In coercive diplomacy, coercers have two basic objectives. First, they want to wrest the largest possible concessions from the target. Second, they want to win without having to pay the costs of war. Yet these two objectives are often in tension: demanding larger concessions carries a greater risk of war. Avoiding war, by contrast, may require backing down in defeat. How do coercers balance these competing pressures when deciding “how much” to demand? This chapter develops a simple model of crisis bargaining to evaluate the strategic choices faced by coercers. How do they weigh their desire for gains against the risk of war? How do power and information play into these calculations? The analysis generates two counterintuitive hypotheses that challenge conventional wisdom about coercion. First, greater military power emboldens coercers to make riskier demands, increasing the likelihood of failure. Second, the model demonstrates that coercers are motivated to attenuate their demands when the target’s resolve is high, thereby making success more likely. In other words, we are more likely to observe coercive success when the target is believed to be highly-resolved. Both insights underscore the importance of the magnitude of demands in the analysis of coercive bargaining.
Introduction

In the autumn of 1939, the Soviet Union engaged in an ambitious coercive campaign against Finland. For over a year Josef Stalin had watched Adolf Hitler apply his expansionist agenda in central Europe, and began preparing for the day the German army turned against the Soviet Union. High on Stalin’s agenda was the security of Leningrad, which would be vulnerable to German attack in the event that Finland was unable – or unwilling – to stop a Nazi invasion. Stalin’s solution was to obtain a series of territorial concessions from Finland, which would allow the Soviets to expand the defense perimeter around Leningrad and install island fortifications in the Gulf of Finland to prevent a naval assault. The Soviets initially attempted to achieve this objective with a strategy of “friendly persuasion,” but soon escalated to threats of war.¹

After months of unsuccessful negotiations, however, the Soviets concluded that coercive diplomacy had failed. The Finnish government refused to yield to Stalin’s demands, and on November 30, 1939, the Soviet Union launched a multi-pronged military assault against Finland. After more than three months of winter combat, Finland sued for peace, ultimately ceding even more territory than the Soviets had originally demanded. The Soviets paid a heavy price for these gains: the war cost the lives of perhaps 127,000 Soviet soldiers,² the Finnish army destroyed thousands of Soviet aircraft and tanks, and the surprising ferocity of Finland’s resistance dealt a severe blow to the prestige and morale of the Red Army.

The Soviet Union’s strategy during the 1939 crisis offers a useful glimpse into the challenges inherent in coercive diplomacy. Stalin found himself with two competing objectives during the negotiations. On one hand, he sought to wring as many concessions as possi-

¹The most complete account is Jakobson 1961. See also Upton 1974; Tanner 1957.
²Krivosheev 1997.
ble from Finnish negotiators. At various points during the crisis, Stalin demanded: the establishment of a Soviet naval base on the strategic peninsula of Hankö; a revision of the Soviet-Finnish border in the Karelian isthmus; anchorage rights in critical Finnish ports; and ownership of a variety of islands in the Gulf of Finland, among other territories. The more territory and bases he could secure, Stalin reasoned, the better he could protect Leningrad against a Nazi assault.

On the other hand, however, Stalin also sought to avoid war with Finland. While he believed the Soviets could easily defeat the Finnish army (an assessment that later proved far too optimistic), he also sought to avoid distractions as the Soviets prepared for a war in Europe. Further, fighting the Finns risked provoking intervention by the United States, Britain, or France – a potentially disastrous outcome. Yet the more Stalin demanded, the more motivated Finland would be to stand firm. Throughout the negotiations Stalin had several opportunities to accept Finnish counteroffers, and at several points even moderated his demands. Ultimately, however, he decided to escalate the crisis rather than accept a suboptimal deal.

The Russo-Finnish crisis of 1939 therefore highlights three important realities about coercion in international relations. First, coercion is fundamentally about bargaining. Achieving one’s coercive objectives entails a process of demands, offers, counteroffers, concessions and trades. In this respect, coercive diplomacy is no different from the sort of negotiations one might observe in a used-car lot: bargaining takes place along a continuum, with each side trying to obtain the best possible deal. The negotiations between Finland and the Soviet Union in 1939 spanned several months, with each side staking out positions that they hoped would maximize their own gains while not pushing the adversary past its breaking point.

What makes coercive diplomacy unique is not the subject matter of the bargaining, but

---

3 Jakobson 1961, pp. 115–16.
4 A variety of studies have evaluated Finland’s decision-making during the crisis. See, for example, Vital 1971; Cohen 1989; Yuen 2009; Sechser 2010.
5 Upton 1974, p. 40; see also Anderson 1954.
6 The literature about coercive diplomacy is extensive. Theoretical and conceptual treatments include George and Simons 1994; Freedman 1998; Byman and Waxman 2001; Art and Cronin 2003. Empirical and historical studies of coercive threats include, for example, Blechman and Kaplan 1978; Petersen 1986; Jakobsen 1998; Schultz 2001; Greenhill 2010; Sechser 2011.
rather the fact that it takes place against the background of the threat of military force.

Second, coercers operate in an environment of imperfect information. Like poker players in a high-stakes game, coercers can see the hand they have been dealt, but others’ hands are hidden to them. Choosing a coercive demand would be straightforward if coercers knew their adversary’s level of resolve: the coercer could then demand precisely enough to maximize its gains while still avoiding war. Leaders, however, cannot read minds. Worse, leaders have reasons to exaggerate their resolve in order to encourage their opponents to make more favorable offers. Stalin’s negotiators in 1939 spent several months trying to establish Finland’s precise level of resolve, at one point pointing to a map and repeatedly asking “would you perhaps give up this island?” Ascertaining the opponent’s level of resolve is a basic feature of coercive diplomacy.

Third, coercers must choose their demands wisely. In coercive diplomacy, coercers have two basic objectives. On one hand, they want to wrest the largest possible concessions from the target. On the other hand, they prefer to win those concessions without having to pay the costs of war, if possible. Yet these two objectives are often in tension: demanding larger concessions carries a greater risk of war. Large demands increase the potential rewards for success, but decrease the likelihood that success will be achieved without war. Coercers therefore must strike a careful balance between maximizing gains and minimizing risk. The Soviet position in 1939 nicely illustrates this fundamental tension. Emboldened by his belief that the Finns would not dare stand up to Soviet power, Stalin budged little from his original demands. Yet had Stalin foreseen the costly war that would follow the 1939 crisis, he might have revised his position. Likewise, the Finnish leadership engaged in extensive debate about just how much risk they could bear – and how much territory it would be worth giving up to mitigate that risk. In the end, each side overplayed its hand.

This chapter considers the implications of the bargaining perspective for coercive diplomacy. In particular, it applies the logic of bargaining to better understand the choices faced by coercers. How do coercers weigh their desire for gains against the risk of war? How do

---

7Cohen 1989, p. 255. Finnish negotiators, realizing immediately what the Soviets were up to, refused to play along.
power and information play into these calculations? And most importantly, how does this perspective reshape our understanding of the dynamics of coercive diplomacy?

To explore the choices and dilemmas faced by coercers, I develop a simple bargaining model of international conflict. One important advantage of utilizing a formal model is that it demands that we consider the perspective of all sides in a coercive bargaining encounter. To completely specify a mathematical model, every player’s strategies, outcomes, and preferences must be stated. This requires imagining a crisis from each player’s viewpoint, envisioning not only choices made but also options that are discarded. The analysis yields several surprising implications. First, the model tells us something about the conditions under which we are likely to observe successful coercive diplomacy in international relations. Most theories of coercive diplomacy tell us that coercive threats are most likely to fail when targets are highly-resolved. But the model turns the conventional wisdom on its head: by accounting for coercers’ incentives to adjust their demands to avoid war, the model shows that successful coercive diplomacy is actually more likely when targets are resolved. Second, the model underscores the double-edged nature of military power in coercive diplomacy. While military power is useful for intimidating adversaries into backing down without a fight, it also motivates coercers to make larger and riskier demands. Powerful coercers therefore may not be more likely to win without a fight; indeed, they may even be less likely to do so.

In the sections that follow, I first discuss the bargaining approach to understanding war and peace, and explain how it can be a useful tool for understanding coercive diplomacy. Second, I develop a simple bargaining model of coercion and introduce two principal modifications to that model. Third, I draw out some empirical implications of that model, and provide some suggestive historical data in support of those implications. The final section links the findings of this chapter to the central themes of this book, offering some thoughts about how a bargaining perspective can help us better understand emerging challenges in coercive diplomacy.
Coercion as a Bargaining Process

In a civil lawsuit, there is typically a period of negotiation in which the parties attempt to settle the lawsuit before it ever reaches a courtroom. The reason is that going to trial is costly: even if one is fortunate enough to win, litigation costs time and money, mitigating the gains from victory at trial. Furthermore, the outcome of a trial is uncertain beforehand. The parties therefore have significant incentives to settle the dispute before expensive attorneys’ fees begin to accumulate. Behind these negotiations, however, is an implicit threat from the plaintiff. If the settlement offered by the defendant is not attractive enough, the plaintiff might decide to take his chances in court, an outcome that could be costly for everyone. The plaintiff, in effect, hopes to use the threat of a courtroom trial to coerce the defendant into offering a lucrative deal.

This picture of pretrial negotiations bears an important resemblance to international conflicts: both typically center around allocation of a scarce resource. In a civil context, that resource generally takes the form of monetary damages. The plaintiff hopes to receive as much money as possible, whereas the defendant prefers to pay nothing at all. In an international setting, disputed issues can include territory, trade policies, treatment of an ethnic group, a country’s leadership, or a variety of other issues. In each case, however, the disputing parties have opposing preferences: each wants to “win” as much of the issue as possible. The central problem therefore is how to allocate that resource.

Characterizing these settings as problems of resource allocation highlights a second important similarity between pretrial negotiations and international disputes: coercion plays a key role in both. In both settings, the disputants hope to use the threat of a costly “outside option” for resolving the dispute – a courtroom trial, or a war – to compel the adversary to agree to a deal beforehand. The ability to successfully compel a settlement stems from one’s ability to shape that costly outcome. For instance, a litigant that appears very likely

---

8The puzzle of why civil legal disputes ever end in trial rather than settlement has been a persistent question in the legal studies field for decades. See, for instance, Mnookin and Kornhauser 1979; Cooter et al. 1982; Priest and Klein 1984; Bebchuk 1984; Gross and Syverud 1991; Loewenstein et al. 1993; Korobkin and Guthrie 1994.
to prevail in a jury trial, or whose legal costs are being covered by a wealthy third party, stands to negotiate a more favorable settlement because he does not fear the costs of a trial. Likewise, a state that is sure to prevail in war, or whose population does not particularly fear the costs of fighting, should also be able to drive a better bargain.

In short, a bargaining perspective sees both threats and military force as central elements of coercive diplomacy. Violence and negotiation are not distinct entities; they are both part of the same bargaining process. As Thomas Schelling wrote nearly 50 years ago: “Coercion requires finding a bargain, arranging for [the adversary] to be better off doing what we want – worse off not doing what we want – when he takes the threatened penalty into account.”

A Basic Model of Coercion

In this section I elaborate the assumptions of the bargaining perspective and explore the dynamics of coercion using a simple model of bargaining. The model draws its inspiration from the famous “ultimatum game,” a bargaining model often used by economists to represent situations in which one actor attempts to impose its will on another. The format of the basic ultimatum game is elegantly simple: one actor proposes a division of a valuable item (usually a pot of money), and the second actor may accept or reject the proposal. If the second actor rejects the offer, however, both parties get nothing.

The ultimatum game is useful for thinking about coercive diplomacy for several reasons. First, it describes a situation in which one actor makes a clear demand and the other must respond. Coercive encounters in international relations often fit this mold, with one side making a demand and the other deciding between accepting or rejecting it. President John F. Kennedy’s demand that the Soviet Union remove its missiles from Cuba in 1962, for example, left little room for a counteroffer and carried an assurance that the only two options available to the Soviets were total compliance or war. Similarly, George W. Bush’s message

---

9Schelling 1966, p. 4.
11For examples of ultimatums in international crises, see Lauren 1972.
to Saddam Hussein prior to the 2003 invasion of Iraq was a clear ultimatum: “Saddam Hussein and his sons must leave Iraq within 48 hours. Their refusal to do so will result in military conflict, commenced at a time of our choosing.”\textsuperscript{12} The ultimatum game allows us to analyze the bargaining dynamics of these encounters.\textsuperscript{13}

Second, the proposer in the ultimatum game must think strategically about its demand. If it demands too much, it risks the possibility that the recipient will reject the demand and leave it with nothing. Indeed, behavioral economists have found that individuals are quite willing to reject ultimatum offers they believe to be unfairly low, even though this means a lower payoff for themselves.\textsuperscript{14} The proposer therefore must weigh the temptation of larger gains against the risk that if it overshoots, it may get nothing at all. At the same time, however, it could be costly to be too cautious. If the proposer asks for too little, its demand will surely be accepted, but the gains from the deal will be small.

A third useful aspect of the ultimatum game is that there is a price to be paid for failing to reach a bargain. In coercive diplomacy, failing to reach a settlement might lead to a costly military conflict. Both sides therefore have an incentive to strike a deal. Similarly, in the ultimatum game, if the recipient rejects the proposer’s offer, the two sides receive nothing, leaving the pot of money on the table.

While the ultimatum format is a useful representation of coercive encounters, it is important to be clear about its limitations. Perhaps the most important one is that it does not allow the recipient to make a counteroffer: it can only accept or reject the original proposal without modifying it. In reality, of course, international negotiations are much more free-wheeling. Leaders in international disputes are not bound by strict rules that stipulate who makes offers, what those offers may look like, and how many rounds of negotiation may take place before the game ends. Further, these rules are not a trivial matter: bargaining models

\textsuperscript{12}Bush 2003.
\textsuperscript{13}It may seem odd to classify the ultimatum game as a bargaining model, since it does not actually contain any back-and-forth haggling. I use the term in its broadest sense, referring to the fact that the model requires the players to decide how to divide a valuable, scarce resource, even though the rules of negotiation are quite restrictive.
\textsuperscript{14}For example, see the reviews in Güth and Tietz 1990 and Camerer 2003, chapter 2.
are notoriously sensitive to minor changes to the rules.\textsuperscript{15} The advantage of the ultimatum model lies in its simplicity and ease of analysis, but it is important to acknowledge that it offers a first glimpse at the problem of coercion, not the final one.\textsuperscript{16}

Before applying the ultimatum model to coercive international bargaining, we need to make a few modifications. First, the model will retain the two-player format of the original ultimatum game, but we will henceforth refer to the players as the “coercer” and the “target.” Next, let us stipulate that the item in dispute is already owned by the second actor. In international relations, there are few equivalents to an unclaimed $100 bill in the middle of the street – most objects of value already have an owner. In this respect, then, the game resembles Schelling’s characterization of a “compellent” threat: one actor demands something from an adversary and threatens some kind of punishment if the adversary does not act.\textsuperscript{17}

Finally, we adjust the outcome of the game. In the standard version of the ultimatum game, the parties receive nothing if they fail to reach an agreement. However, we modify this outcome such that the parties fight a winner-take-all war over the disputed item if the proposer’s offer is rejected. In this war, each side enjoys some probability of winning the entire item, but also pays a fixed cost for fighting whether it wins or loses – though not necessarily the same cost as its opponent.\textsuperscript{18} In other words, a rejection of the coercer’s threat triggers a “costly lottery” in which both sides pay a price for a chance at winning the entire item. This sequence of events is illustrated graphically in Figure 1.\textsuperscript{19}

\textsuperscript{15}See, for example, Fudenberg and Tirole 1991, pp. 398-399, and Banks 1990.

\textsuperscript{16}It is important not to exaggerate this point, however. Even though the ultimatum game employs restrictive rules, changing them would not necessarily alter the findings of the model. Rubinstein, for example, analyzed a model with a more fluid back-and-forth bargaining protocol, but the first-round equilibrium offer turns out to be exactly the same as the equilibrium offer in an ultimatum game. Further, in equilibrium Rubinstein’s model lasts only one round, just as in the ultimatum model. See Rubinstein 1982 and Powell 1996, pp. 252–55.

\textsuperscript{17}Schelling 1960, pp. 195–99.

\textsuperscript{18}The basic model described here was first used in an international relations context by Fearon. See Fearon 1992; Fearon 1995. A number of scholars following in this research program have adopted some form of this model, although others have abandoned it in an effort to gain leverage on problems of wartime and postwar bargaining.

\textsuperscript{19}Reducing war to a single-shot lottery seems unnatural, particularly in light of Clausewitz’s admonition that “war does not consist of a single short blow.” Yet while the model neglects to delve into wartime negotiations, the lottery represents each side’s expectations about the likely payoffs from fighting. Since our focus here is on the pre-war process of coercive bargaining, this simplification is a worthwhile tradeoff. See Clausewitz 1832 [1976], p. 79. Studies that explore the nature of wartime bargaining include Wittman 1979;
In this interpretation of the ultimatum game, “success” occurs when the target accepts the coercer’s demand, and “failure” occurs when the target rejects it. These terms refer solely to the outcome of the threat itself – not whether the overall encounter is a success or failure for the coercer. In coercive diplomacy, an unsuccessful threat does not necessarily imply a foreign policy failure. Indeed, failed threats often precede successful military operations. In 1990 and 1991, for example, the United States issued a coercive threat against Iraq, with President George H.W. Bush demanding that Saddam Hussein order his military forces out of Kuwait. The U.S. threat was unsuccessful, but the ultimate outcome of the crisis was not: U.S. forces expelled Iraqi troops from Kuwait after just six weeks of combat. At the same time, the United States surely would have have preferred to achieve this outcome without any fighting at all. This is precisely the aim of a coercive threat. The outcome of interest in this discussion therefore is the success or failure of coercive threats, rather than broader crisis outcomes.\footnote{Filson and Werner 2002; Slantchev 2003b.}
Some additional assumptions are in order. The first and most basic of these is an assumption about the actors in a coercive encounter: we will assume that the players in our model are rational and that they seek to maximize their gains (or minimize their losses) from the encounter. Stipulating that the actors are rational does not imply that they are all-knowing, but it does mean that they are able to make optimal decisions given the information available to them. Second, we will assume that military conflict only takes one form: “war.” In other words, the coercer does not choose how much military force to employ. Third, we stipulate that the players have full and accurate information about one another’s payoffs. While this assumption will be relaxed shortly, it is useful to begin with this assumption to establish a baseline for our analysis. Finally, let us assume that the players in the model are very slightly risk-averse, such that when peace and war offer identical expected payoffs, they prefer peace.

The model immediately reveals a central feature of coercive diplomacy: the objective of a coercive demand is not to provoke war, but rather to prevent it. Coercion relies on the threat of military punishment to achieve political objectives, but enacting that punishment is costly even if it achieves its ultimate coercive objective. Wars entail substantial human, material, and psychological costs that leaders would prefer to avoid if possible. As Schelling wrote in his seminal book on coercion, *Arms and Influence*, “successful threats are those that do not have to be carried out.”

For a coercer, then, the basic challenge is to devise a threat that attains the greatest gain at the least possible cost.

**Analysis of the Basic Model**

We now turn to a more technical analysis of the model to illuminate its implications for our understanding of coercive diplomacy. We start by assigning mathematical terms to represent its various components. The first step in the model involves the coercer’s decision about how much to demand. For simplicity we will stipulate that the coercer and target are negotiating

---


22 Clausewitz recognized this as well when he wrote: “The aggressor is always peace-loving...he would prefer to take over our country unopposed.” Clausewitz 1832 [1976], p. 370.
over an item with a value of 1. The coercer must choose how much – if any – of that item
to demand. We will use the term $x$ as a placeholder for the coercer’s demand, noting that $x$
can range anywhere from 0 to 1. Using this terminology, we can now define the outcome of
a successful coercive threat: if the target accepts the coercer’s demand, the coercer receives
$x$ and the target keeps $1 - x$.

The other possible outcome of the game is a military conflict. If the parties fight a war
over the disputed item, each stands a chance of winning the entire item. Let us stipulate that
the coercer’s probability of winning is represented by $p$. The target’s probability of winning
is therefore $1 - p$. Since the balance of power plays a central role in determining each side’s
odds of prevailing in a military contest, it is plausible to say that $p$ represents the coercer’s
relative share of military capabilities: the more military power it possesses, the higher $p$ will
be.\textsuperscript{23} Further, each player pays a cost for participating in this military “lottery.” We will use
the notation $c_C$ to represent the coercer’s cost, and $c_T$ to represent the target’s cost. Let us
stipulate that the costs of war are also loosely related to the balance of power, with more
powerful actors able to inflict greater costs on their opponent.\textsuperscript{24} Recall that the players must
pay these costs whether they win or lose. We can now represent each player’s “expected”
payoff from fighting a war: for the coercer, that payoff is $1(p - c_C) + 0(1 - p - c_C)$, which
simplifies to $p - c_C$. For the target, the expected payoff of war is $1(1 - p - c_T) + 0(p - c_T)$,
which simplifies to $1 - p - c_T$.

What is the result of this coercive bargaining encounter? We can answer this question
by solving for the Nash equilibrium of the game, which is described in the appendix for this
chapter. The coercer makes the largest demand that it knows will be accepted, which turns
out to be $x = p + c_T$. The target in turn accepts this demand, and war is avoided.

The first thing to note is that, in equilibrium, this encounter always ends in successful
coercion. In other words, war is never an equilibrium outcome. The reason is that the coercer
has the luxury of complete information: it knows exactly what its opponent’s payoffs are, so

\textsuperscript{23}More formally, stipulate that $p$ is defined as the coercer’s share of military capabilities ($m_i$) in the dyad.
Therefore: $p = \frac{m_C}{m_C + m_T}$.

\textsuperscript{24}To be precise about it, let $c_C$ and $c_T$ be determined by some monotonic function $f_i(p, m_i)$ such that $p$
and $c_T$ are positively related, $p$ and $c_C$ are negatively related, and $m_i$ and $c_i$ are negatively related.
it chooses its demand strategically so that it maximizes the coercer’s gains without provoking a costly war. Because the target expects to receive $1 - p - c_T$ from fighting, and the coercer knows this, the coercer can adjust its demand so that the target is left with exactly this amount (or very slightly more) by acquiescing to the demand. Since the coercer will receive $p + c_T$ from this demand, but only $p - c_C$ from a war, it has no incentive to provoke a conflict by demanding more. The coercer knows how far it can push, and pushes no further. This model therefore underscores the importance of accounting for strategic behavior on the part of the coercer. To understand coercive outcomes, we must understand not only the target’s incentives, but also the coercer’s strategy.

A second observation relates to the role of military power in this coercive encounter. The balance of power shapes the outcome of the crisis, but not in the way that is often assumed in coercive diplomacy scholarship. The standard view is that the balance of power influences the likelihood of coercive success: coercers that can threaten greater punishment are more likely to have their demands accepted. In this view, power enters the scene in the final stage of a crisis, when the target chooses between capitulation and war. But this is not how power works in our ultimatum model. Instead, the primary role of power is to shape the coercer’s demand: coercers with a military advantage – that is, a higher $p$ – make larger demands, whereas weaker coercers settle for small demands. In other words, the coercer incorporates the balance of power into its initial demand, taking into account how its demand will be received by the target. Ultimately, all equilibrium demands have the same likelihood of success, irrespective of the balance of power.

Yet this model is unsatisfying. It is implausible to assume that the coercer knows the target’s level of resolve with such precision. Indeed, this assumption begs the question, since a central question in coercive diplomacy is how coercers and targets can communicate their resolve through signals and other diplomatic measures. The following section relaxes this assumption.

26See also Fearon 1992, p. 24.
27Fearon 1997; Fearon 1994; Morrow 1999b.
The Role of Information in Coercive Diplomacy

One of the most important insights to emerge from international relations scholarship about war is that wars are often a consequence of *incomplete information*. Adversaries in a dispute, in other words, must make educated guesses about the other side’s resolve, military capabilities, and other factors. When leaders miscalculate about these factors, war can be the result.\(^{28}\) Understanding the outcomes of coercive encounters therefore requires that we first understand the information environment in which the actors are operating. In short, we must answer the question: how much does each side know about its opponent?

In our original ultimatum model of coercive bargaining, each player knew all the relevant facts about its opponent: its military capabilities, its costs for fighting, and its value for the issue at hand. But what happens when we relax this assumption and allow for the possibility of error? We can investigate this by introducing a small amount of uncertainty into the game. Let us stipulate that the coercer is facing a target whose resolve it does not know. The target can be one of two types: a “tough” type and a “weak” type. A “tough” target is distinguished by having lower costs for fighting wars, so it is generally more willing to risk war by resisting a coercive threat. But, crucially, the coercer does not know whether it is facing a tough or weak target.\(^{29}\) The coercer is not completely in the dark, however: there is some probability that it is facing each type of target, and it is aware of this probability. But it must devise its strategy without being certain about the other side’s resolve.

We will define the two types of targets as follows. Let us stipulate that a “tough” target is defined as one whose cost for war is \(\xi_T\). A “weak” target’s costs for war are \(\bar{\xi}_T\), where \(\xi_T < \bar{\xi}_T\). Further, both cost terms are tethered to the balance of power, \(p\), such that the ratio of the two cost terms, \(\frac{\xi_T}{\bar{\xi}_T}\), remains constant across all values of \(p\). Lastly, we need to define the probability that the target is “tough.” Call that probability \(\tau\), which implies that \(1 - \tau\) is the probability that the target is weak. At the outset of the game, the target’s type is determined with these probabilities but the outcome is not revealed to the coercer.

\(^{28}\)For example, see Blainey 1973; Morrow 1989; Fearon 1995.
\(^{29}\)The notion of using “types” with different utility functions to represent uncertainty originates from Harsanyi 1967.
How does the outcome change when uncertainty is introduced into our model of coercive diplomacy? Working backward provides the answer. As before, the coercer seeks to demand as much as possible without pushing the target to resist. Yet while the coercer can in principle demand anything between 0 and 1, we can simplify things by recognizing that there are actually only two demands that a coercer might want to make: \( x = p + cT \) and \( x = p + \bar{c}_T \). The logic here is straightforward. From the coercer’s point of view, its demand can have one of just three discrete probabilities of succeeding. First, its demand will never succeed if it demands so much that both tough and weak targets would rather fight than capitulate (i.e., \( x > p + \bar{c}_T \)). This is never optimal since the coercer would prefer to have at least a chance of avoiding the costs of fighting. At the other end of the spectrum, the coercer can be assured of a successful demand if its demands so little that both tough and weak targets will acquiesce (\( x \leq p + cT \)). The coercer’s third option is to demand somewhere in between, where the probability of success is equal to the probability that the target is weak (\( 1 - \tau \)). For a demand in this range (\( p + cT < x \leq p + \bar{c}_T \)), a tough target will refuse while a weak target will accept. The coercer therefore must first decide which risk level it prefers – low, medium, or high – and then select the maximum demand within that range.

How much will the coercer demand? The answer to this question depends on the likelihood that the target is tough. If that likelihood is sufficiently high, then the coercer will make the maximum possible demand within the “low” range, which turns out to be \( \bar{x} = p + c_T \). But if the coercer is confident enough that the target is weak – in other words, if \( \tau \) is sufficiently low – then it will be emboldened to make a demand in the “middle” range: specifically, \( \bar{x} = p + \bar{c}_T \). However, it is never optimal for the coercer to make a demand beyond this level, since it will assuredly trigger a war. Thus, for practical purposes the coercer is left with only two possible demands: for simplicity we will call them “high” (\( \bar{x} \)) and “low” (\( \bar{x} \)). This is illustrated in Figure 2, which tracks the likelihood of successful coercion as the coercer’s demand increases.

Now that we have narrowed down the coercer’s choices to just two possible demands, \( \bar{x} \) to refer to the higher demand and \( \bar{x} \) to refer to the lower one.
Figure 2. The probability of successful coercion as the coercer’s demand varies.

we can determine its strategy. In equilibrium the coercer will issue the high demand (i.e., \( \bar{x} = p + \bar{c}_T \)) if the probability of facing a tough target falls below a critical value, which we will call \( \tau^* \).\(^{31}\) Otherwise, it will make the low demand, \( \bar{x} = p + \underline{c}_T \). In other words, under conditions of uncertainty, the coercer only makes a high demand if it is sufficiently confident – although not necessarily certain – that the target is a weak type and will accept the demand.

**The Effects of Uncertainty on Coercive Outcomes**

The introduction of uncertainty into the model has an immediate and important effect: coercive failures now become possible. The coercer no longer has the luxury of peering into the adversary’s mind to determine exactly how much to demand, so there is real possibility that it will guess incorrectly and issue a demand that is rejected. To be precise, the likelihood of coercive failure is \( \tau \) so long as the coercer is motivated to issue the higher of its two possible

\[^{31}\text{Specifically, this critical value is } \tau^* = \frac{\bar{c}_T - \underline{c}_T}{\bar{c}_C + \bar{c}_T}.\]
equilibrium demands. Under conditions of incomplete information, then, there will always be a chance that coercion will fail. This is not, however, the result of a strategic error on the part of the coercer: indeed, the coercer is playing its optimal strategy, given the information it has. But since its information is necessarily incomplete, it may issue a demand that turns out to exceed the target’s threshold of tolerance.

Coercion and the Shadow of the Future

We now add one final wrinkle into our model of coercive diplomacy. In the models we have investigated so far, there is only one coercive episode. In reality, however, there is no artificial limit on the number of coercive encounters two opponents might experience. The United States, for example, issued coercive demands on multiple occasions against Iraq between 1990 and 2003. Japan engaged in a long string of coercive attempts against China during the 1930s; Germany likewise made multiple coercive demands against Czechoslovakia, Austria, and Poland prior to World War II. In short, coercion is not necessarily a one-time affair. We therefore build on the previous model by incorporating a second coercive episode: once the outcome of the initial confrontation has been decided, the coercer may have the opportunity to make a new demand on another issue. Since the future cannot be known with certainty, however, we stipulate that this second opportunity arises with some known probability. The players do not know beforehand whether a second crisis will occur, but they do know how likely this event is. We also retain the incomplete-information component of the previous game, stipulating that the coercer does not know whether it is facing a tough or weak target.

To account for the shadow of this potential future crisis, we introduce a “discount factor” into the model. Let $\delta$ represent the value of the item in the second dispute, adjusted to account for the uncertainty that a second crisis will even take place. This modification allows each player to calculate its strategy in light of its expectations about the second-round.\footnote{We also need to modify slightly the notation for the coercer’s demands to accommodate this additional round: let $x_n$ represent the coercer’s demand in round $n$.}
Table 1. *Summary of formal notation.*

Table 1 reviews the notation that we have introduced in our models thus far.

The discount factor, represented by $\delta$, deserves a brief discussion here. Because $\delta$ represents the net present value of the second-round pie (i.e., its value to the players at the outset of the game), it is best conceived as the product of three factors. The first factor is the value of the object, which we will stipulate is equal to 1, just as in the first round. The second factor is a discount placed on future gains, reflecting each player’s preference for obtaining gains sooner rather than later. The third factor is the probability that a second opportunity for a dispute will even arise. This uncertainty exerts a discounting influence on the present value of the second-round item because if the likelihood of a second round is very small, then the expected gains from that round must also be proportionately small.

---

33 This assumption does not undermine the generality of the analysis, but does simplify the equilibrium calculations.

34 Note that if $\delta$ is equal to zero – that is, the players are certain that no second round will take place – the game becomes identical to the one-round, incomplete-information model evaluated above.
Coercive Outcomes in the Multi-Round Model

The introduction of a possible second coercive encounter changes the equilibrium of the game in a number of respects. We begin by analyzing the coercer’s strategy. As in the one-round version of the game, if the probability of encountering a tough target is sufficiently low, the coercer will be willing to make a “high” demand. That demand turns out to be $p + \bar{c}_T - \delta(\bar{c}_T - c_T)$. If the likelihood of facing a tough target is too high, however, the coercer will make a smaller first-round demand $(p + c_T)$. This demand is calibrated to persuade all weak targets to capitulate and all tough targets to resist in the first round. The outcome of the first-round therefore helps the coercer identify tough opponents so that it can accurately calibrate its demand in the second round. In game-theoretic terms, this is called a separating equilibrium: the coercer issues a demand that “separates” the two types of targets. In other words, the types identify themselves by their response to the coercer’s first-round demand. This is optimal for the coercer because it can then exploit its newfound knowledge to maximize its gains in the second round. In short, the risk of coercive failure in the first round is a necessary price it must pay in order to obtain information that could be useful in the second encounter.

How do weak targets respond to the coercer’s strategy? Compared to the one-round version of this model, a weak target must adopt a different approach when there is the possibility that a second coercive encounter may occur. Previously, a weak target would have been willing to acquiesce to any demand less than $p + \bar{c}_T$, but now it will reject a larger set of demands – specifically, it will reject any demand greater than $p + \bar{c}_T - \delta(\bar{c}_T - c_T)$. In other words, there is a range of demands that was once agreeable to the weak target, but becomes unacceptable when the weak target develops an expectation that the game might be repeated. This is illustrated in Figure 3, which compares the first-round equilibrium strategies of both types of targets in the one- and two-round games. As the figure demonstrates, weak targets become more difficult to coerce in the two-round game, whereas tough targets maintain the same strategy in both games. To persuade a weak target to acquiesce in the two-round game, a coercer must reduce its demands.
What is going on here? To understand why the weak target has become less agreeable, consider a weak target’s fortunes if it instead played the same strategy as it played in the one-round game. Given a sufficiently low $\tau$, the coercer would begin by demanding $x_1 \leq p + \tau_T$, to which the weak target would agree. The coercer, realizing its target was a weak type, would make the same demand in the second round, giving the target an overall payoff of $(1 - p - \tau_T) + \delta (1 - p - \tau_T)$. In contrast, resisting this demand in the first round causes the coercer to make a lower second-round demand ($x_2 = p + \tau_T$), which the weak target will accept, for a total payoff of $(1 - p - \tau_T) + \delta (1 - p - \tau_T)$. This outcome is strictly better than the alternative of capitulating in the first round.\(^{35}\)

To put it differently, weak targets are motivated to resist otherwise acceptable demands because doing so allows them to masquerade as tough targets. By resisting demands they would really prefer to accept, weak targets can secure lower demands in the second round.

\(^{35}\)Tough targets, by contrast, do not change their equilibrium strategy in this game, and reject all demands greater than $p + \tau_T$ in both rounds.
Fighting, in other words, is a signal of resolve: by paying the price of war today, targets can deter unfavorable offers in the future. The coercer, aware of this incentive, issues a demand in the first round that helps it obtain information about the target’s true resolve. It calibrates its first threat so that tough targets resist while weak targets capitulate. The coercer then uses this new knowledge to make a second-round demand that ensures maximal gains without provoking war. These dynamics play out, however, only when the coercer is bold enough to play a risky first-round strategy. If there is a sufficiently high likelihood that the target is tough to begin with, the coercer will play a conservative strategy in both rounds, gleaning no information whatsoever from the first-round result.

Therefore, despite the coercer’s best efforts, there remains some probability that its threats will be rejected in the first round of the game. We can now calculate the likelihood that this initial encounter will end in coercive failure. That likelihood turns out to be \( \left( \frac{\tau^*}{2} \right)^2 \), which represents the probability of two events occurring in succession: (1) the coercer is sufficiently confident that the target is weak that it issues a “high” demand in the first round, and (2) the target is actually tough, and therefore rejects the demand. In its expanded form, the probability that a coercer’s first-round threat will fail therefore is:

\[
\frac{(\bar{c}_T - c_T)^2}{2(\bar{c}_T + c_C)^2}.
\]

**Implications: Power and Coercion**

The analysis above yields some interesting conjectures about the role of military power in coercive diplomacy. Recall that military power appears in the model indirectly via four terms: \( p \), which represents the balance of military capabilities, \( m_{C+T} \); \( 2c_T \), a tough target’s value for the costs likely to be inflicted on it in wartime; \( \bar{c}_T \), a weak target’s value for an identical amount of military punishment; and \( c_C \), the coercer’s expected costs for fighting. These relationships reflect the fact that greater military power conveys both an enhanced

\[36\text{Because of the information it learns in the first round, however, the coercer is always able to make a successful second-round demand.}\]
ability to inflict costs on one’s opponent, as well as an ability to limit damage to oneself.\textsuperscript{37}

Recall also that we assumed the ratio of the targets’ cost functions – that is, $\frac{c_T}{c_T}$ – to be constant. This assumption is not particularly stringent, tantamount to assuming that a state which values 1,000 battlefield casualties twice as much as another state would also value 10,000 casualties twice as much as the second state.\textsuperscript{38}

In light of these relationships, how does military power influence the likelihood of successful coercion? At first glance, the balance of power appears to have no effect here, since $p$ does not appear in the formula for the likelihood of coercive failure.\textsuperscript{39} Yet, while the balance of power does not appear independently as a parameter in the probability of war, it nonetheless plays a role by influencing the coercer’s costs of fighting – and therefore the demands it is willing to make. More powerful coercers can inflict greater costs on their opponents – and limit damage to themselves. As a result, as a coercer’s military advantage grows, its costs decline while the target’s costs increase.

When we incorporate these effects of military power into the calculations above, we reach a surprising conclusion: \textit{as a coercer’s military power increases, its coercive threats become more likely to fail}. How can this be? The reason is that as the coercer’s power increases, it becomes more profitable for the coercer to assume the risk of a failed threat by playing a “separating” strategy – in other words, a high demand in the first round, instead of a safe demand. War is now less costly for the coercer, so the coercer is more willing to risk it. Further, since the weak and tough target’s costs increase proportionately, the gap between them – that is, the difference between $\underline{c}_T$ and $\bar{c}_T$ – widens. When there is a large difference in resolve between weak and tough targets, determining “who’s who” becomes even more valuable for the coercer. Weak targets can be exploited more in the second round, while tough targets are to be avoided even more diligently. In short: militarily powerful coercers have more to gain and less to lose by making risky, high-stakes threats in coercive diplomacy.

\textsuperscript{37}Slantchev 2003a provides a nice discussion of these two effects of military power. \textsuperscript{38}The conclusions below also hold under some conditions if the assumption of a fixed ratio is relaxed. The appendix contains a proof that derives these conditions. \textsuperscript{39}This is consistent with some previous work on costly-lottery bargaining models in international politics. See, for instance, Fearon 1992 and Powell 1999.
As a coercer’s power increases, it is willing to endure a higher probability of failure because the rewards from success are higher as well. Coercers are willing to launch riskier probes as they grow more powerful.

The idea that military power might be associated with lower rates of coercive success is quite counterintuitive. This conclusion contradicts both traditional scholarship on coercive diplomacy as well as common sense. But it is important to understand the nuance of the argument. Although the balance of power plays a key role in the target’s decision-making process, that is not where the action is. Rather, the effects of military power manifest themselves earlier in the crisis, when the coercer must decide how much to ask for. The coercer has two choices at this stage: issue a high demand that has a large payoff but is also more likely to be rejected, or issue a safe demand that is certain to be accepted but is less lucrative. How does the coercer balance these competing pressures? This is where military power comes into play. When the coercer is weak, it has a low tolerance for risk, and will only issue the “high” demand when it is relatively sure that the target will accept it. This is because the costs of being wrong are high. But as the coercer grows more powerful, those costs dwindle. As a result, its appetite for risk increases, and it becomes willing to issue a “high” demand in situations where it once would have shied away.

Indeed, empirical evidence seems consistent with the notion that coercive challenges from powerful states may be more likely to fail. A database of 210 coercive episodes between 1918 and 2001 indicates that coercers are actually less likely to make successful threats when the target enjoys a military advantage.\(^{40}\) Likewise, coercers with an advantage in nuclear weapons capabilities tend to experience higher failure rates.\(^{41}\) However, this is not because military power is a disadvantage in coercive diplomacy – quite the contrary. Coercers with a military advantage enjoy higher expected payoffs in the crisis bargaining model analyzed above, but they achieve those payoffs in part by taking greater risks.

\(^{40}\) Sechser 2011; see also Sechser 2007, 2016.
Implications: Reputations and Coercive Diplomacy

A second result relates to the role of reputations in international relations. Reputation plays a central role in the model analyzed here: the way a target state behaves in the first round shapes the coercer’s beliefs about its resolve. However, reputation does not always play a role: indeed, the coercer revises its beliefs about the target only when it devises a first-round demand that is specifically designed to generate new information. If it selects the “low” demand, however, the target’s response carries no information about whether it is tough or weak. In other words, under some conditions, targets may capitulate to coercive threats without suffering any damage to their reputations. This carries an important implication for empirical research about reputations: in coercive diplomacy, acquiescence does not necessarily damage a target’s reputation. However, this is not because reputations do not matter in the model; indeed, they matter a great deal. But their primary effect may be to push coercers into courses of action that yield no reputational inferences.

A related observation is even more counterintuitive: coercive success is more likely against targets that have reputations for resolve. The reason is that coercers will only issue a risky “high” demand if it is sufficiently confident that the target is weakly resolved. If the coercer already believes the target to be strong – in other words, if the target’s reputation is good – then it will issue a low demand that is more likely to be accepted. This result obtains because war is an inefficient means of getting what one wants, so the coercer seeks to issue the highest possible demand without prompting a fight. Since coercers have some information about a target’s resolve before formulating a coercive threat, they utilize that information to locate this optimal demand. A successful threat only means that the coercer has done this successfully. In short, successful coercive threats tell us more about the coercer than about the target. We cannot equate capitulation with poor resolve.

The implications of this logic for empirical research are significant. Several scholars

\[ \text{Specifically, as } \tau \text{ increases, it becomes increasingly likely that the coercer will select a pooling strategy, which in equilibrium results in capitulation by both weak and tough targets.} \]

\[ \text{Empirical studies, however, often make precisely this leap, assuming that capitulation implies weak resolve. See, for example, Clare and Danilovic 2010.} \]
have observed that states often do not develop reputations for being irresolute when they capitulate to coercive threats.\textsuperscript{44} When scholars observe a successful coercive demand – in other words, acquiescence – it may only mean that the coercer was intimidated enough to moderate its demands ahead of time. In this scenario, coercers may draw no inferences at all from the conciliatory behavior of their opponents. But again, this does not mean that reputation plays no role. The lesson here is that reputations may indeed matter, but they may operate primarily prior to crises, rather than during them.\textsuperscript{45}

\textit{Implications: Threats and Reassurances}

Schelling observed in \textit{Arms and Influence} that reassurance is critical to success in coercive diplomacy: “any coercive threat requires corresponding assurances” that acquiescence will not invite future depredations.\textsuperscript{46} The analysis here is consistent with this view. The two-round game revealed an often-overlooked dynamic about coercive diplomacy: targets must consider not only the issues at stake in today’s crisis, but also how their actions might influence the outcomes of future crises. In the model, targets that believed future crises to be likely were more difficult to coerce than targets that believed today’s crisis to be the last. It would be a mistake, then, for coercers to think solely about how to make their threats more potent. While threats are undoubtedly a central aspect of coercion, reassurances can also play a critical role. Coercers that can credibly promise not to make further demands are likely to be more successful in coercing their adversaries. The challenge for coercers, then, is to find ways to credibly tie their own hands and make promises of restraint believable.\textsuperscript{47}

\textbf{Limitations and Conclusions}

Explaining coercive outcomes requires that we understand the decisions made both by coercers and by coercive targets. In other words, coercion is a process of strategic bargaining.

\textsuperscript{44}For example, Snyder and Diesing 1977; Hopf 1995; Mercer 1996; Press 2005.
\textsuperscript{45}See also Sechser 2010, which explores the dynamics of reputation in coercive diplomacy in greater detail.
\textsuperscript{46}Schelling 1966, p. 74. For similar views, see Davis 2000, p. 24; Art 2003, p. 366; Knopf 2012.
\textsuperscript{47}Sechser 2016.
In the competitive setting of coercive diplomacy, actors must consider not only their own objectives, but also the anticipated reactions of their adversaries. This chapter has explored a series of models of coercive diplomacy, in which two actors bargain in the shadow of military conflict. In particular, the discussion paid special attention to the strategic calculations of coercers, in an attempt to understand the conditions under which coercion succeeds and fails. The results of this analysis have several implications for how we understand the outcomes of coercive episodes.

First, coercive outcomes depend critically on choices made in the early stages of a crisis. The first of these choices is the coercer’s decision to initiate the crisis, and secondarily, what (or how much) to demand. While a coercer would like to win the largest possible concession from the target, demanding too much could result in a costly conflict. The possibility of war disciplines the coercer’s calculations, compelling it to consider the target’s likely reactions when formulating a threat. Coercive diplomacy is fundamentally about balancing risk and reward.

Second, information plays a starring role in coercive bargaining. Coercers and targets base their decisions around assumptions about the adversary’s capabilities and resolve, but uncertainty can complicate these calculations. In reality, it can be difficult to know the other side’s military strength, willingness to fight, and values. Coercive outcomes therefore depend not simply on the actors’ military capabilities and resolve, but on their ability to learn (or communicate) this information. This insight points scholars of coercive diplomacy in two related directions. First, the study of coercion would benefit greatly from the insights of scholarship on intelligence. A coercer’s ability to obtain information about its opponent can mark the difference between success and failure, so understanding intelligence capabilities is central to explaining coercive outcomes. Second, a better understanding of signaling would improve our theories of coercion. While there is an extensive literature on signaling in international relations, the insights of this literature have not yet been fully integrated into the study of coercive diplomacy. Scholars and practitioners alike would benefit from

---

48 See, for example, Austin Long’s chapter in this volume.

49 See, for example, Jervis 1970; Morrow 1989; Morrow 1999a; Fearon 1994; Fearon 1997. For examples
a sharper understanding of what signals are available to leaders, and how effective they are.

Third, the role of military power in coercive diplomacy is more complex than scholars have acknowledged. In the traditional story about coercion, power plays a role primarily in intimidating the adversary. Actors that possess greater military capabilities can threaten more punishment, in this view, and therefore can more easily compel their adversaries to make concessions. This story, however, is incomplete. It overlooks an important fact: power influences not only the adversary’s calculations, but also one’s own calculations. Military power simultaneously intimidates and emboldens. Actors with more power can make bigger demands and take greater risks. The implications for the study of coercion are profound: instead of being associated with success, military power may be associated with coercive failure. Stronger actors can afford to take greater risks, leading to higher rates of failure.

*Where Do We Go from Here?*

The analysis above also highlights several areas for further exploration in the study of coercive diplomacy. The first area involves coercion involving non-state actors. The modeling exercise in this chapter generally treated the primary actors as states, but this need not be the case. In principle, the model ought to apply equally well to coercive encounters involving insurgent groups, actors in civil wars, terrorist organizations, or other non-state entities. At the same time, it is possible that these types of entities are a poor fit for the assumptions used to derive the conclusions in this chapter. For example, the concept of a “winner-take-all” conflict to resolve a dispute that escalates to war makes little sense when the target is a terrorist organization, since groups such as al Qaeda likely resort to terrorism precisely because they stand no chance of defeated conventional armies in open battle. Other assumptions of the model may prove ill-fitting as well. Nevertheless, the model offers a useful starting point for thinking about how coercion involving non-state organizations may resemble – or deviate

---

50 Several of the chapters in this volume, including contributions by Keren Fraiman; Peter Krause; and James Igoe Walsh, evaluate the logic of coercion in the context of insurgencies.
from – models of interstate coercive diplomacy.

On a similar note, more research is needed in the area of non-military coercion. Whereas the analysis in this chapter emphasized coercive bargaining in the shadow of military force, states in fact have many coercive tools at their disposal, including economic sanctions, cyber warfare, and other forms of economic and political punishment. Do standard models of military coercion apply to these scenarios? If not, how do these models need to be revised in order to account for non-military coercive tools?

Finally, one limitation of the foregoing analysis is the absence of third-party actors. Third parties – including allies, other adversaries, neutral observers, and domestic political actors – often play a central role in coercive encounters. The two-actor engagements analyzed in this chapter surely are too simple, overlooking a rich set of dynamics involving outside actors. A fruitful avenue for coercive diplomacy research would be to begin incorporating these actors into our models of coercion.

51 For sophisticated discussions of other forms of coercion, see the chapters by Alexander B. Downes, Daniel W. Drezner; Kelly M. Greenhill; and Jon R. Lindsay and Erik A. Gartzke in this volume.
52 See, for instance, Timothy Crawford’s chapter in this volume.
Appendix

This appendix describes the formal solutions to the three versions of the bargaining games presented in the chapter text. The solution concept employed in the first two cases is a subgame perfect Nash equilibrium. The solution in the third case employs a perfect Bayesian equilibrium.

Equilibrium Solution for the Basic Model

**Claim.** In equilibrium the coercer’s demand, $x$, is equal to $p + c_T$. The target accepts if $x \leq p + c_T$ and rejects otherwise.

**Proof.** We begin at the end with $T$’s decision. $T$ may either accept $C$’s demand $x$ or reject, thereby selecting the outside option of war, which carries an expected payoff of $1 - p - c_T$. Assuming slight risk-aversion, we stipulate that both players prefer the peaceful option when indifferent, deducing that $T$ will accept any offer carrying a payoff equal to or greater than war. $T$ therefore capitulates whenever $x \leq p + c_T$.

Working backward, $C$ must now choose between the best available peaceful settlement and war. War carries an expected payoff of $p - c_C$, which $C$ recognizes will result whenever $x > p + c_T$. Also being slightly risk-averse, $C$ will therefore choose a peaceful settlement so long as $x \geq p - c_C$.

For a peaceful settlement to obtain, $C$’s demand must now meet two criteria: $x \geq p - c_C$ and $x \leq p + c_T$. Since both $c_C$ and $c_T$ are both positive by assumption, $-c_C \leq c_T$ and the set of demands such that $p - c_C \leq x \leq p + c_T$ is therefore non-empty. With the privilege of issuing a take-it-or-leave-it demand, $C$ chooses the maximal demand acceptable to $T$, $x = p + c_T$, which $T$ accepts.

Equilibrium Solution for the First Modification (One Round, Incomplete Information)

**Claim.** In equilibrium, the coercer demands $\bar{x} = p + \xi_T$ if

$$\tau < \frac{\xi_T - \xi_T}{\xi_T + c_C}$$

and $\bar{x} = p + \xi_T$ otherwise. A weak target (i.e., $c_T = \xi_T$) accepts the demand if $x \leq \bar{x}$ and rejects otherwise; similarly, a tough target ($c_T = \xi_T$) accepts the demand if $x \leq \bar{x}$ and rejects otherwise.

**Proof.** Extrapolating from the logic in the basic model, a “tough” target (i.e., $c_T = p + \xi_T$) will accept the coercer’s demand if $x \leq p + \xi_T$, whereas a “weak” target will accept if $x \leq p + \xi_T$.

Likewise extrapolating from the solution to the basic model, $C$’s equilibrium demand would be $x = p + \xi_T$ if it knew for certain that $T$ was tough, and $x = p + \bar{x}_T$ if it knew $T$ was weak. However, since it only knows that $T$ is tough with probability $\tau$, we must determine $C$’s strategy when $0 < \tau < 1$. At the critical value $\tau^*$, the coercer is indifferent between the only two non-dominated strategies available: $\bar{x} = p + \xi_T$ and $\bar{x} = p + \bar{x}_T$. We locate the critical value by setting the expected utilities from these demands equal and solving for $\tau^*$:

$$\tau^* (p - c_C) + (1 - \tau^*) (p + \xi_T) = \frac{p + \xi_T}{\xi_T - \xi_T + c_C}.$$

Henceforth $\bar{T}$ will refer to a tough target, and $\bar{T}$ to a weak one.

---

---
Equilibrium Solution for the Second Modification (Two Rounds, Incomplete Information)

Claim.

1. Targets’ Strategies. A tough target accepts in either round if \( x_n \leq p + \xi_T \) and rejects otherwise. A weak target accepts in the first round if \( x_1 \leq p + \xi_T - \delta (\xi_T - \xi_T) \) and rejects otherwise. In the second round, a weak target capitulates if \( x_2 \leq \xi_T \) and rejects otherwise. A tough target accepts in round 2 if \( x_2 \leq \xi_T \) and rejects otherwise.

2. Separating Equilibrium. \( C \) issues the separating demand \( x_1 = p + \xi_T - \delta (\xi_T - \xi_T) \) if \( \tau < \frac{\xi_T - \xi_T}{\xi_T + \xi_T} \).

3. Pooling Equilibrium. \( C \) issues the pooling demand \( x_1 = p + \xi_T \) if \( \tau \geq \frac{\xi_T - \xi_T}{\xi_T + \xi_T} \).

Proof. The appropriate solution concept here is a perfect Bayesian equilibrium. We begin by conjecturing a separating equilibrium and locating \( C \)'s optimal second round demand, \( x_2 \). If \( C \) is able to separate in the first round, its strategic options mimic those in the single-stage, complete-information model: if \( T \) rejects in round 1, then \( x_2^* = p + \xi_T \), whereas \( x_2^* = p + \xi_T \) if \( T \) capitulates in the first round. In both cases, the target will capitulate in equilibrium (using the solution to Variant 1 above).

To characterize a separating equilibrium, we must locate a demand that a tough target would resist but a weak target would accept. First consider the decision of a weak target in the first round. The coercer’s optimal first round demand should make a weak target exactly indifferent between capitulation and fighting. This optimal demand, which we label \( x_1^* \), is calculated as follows:

\[
EU_T (Accept) = EU_T (Reject) \\
(1 - x_1^*) + \delta (1 - p - \xi_T) = (1 - p - \xi_T) + \delta (1 - p - \xi_T) \\
x_1^* = p + \xi_T - \delta (\xi_T - \xi_T) .
\]

How does a tough target react to this demand? Recall that a tough target will fight in the first round (knowing that \( C \) will demand \( x \) in the second round) if the following condition holds:

\[
EU_T (Accept) < EU_T (Reject) \\
(1 - x_1^*) + \delta (1 - p - \xi_T) < (1 - p - \xi_T) + \delta (1 - p - \xi_T) \\
x_1^* > p + \xi_T.
\]

We can now substitute \( p + \xi_T - \delta (\xi_T - \xi_T) \) for \( x_1^* \) and simplify:

\[
x_1^* > p + \xi_T \\
p + \xi_T + \delta (\xi_T - \xi_T) = p + \xi_T \\
(1 + \delta) (\xi_T - \xi_T) = 0 \\
\xi_T > \xi_T .
\]

Since \( \xi_T > \xi_T \) by assumption, this condition holds. We therefore know that \( x_1^* \) successfully separates the target types because a weak target will capitulate while a tough target will fight.

The next step is to determine when \( C \) will wish to separate (rather than pool), given its uncertainty about \( T \)'s type (represented by \( \tau \), the probability that \( c_T = \xi_T \)).

In a pooling equilibrium, both target types are willing to accept \( C \)'s demand in the first round. Having already stipulated that a tough \( T \) will only accept demands up to \( p + \xi_T \) in either round, we can deduce that \( C \) will demand precisely that amount in both rounds. Thus the expected payoff for either type of \( T \) in the pooling equilibrium is \( (1 + \delta) (1 - p - \xi_T) \). For \( C \), the expected payoff of such an equilibrium is \( (1 + \delta) (p + \xi_T) \).
We can now determine the critical value $\tau^*$ which makes $C$ indifferent between the pooling and separating equilibria.

\[
\begin{align*}
\text{EU}_C (\text{Separate}) &= \text{EU}_C (\text{Pool}) \\
\tau^* \left[ p - c_C + \delta (p + \xi_T) \right] + \\
(1 - \tau^*) \left[ p + \bar{\sigma}_T - \delta (\bar{\sigma}_T - \xi_T) + \delta (p + \bar{\sigma}_T) \right] &= (1 + \delta) (p + \xi_T) \\
\therefore \quad \tau^* &= \frac{p - \xi_T}{c_C + \bar{\sigma}_T}.
\end{align*}
\]

\[\square\]

**The Probability of Coercion Failure**

**Claim.** The *ex ante* probability of coercion failure in the first round of the two-round, incomplete information model is $\frac{(\tau^*)^2}{2}$, or

\[
\frac{(\bar{\sigma}_T - \xi_T)^2}{2 (\bar{\sigma}_T + c_C)^2}.
\]

**Proof.** Coercion failure occurs only in first round of the separating equilibrium, so the probability of war is equal to $\Pr(\text{Separation}) \times \Pr(\text{Failure}|\text{Separation})$.

The probability of separation is easily determined. Separation occurs below $\tau^*$, and since $\tau$ is drawn from an even distribution across $[0, 1]$, the probability that $\tau < \tau^*$ is simply $\tau^*$.

Given that $C$ opts for separating strategy, the expected probability that the target is tough is equal to the expected value of $\tau$ at this stage. Since $\tau$ must be drawn from an even distribution across $[0, \tau^*)$, given that $C$ separates, the expected value of $\tau$ is simply $\frac{\tau^*}{2}$.

The *ex ante* probability of coercion failure is the product of these two factors: $\frac{(\tau^*)^2}{2}$.

\[\square\]

**Power and Coercive Outcomes**

The first two claims deal with the effects of an increase in the coercer’s power, caused either by an increase in the coercer’s absolute power $m_C$, a decline in the target’s absolute power $m_T$, or both. Two effects obtain from this change: first, a (proportional) increase in weak and strong targets’ costs for war, and second, a reduction in the coercer’s costs for war. Both independently increase the likelihood of war.

**Claim.** If $\bar{\sigma}_T$ and $\xi_T$ increase by a proportional amount, the probability of coercion failure $\frac{(\tau^*)^2}{2}$ will increase.

**Proof.** Recall that the ratio of costs $\frac{\bar{\sigma}_T}{\xi_T}$ is assumed to be constant. Since we are considering the effect of a proportional increase in the cost terms $\bar{\sigma}_T$ and $\xi_T$, let $\alpha$ be the multiple by which both cost terms increase, such that $\alpha > 1$.

The proof proceeds by contradiction. Begin by assuming the contrary: a proportional increase in $\bar{\sigma}_T$ and $\xi_T$ does not increase the probability of coercion failure:

\[
\frac{(\bar{\sigma}_T - \xi_T)^2}{2 (\bar{\sigma}_T + c_C)^2} \geq \frac{(\alpha \bar{\sigma}_T - \alpha \xi_T)^2}{2 (\alpha \bar{\sigma}_T + c_C)^2}.
\]

Since all terms in the inequality are positive by definition and $\bar{\sigma}_T > \xi_T$, the numerators and denominators must all be positive. The statement therefore simplifies to:

\[\square\]
A Bargaining Theory of Coercion

\[ \frac{\bar{c}_T - c_T}{\bar{c}_T + c_C} \geq \frac{\alpha \bar{c}_T - \alpha c_T}{\alpha \bar{c}_T + \alpha c_C}. \]

This reduces to the inequality \( \alpha \leq 1 \), which contradicts the original assumption \( \alpha > 1 \).

Claim. If \( c_C \) declines, other things being equal, the overall probability of coercion failure \( \frac{(\tau^*)^2}{2} \) will increase.

Proof. Again, the proof proceeds by contradiction. Begin by assuming the contrary: a decrease in \( c_C \) by the factor of \( \beta \) (such that \( \beta < 1 \)) does not increase the probability of failed coercion:

\[ \frac{(\bar{c}_T - c_T)^2}{2(\bar{c}_T + c_C)^2} \geq \frac{(\bar{c}_T - c_T)^2}{2(\bar{c}_T + \beta c_C)^2}. \]

The inequality reduces to:

\[ \frac{1}{c_C} \geq \frac{1}{\beta c_C}, \]

therefore implying that \( \beta \geq 1 \), which contradicts the original assumption.

The assumption thus far has been that the ratio of costs for tough and weak targets remains fixed – in other words, that \( \frac{\bar{c}_T}{c_T} \) stays constant. But what happens if the relative costs of war for tough and weak targets change in a non-linear fashion – that is, if \( \frac{\bar{c}_T}{c_T} \) also changes as the costs of war change?

Claim. For the probability of coercion failure, \( \frac{(\tau^*)^2}{2} \) to increase, it must be the case that:

\[ c_C \left[ (\bar{c}_T - c_T) - (\xi_T - \xi_T) \right] + \bar{c}_T \xi_T - c_T \xi_T > 0. \]

Proof. As established above, for the probability of coercion failure to increase, the following statement must be true:

\[ \frac{(\bar{c}_T - c_T)^2}{2(\bar{c}_T + c_C)^2} > \frac{(\bar{c}_T - c_T)^2}{2(\bar{c}_T + \xi_T)^2}. \]

Also as before, since the numerators and denominators are all positive, we have:

\[ \frac{\bar{c}_T - c_T}{\bar{c}_T + c_C} > \frac{\bar{c}_T - c_T}{\bar{c}_T + \xi_T}. \]

This inequality ultimately reduces to

\[ c_C \left[ (\bar{c}_T - c_T) - (\xi_T - \xi_T) \right] + \bar{c}_T \xi_T - c_T \xi_T > 0. \]
References


Fraidman, Keren. 2015. “The Opportunities of Base State Coercion.” Typescript, MIT.


A Bargaining Theory of Coercion


