their declarations.” Yet even this document evinced a clear concern for provoking without overtly attacking, mainly recommending measures such as crossing the paths of Soviet ships to force them to turn to avoid collisions and training guns on them without firing.\textsuperscript{150} The Joint Staff produced a number of documents along these lines, i.e., outlining a policy considerably more aggressive than that favored in the White House. Other documents expressed fears that the Soviets would try to run the blockade in order to provoke an incident that would justify retaliatory measures, e.g., in Berlin.\textsuperscript{151} Soviet leaders were not thinking along these lines.

Although the documentary record is thinner on the Soviet side, unfortunately lacking a detailed discussion in which Soviet leaders predict how an intercept would play out, Soviet documents reveal intense fears about how an intercept would go. The Soviet ships carrying nuclear weapons sailed with concealed 37mm machine guns. Soviet ships with such sensitive cargos were also fitted with explosive charges and given orders to self-destruct rather than allow American forces to seize the weapons.\textsuperscript{152} These sorts of steps, ordered before the crisis began, were countermeasures to boarding operations by the U.S. Navy. They offered no meaningful defense against an American decision to simply destroy Soviet ships. Soviet soldiers sailing under these orders did so with great trepidation, because their orders put them in an unenviable situation if up against a boarding operations. The options: scuttle the ship before surrendering or fire first on the boarders, violating that focal point first, and then lose the resultant engagement. The fact that these options were so unpalatable – to both Soviet sailors and Soviet leaders – offers the most compelling explanation for the Soviet reluctance to challenge the blockade. The Soviet Union knew it would need to shoot to kill first to win an intercept encounter and so believed it would not win.

The success of the quarantine came down to fact that the United States had the upper hand in an intercept encounter. That upper hand took the form of being able to stop Soviet ships without overtly firing on them, which neither side wished to do. Disabling and/or boarding fell in a gray area in the Soviet red line, and the United States both sides understood that the U.S. Navy would exploit this

\textsuperscript{148} The exception might be a Soviet ship carrying sensitive materials with orders not to allow them to fall into American hands. Such a ship might prefer to evacuate most of its crew and then scuttle or be destroyed.\textsuperscript{149} Ibid.\textsuperscript{150} “Blockade Measures,” n.d. [25 October 1962] [GWU NSA].\textsuperscript{151} “Smith to McCone [Soviet Challenge to the Quarantine],” October 23, 1962.\textsuperscript{152} “Malinovsky to Pliev,” October 25, 1962 [WC]; Nikolai Beloborodov, “The War Was Averted,” Early 1990s [WC].

161
opening. Consequently, the United States was likely to prevail in an intercept encounter. This was so not merely because the U.S. Navy was so much larger than the Soviet Navy, but because of the weakness in the Soviet red line protecting their ships. The decision not to run the blockade was a decision to avoid confrontation where the United States could prevail without shooting first.

**Signaling Resolve and the Soviet Decision to Withdraw**

The success of the blockade in forcing the Soviets to “blink” did not win the Cuban Missile Crisis for the United States. The blockade, in itself, did not remove the missiles. Instead, the United States made a compellent threat demanding the withdrawal all “offensive weapons” from Cuba and successfully signaled its resolve to carry out the threat.

What convinced the Soviet Union that the United States was likely to use force to remove the missiles from Cuba if no concession was forthcoming? Neither side accomplished this in the Berlin case, so how did it happen here? In this section, I analyze four potential causes of this perception. Because the Soviet Union believed that the United States would relent to the deployment prior to the crisis, I exclude all explanations that do not involve American actions taken during the crisis; only these actions can explain the sudden change in Soviet perceptions. First, did the general mobilization of American military forces and/or their forward deployment to the Caribbean convince the Soviet Union of American resolve? Second, did the firm public statements of the Kennedy Administration signal resolve effectively? Third, did the quarantine do so? Finally, did intensive American reconnaissance overflights play a role in changing Soviet perceptions?

Overall, the Soviet decision to withdraw the missiles comprises a unique observable implication for the conventional wisdom’s emphasis on signaling in crisis strategy. In his October 22nd letter to Khrushchev, President Kennedy wrote, “In our discussions and exchanges on Berlin and other international questions, the one thing that has most concerned me has been the possibility that your Government would not correctly understand the will and determination of the United States.”

The evidence does not indicate that any single method of signaling was itself sufficient, but traditional types of signals such as military mobilization and firm public statements appear to have been – at minimum – necessary ingredients in the success of signaling. This conclusion is not surprising. The conventional wisdom draws its inspiration from this case and specifically this event more than any other.

Nonetheless, red lines theory makes two significant contributions to better understanding the Soviet decision to withdraw the missiles. First, the threat whose credibility the United States established was not a threat to start a general war – a threat that would not have been perceived as credible – but rather the threat to strike Cuba in a *fait accompli*. As it protected an isolated, distant island, the Soviet extended deterrent red line around Cuba was quite weak with respect to the

---

153 “Kennedy to Khrushchev,” October 22, 1962 [FRUS #44].
geography (as Chapter Four implies). It relied instead on the firing-on-forces focal point. Soviet leaders harbored doubts about whether this was sufficient to deter, because American leaders could hope to get away with using force in Cuba without escalation to general war in Europe. Some Americans did hold this belief that the Soviet Union would ultimately relent to a *fait accompli* strike.

Second, the reconnaissance overflights proved to be a surprisingly important part of the brinksmanship which made American signals of resolve convincing. This occurred due to the intersection of the conventional emphasis on brinksmanship-type signaling and red lines theory’s emphasis on the significance of overt uses of force, even small and localized ones. American overflights put great pressure on Soviet units and the Cuban government to open fire, which they did at times even without orders to do so from Moscow. Although Cuba held the legal right to fire on planes violating its airspace, Soviet leaders feared that the United States would react to the violation of this focal point even against planes that were themselves violating Cuban sovereignty. Worried about possible American retaliation for the downing of one of their planes and difficulty of putting a lid on escalation once the firing-on-forces focal point was crossed, Soviet leaders decided to withdraw the missiles rather than allow these risks to continue. By risking violations of that focal point, the reconnaissance flights were integral to American brinksmanship. The mechanism underpinning brinksmanship relied on the existence of a focal point and the ability to incur a calculated uncontrollable risk that it would be violated. The risk of a small violation of this focal point provided the element of chance in the “threat that leaves something to chance.”

**THE CHANGE IN SOVIET PERCEPTIONS**

On October 21st, Khrushchev expected the United States to relent to the missile deployment. On October 27th, Khrushchev saw an American attack as sufficiently likely and imminent that he ordered an immediate withdrawal of the missiles without waiting to consult with the Cuban leadership. In those few days, the United States signaled resolve effectively enough to bring about this change. Unknowingly underscoring how much Soviet perceptions were about to change, Soviet Foreign Minister Gromyko reported back to Moscow on an October 19th meeting with President Kennedy. He described the American position on Cuba as “completely satisfactory,” adding that it is possible to say that “a USA military adventure against Cuba is almost impossible to imagine.”

On the 19th, Kennedy knew of the missiles, but Gromyko did not know that Kennedy knew. Consequently, Gromyko glibly repeated Soviet denials about offensive weapons in Cuba.

Between the 22nd and the 28th, Soviet perceptions changed markedly. By the 25th, Ambassador Dobrynin had concluded, “Judging from available data, the administration sets itself, as a minimal aim, the object of not allowing the emplacement in Cuba the aforementioned missile launchers.” He regarded airstrikes as a real possibility. Castro cabled Khrushchev on October 26th that “aggression in

---

154 “Gromyko to the CC CPSU,” October 19, 1962 [WC].
the next 24 to 72 hours is almost inevitable."156 Eastern bloc embassies in Washington burned their sensitive documents as a precaution in case of war. 157 By the 28th, Khrushchev ordered the immediate removal of the missiles. The U.S. Embassy in Moscow explained it this way, “We believe Khrushchev personally convinced situation was critically dangerous and had to be liquidated. Factors producing dramatic decision withdraw included ... striking display US determination.”158

MOBILIZATION AND FORWARD DEPLOYMENT

The mobilization of American forces for an attack on Cuba and their forward deployment to the Caribbean theater was a necessary but not sufficient part of the American success at signaling resolve. Without mobilization, no threat of attack would be credible. Mobilization alone did not establish credibility.

Without large-scale the mobilization and deployment of the forces needed for an airstrike and an invasion, American threats would have lacked credibility. An October 23rd remark by Soviet Defense Minister Malinovsky reflects this: “I do not think that the USA right now could embark on blitzkrieg operations... If an invasion of Cuba will be declared, this will be after another 24 hours has passed in order to get ready.”159 Often relying on the American media, Soviet and Cuban intelligence watched American force levels in the Caribbean with great care and from the standpoint of regarding an invasion as unlikely until forces reached levels they deemed sufficient.160 On the American side, the value of mobilization was clear to all. Even Maxwell Taylor, a hawk so determined that he was involved in searching for ways to rig a provocative incident to justify an invasion, advocated holding off on airstrikes until sufficient military capability was in place.161

However, American policymakers discussed mobilization measures much more in terms of their necessity for enabling military options than in terms of their signaling value. Although the United States quickly concentrated a massive force in the theater, some mobilization measures essential for an invasion had not occurred by the time that Soviets agreed to remove the missiles. Because a massive campaign of airstrikes would precede any invasion to soften the defenses, there was no military need to have all the invasion pieces in place, despite the signaling value of doing so. As early as October 16th, McNamara anticipated most mobilization measures taking place during or after airstrikes.162 As late as October 27th – hours before the Soviet Union agreed to remove the missiles – McNamara met with the

156 “Castro to Khrushchev,” October 26, 1962 [WC].
158 “The Embassy in the Soviet Union to the Department of State,” October 31, 1962 [FRUS #121].
161 “Minutes of the 505th Meeting of the National Security Council, Washington, 2:30-5:10 p.m.,” October 20, 1962, [FRUS #34].
162 “White House Meeting Transcript,” October 16, 1962 [FRUS #18].
President to receive approval to call up forces needed for the invasion, including 24 air reserve squadrons.\textsuperscript{163} The mobilization was designed and conducted by the military to support the military operations under consideration – impressing (i.e., intimidating) the Soviet Union was beneficial but not the primary aim.

Curtis Lemay, perhaps the most hawkish policymaker in Washington, ordered improved readiness levels for U.S. nuclear forces at the Strategic Air Command (SAC) and aimed for a maximum readiness posture. Although Lemay expressed his view that the “readiness of U.S. military forces comprises a clear warning to the Sino-Soviet Bloc,” he also ordered that the increase in readiness be carried out “quietly and gradually,” which hardly seems in line with a signaling strategy.\textsuperscript{164}

The Soviet Union carefully monitored American military readiness and deployments. On October 23\textsuperscript{rd}, Ambassador Dobrynin briefed his government, “For the practical implementation of the quarantine in the area of Cuba, there has been assembled, according to the reports of military observers, around 450 military ships, more than 1,200 airplanes and around 200 thousand soldiers.”\textsuperscript{165}

On the other side of the crisis, Soviet forces in Cuba prepared for war.\textsuperscript{166} Soviet forces had orders to “undertake urgent measures to increase combat readiness, and to repel the enemy by joint efforts of the Cuban army and all units of the Soviet troops, excluding the [ballistic missiles and nuclear warheads].”\textsuperscript{167} Soviet strategic forces in Cuba ramped up efforts to achieve readiness to launch.\textsuperscript{168} Soviet forces in Eurasia increased their readiness levels.\textsuperscript{169} The Cuban military mobilized in full, 350,000 men in total. Soviet documents underscored the toll that prolonged full mobilization would take on the already frail Cuban economy.\textsuperscript{170} Although this level of cost is generally thought to make signals more credible, in this case it was immaterial. American policymakers already assumed that Castro would resist an invasion to remove his regime with all means at his disposal.

Was American mobilization a credible “costly signal”? The Joint Staff described the material costs, “Maintenance of current alert levels degrades training and equipment maintenance in the units concerned. If maintained for protracted periods, the capability of the force may be reduced...”\textsuperscript{171} Although a nontrivial set of costs, when set against the stakes of nuclear weapons in Cuba and the

\textsuperscript{163} “Transcript of Conversation between President Kennedy and McNamara,” October 27, 1962 in Widmer and Kennedy, Listening In; “Summary Record of the Ninth Meeting of the Executive Committee of the National Security Council,” October 27, 1962 [FRUS #97].

\textsuperscript{164} “Lemay to Taylor,” October 22, 1962 [GWU NSA].

\textsuperscript{165} “Dobrynin to the Soviet Foreign Ministry,” October 23, 1962 [WC].

\textsuperscript{166} “Central Committee of the Communist Party of the Soviet Union Presidium Protocol 60,” October 23, 1962 [WC].

\textsuperscript{167} “Malinovsky to Pliev,” October 22, 1962 [WC].


\textsuperscript{170} “Aleksseev to Soviet Foreign Ministry,” October 23, 1962 [WC].

\textsuperscript{171} Joint Staff, “Military Aspects of the Cuban Situation,” November 16, 1962 [GWU NSA].
potential for a nuclear war, it is hard to believe that the costs of mobilization for weeks or months did anything other than pale in comparison. 172

Overall, mobilization was an essential element in the success of American signaling, but it was not alone sufficient. As recently as April of 1962, the United States conducted a massive military exercise (Lantphibex-62) simulating an invasion of Cuba that included 40,000 men and 82 ships. 173 Another such operation was scheduled for later in 1962. Although these actions in isolation were perceived as threatening in Moscow and Havana, it required more than these actions alone to establish the credibility of the threat to attack large numbers of Soviet troops during the crisis.

PUBLIC STATEMENTS

As with mobilization, public statements of resolve were a necessary but not sufficient ingredient in the success of signaling. President Kennedy’s speech of October 22nd announcing the quarantine was the primary such statement, but it was supplemented by many others and by official private statements (the contents of which could easily become public) between the United States and the Soviet Union. President Kennedy understood and remarked upon the value of taking a hard line in public comments in order to signal resolve, even if the intent was to eventually soften the position to reach a deal. 174 When State Department spokesman Lincoln White exceeded his instructions by saying that continued construction of the missiles could mean that “further action will be justified,” Kennedy was initially angry before quickly coming to see advantage of the comment. 175

The Cuban Missile Crisis yields a certain amount of evidence in favor of the salience of audience costs. In a major critique of audience costs theory, Snyder and Borghard cite this case as evidence for the conclusion, “Authoritarian targets of democratic threats do not perceive audience costs dynamics in the same way that audience costs theorists do.” 176 However, writing to Moscow on October 25th, Ambassador Dobrynin begged to differ,

A certain danger of the situation is that the President has largely engaged himself before the public opinion of America and not only America. In essence, he, as a hot-tempered gambler, has put at stake his reputation as a statesman and politician, and thus his prospects for re-election in 1964, what--being an ambitious man--he passionately seeks. This is why it is not possible to exclude completely the possibility that he can, especially taking into consideration his circle, undertake such an adventurist step as an invasion of Cuba. 177

172 Mobilization – like public statements – also may have raised the domestic political costs of backing down for President Kennedy. The next section, although focusing on the statements, considers this type of signal.
173 Fursenko and Naftali, One Hell of a Gamble, p. 166.
174 “Minutes of the 506th Meeting of the National Security Council,” October 21, 1962 [FRUS #38].
175 Sorensen, Kennedy, p. 712.
Although this Dobrynin cable provides a clear instance of an authoritarian state perceiving audience costs in line with the theory, it is not difficult to find a senior policymaker making emphatic statements like this in favor of one or another type of signal or motive. The statement suggests a role for audience costs, but not necessarily a decisive one. Dobrynin wrote a separate cable reflecting on Kennedy’s speech announcing the quarantine in which he concluded that strategic considerations superseded domestic politics:

On the domestic political plane, Kennedy obviously is counting on his last step [the blockade] to pull the rug out from under the legs of the Republicans, whose leadership in recent days officially announced that they consider the Cuban issue a fundamental issue of the election campaign, having in essence accused the administration of inactivity on that issue. However, it is necessary to stress that the events connected with Kennedy’s announcement yesterday obviously have overtaken the significance of electoral considerations and that these considerations now are moving to the background.\(^{178}\)

Robert Kennedy was quite open with the Soviet government about the domestic political punishment that awaited him and his brother if they backed down over Cuba. Explaining why he would not put on paper the agreement concerning the withdrawal of Jupiter missiles from Turkey, Robert Kennedy said the following to Ambassador Dobrynin,

Speaking in all candor, I myself, for example, do not want to risk getting involved in the transmission of this sort of letter, since who knows where and when such letters can surface or be somehow published--not now, but in the future--and any changes in the course of events are possible. The appearance of such a document could cause irreparable harm to my political career in the future.\(^ {179}\)

Both the Soviets and the Americans took heed of the extent to which President Kennedy staked his reputation on removing the Soviet missiles from Cuban soil. So too did Congress and the American media. Even absent these statements, Kennedy stood to suffer at the ballot box for allowing the introduction of Soviet missiles to Cuba even had he made no statements prohibiting it or demanding they be withdrawn. Nonetheless, the general assumption seems to have been that his statements raised these costs to some degree. At a more basic level, American threats to use force to remove the missiles would not have been credible without strong public statements demanding that outcome.

THE BLOCKADE

It is difficult to credit the blockade for the success of American signaling, because – much like the airlift – the blockade was a signal of restraint as well as a signal of resolve. On the 22\(^{\text{nd}}\), Soviet


leaders learned that President Kennedy would give a major television address relating to Cuba amidst an atmosphere in Washington of sudden tension and crisis. Soviet leaders immediately feared the worst – the announcement of an airstrike, an invasion, or an explicit ultimatum threatening them. Upon learning that the announcement was a blockade, the reaction – while still laden with deep concern – also involved an element of relief that the United States had foregone those more aggressive options.  

The American choice for a blockade had two main advantages. It conveyed to the Soviet Union that the United States continued to desire peace and it left leeway for a negotiated settlement. The downside, however, was revealing reluctance to use force against the missiles. This called into question whether the United States would become willing to do so if coercive bargaining failed. Ambassador Dobrynin conveyed what he referred to as the “American fixation” with the blockade as a “test of strength” to Moscow, but there was no escaping its status as a signal of the unwillingness to immediately remove the missiles in a fait accompli airstrike.  

The White House feared this perception. Alongside his public speech on the 22nd, Kennedy wrote a letter to Khrushchev which said in part,

I wish to point out that the action we are taking is the minimum necessary to remove the threat to the security of the nations of this hemisphere. The fact of this minimum response should not be taken as a basis, however, for any misjudgment on your part.

In the White House, confidence was low that the blockade would send a sufficient signal of resolve as opposed to a mixed signal of resolve and restraint. Writing on October 25th, Dean Rusk summed up the situation immediately after Soviet ships turned back,

[The Soviet Union] is probably waiting (1) to see how much momentum will develop for a trade involving US strategic bases, and (2) whether the US is prepared to exert further force to achieve removal of missiles already in Cuba. Until it obtains greater certainty on these points, it will probably keep its missiles in place.

In hindsight, this analysis is largely accurate. The blockade may have contributed to the success of American signal at the margins, but it was far from a silver bullet.

RECONNAISSANCE OVERFLIGHTS

180 Sergei N. Khrushchev, Nikita Khrushchev and the Creation of a Superpower (University Park: Penn State University Press, 2010), pp. 555, 559.
181 “Dobrynin to the Soviet Foreign Ministry,” October 23, 1962 [WC]. Dobrynin was correct on this point. According to Dean Acheson, “We were involved in a test of wills, and the sooner we got to a showdown the better.” “Record of Meeting, Washington, 11 a.m.,” October 19, 1962 [FRUS #31].
182 “Kennedy to Khrushchev,” October 22, 1962 [FRUS #44].
183 Hilsman to Rusk, October 25, 1962 [FRUS #76].
Perhaps the most surprising contributor to the credibility of American signaling was the unrelenting series of reconnaissance overflights of Cuba that began on October 22\textsuperscript{nd}. Unlike the more traditional policy tools for signaling — mobilization, statements, etc. — reconnaissance flights receive little mention in the literature. Their significance speaks both to the utility of brinksmanship in this case and the role of strong red lines (focal points) therein.

More than any other U.S. policy, these overflights created the risk of an accidental overt use of force. Because Khrushchev mistrusted his ability to ensure restraint from Soviet and (especially) Cuban personnel manning anti-aircraft batteries, he feared that the overflights would produce an overt use of force. With the critical red line against escalation crossed, the United States would be likely to retaliate and no equivalent barrier to additional escalation would exist.

The two theoretical frameworks therefore collide in the form of these reconnaissance overflights. Because the overt attack line has such salience, it both creates a great deal of strategic maneuver to make gains without crossing it and endows brinkmanship-type actions which risk its violation with credibility as signals of resolve.

As of the morning of October 27\textsuperscript{th}, it would have been reasonable for the Soviet Union to hope that an overall favorable stalemate had with some degree of stability had emerged. The United States had successfully blockaded Cuba, but was reluctant to launch an outright attack. The Soviets had every reason to stall as the missiles neared completion. Then something changed.

On October 27\textsuperscript{th}, a Soviet surface-to-air missile site fired two missiles, downing an American U-2 spyplane and killing Major Rudolf Anderson, Jr. Soviet forces reported this to Moscow, describing their intent in firing the missiles as preventing images of the missiles from reaching the United States. Soviet General Stepan Grechko ordered the missiles fired on his own initiative, but Washington did not know this. Soviet Defense Minister Malinovsky reproached his subordinates in Cuba for this decision, “We believe that you were too hasty in shooting down the US U-2 reconnaissance plane.” This same cable ordered the withdrawal of the missiles. Later that day, Malinovsky sent a follow-on cable, “In addition to the order not to use S-75s [SAMs], you are ordered not to dispatch fighter aircraft in order to avoid collisions with US reconnaissance planes.”

The strongest evidence in favor of the importance of the overflights in changing Soviet perceptions comes from the accounts of Sergei Khrushchev, son and biographer of Nikita Khrushchev.

---


\textsuperscript{185} I intend no criticism here. The role of these flights may be unique to this case.

\textsuperscript{186} The harassment of Soviet submarines is the exception to this claim, but note that this harassment likely exceeded President Kennedy’s intent.

\textsuperscript{187} “Ivanov and Malinovsky to Khrushchev,” October 27, 1962 [WC].

\textsuperscript{188} “Malinovsky to Pliev,” October 28, 1962 [WC].

\textsuperscript{189} “Malinovsky to Pliev,” October 28, 1962 [WC].
He describes his father’s reaction to news of the shootdown: “It was at that very moment – not before or after – that Father felt the situation slipping out of his control.” Khrushchev was furious that this action had been taken without his orders. He feared how the United States would respond, and the possibility that further firing without high-level orders (e.g., by the Cubans) would lead to disaster. Sergei Khrushchev regards this moment and these fears emanating from the shootdown incident as the tipping point for Soviet policy.

How would the United States react to the downing of one of its aircraft – aircraft undeniably violating Cuban airspace? Contingency planning on October 23rd called for a proportional response: the destruction of the responsible SAM site in retaliation. Briefings given to NATO allies described this as American policy. McNamara advocated this policy only hours before learning of Major Anderson’s death. General Taylor reported the U-2 shootdown on the afternoon of the 27th along with a recommendation for an airstrike on that SAM site the next day, in line with the stated policy. Both he and McNamara expected this strike to escalate to a larger aerial campaign and – most likely – an invasion.

Fearful that the U-2 shootdown and a retaliatory airstrike would lead to uncontrolled escalation, President Kennedy decided to give diplomacy one last chance. Even if the initial attack was limited to the SAM site and thus proportional, the attacking planes would likely come under attack, and escalation could continue from there. In a White House meeting on the evening of October 27th, Kennedy ordered retaliatory attacks on any SAM sites firing on American aircraft the next day provided that diplomatic negotiations did not appear to be about to reach fruition. He decided, “If they fire on us, tell them we’ll take them out.” Although Kennedy did not immediately retaliate for the first shootdown, the United States was very much on the brink of retaliatory strikes in its aftermath. As with the Soviet side, fears of imminent unwanted escalation peaked at this moment and for this reason.

---

190 Note that it is difficult to analytically separate the concerns about the shootdown itself and fears about future aggressive actions without orders. Sergei N. Khrushchev, Nikita Khrushchev and the Creation of a Superpower (University Park: Penn State University Press, 2010), p. 608.
192 “Minutes of Meeting of the Executive Committee of the National Security Council,” October 23, 1962 [FRUS #47].
193 “Knappstein to the German Foreign Ministry,” October 24, 1962 [WC]. There is no evidence suggesting that Soviet intelligence had knowledge of this contingency planning.
194 “Summary Record of the Eighth Meeting of the Executive Committee of the National Security Council,” October 27, 1962 [FRUS #94].
195 Ibid.
196 “Summary Record of the Ninth Meeting of the Executive Committee of the National Security Council,” October 27, 1962 [FRUS #97].
197 “Transcript of Conversation between President Kennedy and McNamara,” October 27, 1962 in Widmer and Kennedy, Listening In.
Sensitivity to the risks of these overflights predated the crisis. In September, during the planning that led to the flights which initially spotted the missiles, Lyman Kirkpatrick asked in a memo, “Is there anyone in the planning of these missions who might wish to provoke an incident?” On October 26th, at the height of the crisis, a Joint Staff memo stated, “They [reconnaissance pilots] should be encouraged to maneuver their planes as a fighter bomber pilot might be expected to do in attacking a site.” Just as Khrushchev came to doubt his control over whether to fire on the overflights, supporters of an attack on Cuba within the Pentagon sought to make these flights as provocative as possible in order to engineer a violation of the firing-on-forces focal point.

The Soviet government repeatedly butted heads with the Cuban government over the question of whether to shoot at American planes, which the Cubans considered a humiliating violation of their hard-won sovereignty. On October 24th, Castro argued that “it is necessary to start to fire some shots, because the ... [Americans] are flying within our territory, at 200 feet.” Castro’s desire was relayed to Moscow, but the Kremlin disagreed and asked him to “show patience, restraint, and more restraint” and to “not let yourselves be provoked ... by frenzied military men in the Pentagon.”

Castro was sufficiently enraged by the Soviet decision to withdraw the missiles that he ordered Cuba anti-aircraft gunners to fire. According to Castro, “It is difficult to explain to our people this concession to the enemy. ... All we need now is for American planes to land on our territory to refuel. And what are we doing? We are enabling them. In effect, we are allowing the enemy to violate our airspace.” The Soviet Union was eventually able to again pressure him into restraint on that issue. In fact, not firing at American aircraft over Cuba was of such importance to Khrushchev that he was potentially willing to abandon the Soviet-Cuban relationship to get his way:

If the Cuban comrades do not want to cooperate with us on this issue and do not want to undertake measures which would help us resolve this issue and avoid being pulled into a war together with us, then apparently the conclusion that we see is that our presence in Cuba is not helpful for our friends now. Then let them state that openly, and we will have to make conclusions for ourselves.

199 “Taylor to Riley,” October 26, 1962 [NSA GWU].
201 “Record of Meeting between Fidel Castro and Military Chiefs,” October 24, 1962 [WC].
204 “Record of Conversation between Mikoyan and Castro,” November 13, 1962 [WC].
Khrushchev’s instructions to Mikoyan with respect to persuading the Cubans to forgo firing on American reconnaissance aircraft are particularly revealing:

We believe—and this is very important—that, even if they opened fire against the American aircraft, and we would regret if such a development occurred, if that would have been done, that fire would not be effective. It would not result in a real strengthening of Cuban security by military means. But it could cause an onset of U.S. military actions against Cuba.\(^{206}\)

Khrushchev brings to light the most informative aspect of the importance of the issue of firing on these overflights. He stipulates that shooting down one plane would not meaningfully add to Cuban security. The point extends to Soviet interests and even to American interests – one plane and one life, while precious, pale in comparison to the risks of nuclear war. Yet despite the lack of a vital intrinsic interest at play in a shootdown incident, Khrushchev saw the risk of provoking a war as grave.

Concurrent to his decision to withdrawal of the missiles, Khrushchev also requested in writing that the United States desist from these flights, “I should like you to consider, Mr. President, that violation of Cuban airspace by American planes could also lead to dangerous consequences. And if you do not want this to happen, it would be better if no cause is given for a dangerous situation to arise.”\(^{207}\) Independent of this request, President Kennedy ordered that no air reconnaissance take place on October 28\(^{th}\) in order to prevent an incident that could disrupt the deal.\(^{208}\) The United States resumed intensive reconnaissance flights shortly thereafter once tensions eased with the aim of confirming the Soviet withdrawal.\(^{209}\)

Robert Kennedy’s October 27\(^{th}\) ultimatum to Ambassador Dobrynin during their pivotal meeting is now well-known, although Kennedy made sure to describe it as a request and not an ultimatum. The exact context in which he said it has received less attention, and lends some support to the importance of the overflights:

I explained to him that in the last two hours we had found that our planes flying over Cuba had been fired upon and that one of our U-2’s had been shot down and the pilot killed. ... I told him that this was an extremely serious turn in events. We would have to make certain decisions within the next 12 or possibly 24 hours. There was a very little time left. If the Cubans were shooting at our planes, then we were going to shoot back. This could not help but bring on further incidents and that he had better understand the full implications of this matter.\(^{210}\)

\(^{207}\) “Khrushchev to Kennedy,” October 28, 1962 [FRUS #102].
\(^{208}\) “Summary Record of the Tenth Meeting of the Executive Committee of the National Security Council,” October 28, 1962 [FRUS #103].
\(^{209}\) Castro’s refusal to tolerate inspections made this necessary
\(^{210}\) “Robert Kennedy to Rusk,” October 30, 1962 [FRUS #96].
Dobrynin reported this conversation to Moscow – perhaps the single most important moment of the crisis – in a similar but not identical manner:

The Cuban crisis, R. Kennedy began, continues to quickly worsen. We have just received a report that an unarmed American plane was shot down while carrying out a reconnaissance flight over Cuba. The military is demanding that the President arm such planes and respond to fire with fire. The USA government will have to do this. ... Because of the plane that was shot down, there is now strong pressure on the president to give an order to respond with fire if fired upon when American reconnaissance planes are flying over Cuba. The USA can’t stop these flights, because this is the only way we can quickly get information about the state of construction of the missile bases in Cuba, which we believe pose a very serious threat to our national security. But if we start to fire in response—a chain reaction will quickly start that will be very hard to stop. 

Both accounts agree that the meeting began with a discussion of the U-2 shootdown. Dobrynin’s account includes a sustained discussion of terms before Kennedy requests an urgent reply the next day. In contrast, Kennedy’s account moves directly from the U-2 discussion to the deadline for a reply. Dobrynin’s account is the more reliable – most of the difference stems from Kennedy’s deliberate omission of the intervening discussion of removing the missiles in Turkey. Dobrynin’s account, however, still bears out the heavy emphasis on overflights and the potential for a shootdown incident in this pivotal meeting. Far from being a minor part of the crisis, the U-2 shootdown was a potential catalyst for escalation that made the imminent use of force more likely.

Conclusion

The conventional wisdom is right in regarding the Cuban Missile Crisis as a milestone success for signaling. By and large it also represents a success for compellence, with the victory tempered only by the concessions the Americans gave in return for the withdrawal of the missiles. This is exactly the sort of coerced concession that Chapter Four shows to be quite rare in territorial crises. Nonetheless – and despite the status of the Cuban Missile Crisis as the flagship case the flagship case that motivates so much of the conventional wisdom – red lines theory uniquely explains key aspects of the case.

Red lines theory makes four main contributions. First, the crisis started because Khrushchev perceived an opening to sneak missiles into Cuba undetected. Covertness was so integral to the Soviet decision to deploy the missiles by fait accompli that it is hard to envision the crisis occurring without the element of unverifiability in this American red line.

Second, the opinions of President Kennedy and the majority of his advisors shifted from their initial inclination for an airstrike to the decision for a blockade due to the fear that killing Soviet soldiers

would result in escalation to war. The blockade was a strategy with an obvious – seemingly crippling – deficiency: it did not remove the missiles. Yet it was the most aggressive course of action available that respected the Soviet red line against an overt attack on their forces. That consideration won the day.

Third, the decisive tactical consideration underpinning the success of the American blockade was the ability of the U.S. Navy to prevail in an intercept encounter without having to overtly attack (to kill or destroy first). The options to board or disable Soviet ships without sinking them was crucial – an option that the Soviet Union lacked against the Berlin Airlift. For this reason, Soviet submarines and aircraft continued to reach Cuba during the blockade while the freighters failed.

Finally, red lines theory provides a piece that fits into literature’s understanding of brinksmanship. That piece concerns the question of why one accidental incident can lead to such costly escalation when neither side desires war. U.S. reconnaissance overflights generated much more risk of nuclear war than the value of individual planes or pilots would suggest. The problem was that one the firing-on-forces focal point was crossed, there was no logical stopping point in which leaders would have confidence. Moreover, the United States would (and did) experience amplified pressure to respond at least proportionately or risk the broader red line against Soviet attacks on American forces. These considerations magnified to a grave extent the difficulties Khrushchev confronted with preventing both Cuban and Soviet forces from firing on reconnaissance aircraft.

Overall, approaching the Cuban Missile Crisis from the standpoint of red lines theory contributes to a deeper understanding of the case. At both the strategic and tactical levels, the United States and the Soviet Union at times sought to exploit vulnerable red lines. Each side took advantage of opportunities to advance unilaterally without overtly attacking the other. The result was a game largely played by the rule of respecting the firing-on-forces focal point. That rule was broken only once, in the process spurring the two sides to end the crisis before it got out of control. If red lines theory can shed light on important aspects of the flagship case for the crisis signaling, this speaks well of its capacity for aiding in understanding crises in general.
Ch. 7 Conclusion

If crises are “wars of nerves,” how are these wars fought? The prevailing answer is that states rely on brinksmanship and other means of signaling resolve in order to make credible threats and coerce adversaries into adopting the desired policy. This book has explored crises from another perspective. For the challenger, making a gain involves finding the right spot to get away with unilaterally taking something of value by fait accompli. For the deterrer, the task is to deter not just all-out invasion, but to set red lines that deter these smaller-scale aggressions. The strategic logic of red lines and faits accomplis hinges on finding vulnerabilities in red lines. States seeking to protect many small interests at once try to do so by grouping them into one inviolable whole encapsulated by an unmistakable line. Where this line has gray areas or loopholes, however, adversaries find openings to get away with taking unilateral gains. This is a central problem of deterrence at issue in interstate crises.

Case studies of the Berlin Blockade Crisis and Cuban Missile Crisis document the strategic game that unfolds in crises from the perspective of this theoretical framework. Both the United States and the Soviet Union avoided violating strong red lines while seeking out ways to make gains by working around them. The analysis of territorial crises from 1918 to 2007 establishes the existence of what is ultimately a similar dynamic in that domain. States have tended to avoid seizing territory unless they can exploit a gray area to increase their chances of making the gain while escaping war. This chapter concludes by reviewing the findings and identifying some of the most significant implications that flow from them.

Reviewing the Findings

Chapter Two developed a theory of deterrent red lines and faits accomplis. This theory explains why gray areas and other vulnerabilities in red lines make faits accomplis more effective and therefore more common. Here I return to the four hypotheses derived from the theory in Chapter Two and review the evidence that the empirical chapters provide with respect to them.

H1: Vulnerable red lines make faits accomplis more likely.

The connection between gray areas and other vulnerabilities in red lines and faits accomplis targeting them is the strongest prediction of the theory. Many other factors intercede that influence crisis outcomes, but the implications of the theory for crisis behavior has far fewer intervening factors that could mask the relationship. Consequently, this hypothesis is the most valuable for theory-testing purposes.

I find support for this hypothesis in both the case studies and the broader analysis of territorial crises since 1918. The sources of strength and weakness in red lines vary by issue area, and these chapters focus on different types of focal points that make for strong red lines. In a “garden variety” crisis where deterrence aims to stop an advance by the armed forces of the adversary, however, there

175
are three consistent sources of the focal points that make for strong red lines: physical geography (waterways), political geography (established boundaries), and the politically sanctioned use of force against a deterrer’s forces (the firing-on-forces focal point). The analysis of territorial conflicts focuses on the first two, while the third provided the most leverage in the case studies.¹ Deterrent red lines are stronger when the adversary must violate these focal points to make a unilateral gain. These unilateral violations often – but not always – take the form of faits accomplis. A great deal of the “action” in interstate crises, correspondingly, surrounds efforts to make unilateral gains while avoiding overt violations of these focal points.

The Berlin and Cuba case studies demonstrate just how creative states can be when trying to find unilateral ways to advance their interests without having to use force. In Berlin, the Soviet Union declined to attack American forces in Berlin, deciding instead to block the transit corridors in its own territory in East Germany. American forces were deterred from trying to force their way through these Soviet forces blocking the roads, but persevered with the Berlin airlift. The decisive advantage: in the air, unlike on the ground, Soviet forces had to attack first to prevail.

In the Cuban Missile Crisis, the United States chose blockade over airstrike and invasion out of a deep reluctance to attack Soviet forces in Cuba. At the tactical level, the blockade line deterred Soviet ships because freighters could be disarmed or boarded without being destroyed by the U.S. Navy. These actions exploited gray areas in the Soviet red line against overt attack on their forces. In contrast, Soviet aircraft and submarines that could have carried nuclear warheads continued to reach Cuba. As in the Berlin case, aircraft and submarines could be destroyed but not boarded or disabled. Like the Soviet Union in the Berlin Blockade Crisis, the United States was not willing to take that step. Time and again, the strategic logic of these crises was finding unilateral moves that provided some gain or advantage without having to overtly attack.

These cases are no aberrations. Finding and exploiting clever ways to make unilateral gains without firing a shot has been a fixture of interstate crises. Consider the Cod Wars, two crises between Britain and Iceland over fishing rights that should seemingly bear little in common with these two high-stakes Cold War crises between superpowers.² And yet, the same basic dynamics emerged. Reluctant to attack, Icelandic and British ships reverted to the antiquated tactic of ramming rather than shooting – another gray area in the red line against overt attack. Between the two “wars,” the Icelandic Coast Guard invented a new weapon system, the net-cutter. Resembling an anchor in appearance, this device would drag behind a ship before snaring and severing the lines connecting fishing nets to their trawlers.

¹ Because geography is static, it is far less able to provide the unique observable implications that make for strong qualitative theory testing. The geographic status of Berlin as a gray area was likely quite important, but it is difficult to disentangle from any other structural variable (e.g., the value of Berlin). The case studies focus on the firing-on-forces focal point for this reason.

² One interesting commonality across these cases is that the costs of escalation to war overshadowed the value of the stakes.
preventing the fishing without directly attacking the trawler. In the terminology of the theory, this tactic flanked an incomplete red line protecting the trawlers.

In the Fashoda Crisis, Britain and France raced to Fashoda precisely because the first to arrive would have the advantage of putting the adversary in a position of having to fire first to prevail. The same basic dynamic emerged in 1999 when NATO and Russian forces raced for control of Pristina Airport after the end of the Kosovo conflict. Racing to arrive first and thereby compel a challenger to overtly attack to take a territory can even happen on a broader scale. The Spratly Islands are controlled today by a bizarre patchwork of intermingled Chinese, Vietnamese, Filipino, Malaysian, and even (one) Taiwanese outposts. Lying between but separate from these states, the islands are a clear gray area. With geography provided few constraints, the islands were taken and retained by whichever state arrived first – often more in order of size and value than location – with the other disputants then reluctant to attack the occupying forces.

In territorial crises, time and again states target gray areas with land grabs. Gray areas are less than half of all territories disputed in a crises, but two-thirds of land grabs strike a gray area. It is not enough for a territory to be small for it to be vulnerable to a fait accompli. Any disputed territory contains individual small areas that a challenger can seize. Yet land grabs striking non-gray area territories are rare. Overall, the statistical models suggest that changing a disputed territory from non-gray to gray doubles the likelihood of a land grab, holding all else constant.

**H2: Vulnerable red lines make deterrers more likely to endure losses to adversary faits accomplis without retaliating.**

The second hypothesis suggests that faits accomplis striking vulnerable red lines are more effective. States rely on strong red lines to deter most limited predations, while begrudgingly relenting to those that target gray areas. Losing a small island off the coast or a contested region where the border was drawn ambiguously is undesirable, but it does not call into question the willingness to defend the core territory in the way that letting a piece of the core territory go without a fight would. Even where states challenge a fait accompli that made a gain in a gray area, they do so in a manner to keep the conflict limited to that gray area, making escalation less likely.

Whereas the next hypothesis concerns the likelihood of war, this hypothesis examines whether or not the deterrer loses something of value without war. These are the two forms of losing in a crisis. If the challenger’s fait accompli is successful, the deterrer loses the stakes without war breaking out. Because the analysis of territorial crises assesses crisis outcomes only in terms of whether war breaks out, the direct evidence for the second hypothesis comes from the case studies.

---


Both cases largely support the hypothesis, with the important caveat that many factors intercede between the disadvantage created by a vulnerable red line and the final outcome of the crisis. In the Berlin case, the Soviet Union was unwilling to cross the red line against seizing the Western Sectors of Berlin but was able to work around it by blocking the transit corridors. In the words of Army Chief of Staff Bradley, “At present with our passenger trains completely stopped, Russians in effect have won the first round.” If the Soviets had been correct in assuming that no airlift could supply the Western Sectors on an indefinite basis, this advantage might well have led to a favorable outcome. However, to the surprise of both Moscow and Washington, the airlift was logistically feasible. Because the airlift worked around the Soviet red line against pushing through their troops blocking the transit corridors on land, the Soviet Union relented to it. Conversely, the fact that the United States could resupply Berlin without having to push through Soviet forces was crucial for Soviet defeat in the crisis.

In the Cuban Missile Crisis, Washington did not doubt that the Soviet Union’s *fait accompli* deployment of missiles to Cuba gave them an initial advantage in the crisis. President Kennedy characterized it as “a rather dangerous but rather useful play of theirs” that resulted in “a very satisfactory position from their point of view.” This is consistent with the hypothesis. However, the Soviet Union did not then keep the gain as the hypothesis would predict. Instead, the credible threat of U.S. military action coerced the Soviets into a bargain in which they withdrew the missiles but received in exchange a non-invasion pledge and an assurance that the United States would remove its missiles from Turkey after a decent interval.

In sum, the cases support the hypothesis insofar as vulnerable red lines created disadvantages for the states setting them. However, in these cases at least, this disadvantage was not always decisive. The outcomes emerged through a process of strategic interaction with many causes, of which this was but one.

*H3: Faits accomplis that strike vulnerable red lines are less likely to lead to war than faits accomplis that strike strong red lines.*

The analysis of territorial crises shows how salient the geographic gray areas can be in determining whether *faits accomplis* occur and whether they lead to war. States are twice as likely to get away with a land grab *fait accompli* when they seize a gray area territory, i.e. a small island or island-like peninsula located awkwardly between the two core territories or a region where borders were drawn ambiguously.

Although neither the Berlin nor Cuba cases ended in war, both provide important evidence for this hypothesis. This evidence comes from the perceptions of which actions would vs. would not have provoked war. In the Berlin case, American policymakers expected an attempt to push a truck convoy through Soviet forces to drastically increase the chances of war. The airlift, however, worked around this constraint by flanking it through the skies. Had a Soviet fighter or anti-aircraft battery shot down an American aircraft, the result would have been a grave risk of war. The Soviet Union refrained from firing on the airlift.
In contrast, had the Soviet Union ‘shot down’ an American aircraft by firing high-powered searchlights into its pilots’ eyes as they descended to land in Berlin, little would likely have come of it. The Soviet Union did use this tactic, albeit without successfully bringing down an aircraft. Why is shooting down an aircraft with bullets intolerable but shooting it down with beams of light tolerable? The only plausible explanation is that beams of light – not in themselves deadly – fall into a gray area in the focal point against the overt use of force. More generally, dozens of Americans died in plane crashes due to Soviet policy (the blockade), but these deaths posed no significant risk of escalation. An overt attack would have crossed the line and had far more dangerous implications, even if the casualties were no higher.

Takin a larger view of the Berlin Crisis, it becomes clear that the strength of the red lines was a more potent predictor of the willingness to fight than even the proximate interests at stake. The Soviet Union was unwilling to fight or threaten to fight for its primary objective: halting the process of state formation in what became West Germany. The Soviet Union was also unwilling to fight for its secondary objective – control of the Western Sectors of Berlin – if its blockade strategy was negated by airlift. However, if the United States were to have tried to accomplish that same objective of defeating the blockade by truck convoy, the Soviet response was expected to be far stronger. In Washington as in London and Paris, the Soviets were perceived to be willing to fight if overtly attacked on even a small scale, but not for the actual stakes of the crisis in themselves. Similarly, the Soviet Union hoped that the United States would be willing to lose Berlin by pressure from the blockade, but was unwilling to seize it directly or to engage American aircraft breaking that blockade. For both sides, the firing-on-forces focal influenced the willingness to fight more than the value of the immediate stakes.

In the Cuban Missile Crisis, President Kennedy resisted strong pressure to resort to airstrikes and a follow-on invasion because he believed these actions would lead to war. The key to this Soviet red line was the need to spill the blood of Soviet soldiers to impose the desired outcome. In Washington, this was widely understood to be the pivotal red line that meant escalation if crossed. Kennedy opted instead for the blockade and a coercive strategy.

**H4: Vulnerable red lines make war more likely by making faits accomplis more likely.**

Vulnerable red lines make *faits accomplis* more likely. *Faits accomplis* are an aggressive and risky tactic that makes war more likely. Therefore, it seems reasonable to expect that vulnerable red lines make war more likely. The catch, however, is the third hypothesis, which predicts that vulnerable red lines make *faits accomplis* more common but individually less perilous. As Chapter Four shows, these two effects tend to offset each other, resulting in no consistent relationship between gray areas in red lines and the onset of war.

Despite the lack of a correlation between gray areas in red lines and the onset of war, there are cases within the analysis that fit this hypothesis. Without the Falkland Islands, no territorial crisis or war between Britain and Argentina would have occurred. Israel might not have attacked Egypt in 1956 if not
for the ability to limit its operations to seizing control of the Sinai Peninsula. Without border ambiguities, the Chaco War between Bolivia and Paraguay, the Sino-Indian War, and the Badme War between Ethiopia and Eritrea might have been avoided. Although most land grabs striking gray areas have not led to war, these cases suggest that the hypothesis is vindicated in certain cases even if not as a general predictor of war.

In the case studies, policymakers harbored grave fears that a chain of events would unfold along the lines predicted by this hypothesis. Although none did, these fears remain plausible to this day. Throughout the Cold War, American leaders lost sleep over the possibility that the Soviet Union might one day decide it could get away with seizing West Berlin by force in a fait accompli. As an isolated enclave in East Germany, West Berlin is exactly the sort of gray area that has so often fallen victim to a fait accompli across the history of territorial crises. Only the firing-on-forces focal supported the U.S. red line in Berlin.

Similarly, in the Cuban Missile Crisis, the Soviet Union exploited an unverifiable red line to covertly deploy nuclear missiles in a fait accompli that brought the world as close to nuclear war as it has ever come. If the Soviet Union had not stepped back from the deployment, it would have been difficult for President Kennedy to resist the pressure for a forceful response that might have led to war.

On the whole, the evidence for the fourth hypothesis is weaker than the evidence for the first three. Gray areas in particular and weak red lines in general are not reliable predictors of the onset of war. Nonetheless, the evidence suggests that understanding the strengths and weaknesses of the red lines in play is useful in forecasting the chain of events through which a crisis is likely to unfold, including how that crisis might escalate to war.

Implications

This book has sought to develop a theoretical framework to better understand how states win crises and other coercive conflicts. This framework has implications for predicting future crises, for managing those crises should they arise, and for avoiding these crises altogether by proactively eliminating vulnerabilities in red lines.

First, applying the concept of vulnerable red lines to current and potential hotspots facilitates forecasting where and how future crises might unfold across the globe. In Northeast Asia, for instance, the prospect of a future crisis between Japan and China over the Senkaku Islands escalating to a land grab is, unfortunately, all too plausible. The Falklands War notwithstanding, however, few of these land grabs over islands escalate to war. With respect to a future crisis in the Taiwan Strait, similar concerns must surround the Taiwanese-held islands of Kinmen and Mazu off the coast of mainland China. As the 2014 Ukraine Crisis suggests, what starts as a crisis over one issue (e.g., domestic political turmoil) can become a territorial crisis when one side sees an opportunity to exploit a gray area to make a territorial gain. On the Korean Peninsula, a North Korean fait accompli attempt across the Demilitarized Zone is
relatively unlikely. Were North Korea to make such an attempt, the results suggest that war would be the likely outcome. If a land grab takes place in a future crisis in Korea, it is most likely to target one of the small islands held by South Korea near the disputed Northern Limit Line. These examples from one region are a microcosm of how identifying gray areas can facilitate planning for future crises around the world.

In Western Europe, the 2014 Russian fait accompli in Crimea was foreseeable even though Russia had not pressed its claims until the opportunity arose. The underlying geography – Crimea’s position as distinctly separate from the rest of Ukraine – meant that the opening for a fait accompli would remain available to Russia. Although the theory could not predict when a crisis would arise between Russia and Ukraine, it pointed clearly to Crimea as an area for long-term concern from the moment that Ukraine became independent. Elsewhere in Europe, however, this book provides some basis for optimism. With the notable exception of the South Ossetia and Abkhazia in Georgia, there remain few significant islands or border ambiguities between Russia and its neighbors. Despite tensions with the Baltic states, a similar crisis is less likely to unfold because no tempting gray area exists.

Second, this book provides guidance for managing crises, including where to press for gains and where to exercise restraint. Although the four types of weaknesses in red lines yield many diverse openings for unilateral gains, the most consistent opportunity is to find a way to take something without having to overtly attack the other side’s forces. Suppose, for instance, that China takes advantage of absence of Japanese forces in the Senkaku Islands to land forces there, seizing at least the main island in a fait accompli. Japan would then find itself in the unenviable position of deciding between escalating to retake the islands and suffering the humiliation of tacitly relenting to the Chinese gains. History suggests that threats alone are unlikely to dislodge Chinese forces in a scenario along these lines, so these would seem to be the two main options. However, when confronted with Greece’s occupation of the disputed island of Imia in 1996, Turkey found a third option that worked around the Greek red line against attacking its marines. Rather than attack or back down, the Turkish military seized a nearby unoccupied – but also disputed – island. From there both stared each other down until they mutually agreed to withdraw. Neither side lost face. Escalation was avoided. The same option, an option which maneuvers around the firing-on-forces focal point – would be available to Japan if China declined to occupy some of the smaller islands in the group. This is the sort of strategy that this book commends. More generally, identifying and exploiting these unilateral ways to flank the adversary’s red line and/or exploit a gray area in it can be the key to victory in a crisis.

To augment deterrence against this tactic, it is advantageous for a state to “stack” focal points so that they overlap to coincide with each other as the basis for the strongest possible deterrent red line. In the short run, this means avoiding leaving sensitive areas unoccupied absent a policy choice to be willing to sacrifice them. Tripwires have real deterrent value, especially when set on top of geographic focal points, something which is not always possible. The strategic game surrounding these tripwires is more intricate and more consequential for crisis outcomes than has generally been appreciated.
Finally, avoiding the creation of unnecessary gray areas such as West Berlin should make it possible to prevent future crises and *faits accomplis*, strengthening deterrence. Islands may be immutable, but border ambiguities and enclaves are not. Taking advantage of moments when diplomatic relations are unusually harmonious to negotiate clear settlements to existing border ambiguities can help to prevent future crises. Too many crises grew out of ambiguities in boundaries or areas that seemed unimportant at the time (e.g., in the colonial era) but became more significant later. The same point can extend to any of the four types of red line vulnerabilities. Although its practicality is open to question, in principle a more intensive surveillance effort directed at Cuba in 1962 could have reduced the unverifiability of the red line and, consequently, bolstered deterrence.

It is all too often impossible to rectify weaknesses in red lines that leave openings for *faits accomplis* once a crisis begins. However, an appreciation of these weaknesses years or decades in advance can enable their removal when doing so is still possible. Proactively eliminating weaknesses in what could become important red lines should prevent a significant number of crises from ever beginning. The Siachen Glacier region in 1974 and the exact nature of the division of Germany during World War II both seemed like secondary issues to policymakers at the time. Both became gray areas as a result. Both were necessary causes of future crises that could have been avoided.

The International Relations field has too long neglected the roles of both weaknesses in red lines and the *faits accomplis* which exploit them. Through these means, it is quite possible to make gains and win crises without signaling resolve, changing adversary perceptions, or extracting concessions. Red lines and *faits accomplis* are vital tools of statecraft, and both are essential to understanding how states win crises.