

Introduction

When Do Threats Work?

Coercion is the practice of convincing a target by the use of threats to bend to your will. From a library imposing late fees on tardy patrons to a parent sending an insolent child to their room, everyday life brims with the logic of coercion: obey me, or else. In their foreign policies, states often make grave economic or military threats to try to affect others' behavior rather than resort to costly war. It is little wonder, then, that states frequently choose coercion.

What is more surprising is how bad they are at it. Depending on how one counts, US threats, historically, have succeeded just 18 percent,¹ 29 percent,² or 31 percent³ of the time.⁴ More broadly, "stronger" coercers have achieved their compellent aims only 36 percent of the time.⁵ Material power, it would seem, does not portend coercive success.

These striking figures fly in the face of how coercion is typically understood to work. A dominant paradigm explains coercive outcomes by pointing to the credibility and severity of threats. Targets will defy a coercive demand if they think the threat is a bluff. And even if a target believes that a threat is credible, it may still defy it if it thinks the expected pain is bearable. Threats must therefore be serious and severe.

This threat-centric paradigm is clear in theory and intuited by practitioners. But it is woefully incomplete. This book advances another paradigm, which points to the implicit logic of contingent action at the heart of all coercion: *coercive assurance*.⁶ Threats must be perceived by their targets as *conditional* upon their behavior. "They'll have me whipp'd for speaking true; thou'lt have me whipp'd for lying; and sometimes I am whipp'd for holding my peace," laments King Lear's fool.⁷ Far from being cowed, he finds freedom in his lot—jesting, boasting, and speaking truth to power. Pain sheds its coercive power when it cannot be avoided. Even credible and severe threats will fail if they are not perceived as sufficiently conditional.⁸

To develop the concept of coercive assurance in the study of international relations, this book pursues three questions. First, what is the relationship between threats and assurances in coercion? Second, why do coercers punish after receiving compliance? Third, why do targets of coercion fear unconditional pain? Addressing these questions involves investigating coercive assurance from the perspective of target states: how they receive coercive threats, how they evaluate their choices in light of coercion. But it also involves recognizing that coercers confront, are stymied by, and try to overcome a predicament inherent in cases of coercion: the *assurance dilemma*.

Argument in Brief

The assurance dilemma is a situation in which the actions coercers take to bolster the credibility of a threat undermine the credibility of their assurance not to punish. It is a trade-off that coercers must navigate to wrest concessions. Simply put, when leaders augment their threats, they do so at the expense of their necessary assurances, thereby unwittingly compromising their own coercive strategies.

Establishing the credibility of coercive assurance is challenging because coercers may end up punishing their targets unconditionally for several rational and nonrational reasons. First, targets of coercion typically face multiple coercive demands at once, creating the opportunity for coercers to continue threatening targets who have made concessions on some issues but not others. Second, targets often confront multiple coercers with different interests and independent capacities to punish. They therefore worry whether conceding to one coercer will actually avoid a punishment imposed by another. Third, when considering whether to concede to coercion, targets fear that they will reveal new information to their coercers through their concessions and only invite further threats.

From the perspective of targets, these sources of unconditional punishment are impediments to making concessions and avoiding pain. Yet coercers can mitigate all of these assurance challenges. From the perspective of coercers, this book demonstrates the effectiveness of three corresponding strategies of signaling coercive assurance. First, coercers can disentangle multiple demands from the same punishment such that partial concessions still avoid some pain. I call this “disentangling demands.” Second, coercers can demonstrate control over potential spoilers to signal that they alone will decide whether to impose or relieve punishments. I call this “managing spoilers.” Third, coercers can share with a target their knowledge of its misdeeds to communicate that its concessions will not reveal new information to the coercer. I call this “sharing knowledge.”

Finally, assurances need not be perfectly credible to affect the outcome of coercion. Every concession includes a bet that punishment will be

avoided. Targets cannot be truly certain that their coercer is sincere, but they accept bigger risks of unconditional harm to avoid severer punishments. Nevertheless, some coercive assurance is always necessary because targets do not make concessions to coercion if they conclude that they will be punished anyway. In a dynamic coercive bargaining process, targets look for assuring signals that their coercers will withhold the threatened pain. These powerful dynamics explain when and why targets of coercion concede or defy.

Improving Our Understanding of Coercion

Coercive assurance matters a great deal to the study of international politics because coercion is so prevalent. An international order underwritten by great powers that provide public goods to incentivize cooperation also relies heavily on coercive tools.⁹ And much more of international competition is coercive than directly violent. States tend to issue threats before turning to force. But the consequences of failed coercion are grave: crises, tit-for-tat escalation, war.

More broadly, coercion is everywhere around us. Most of what we mean when we say “you can’t park there” or “you must pay your taxes” is not really that you cannot or must not—it is that you would be punished if you did not comply. When we say, “I don’t have a choice but to drive on the correct side of the road,” we have it exactly backwards. Credibly conditional pain has presented us with a clear choice.

The necessity of coercive assurance has been underappreciated in the study and practice of coercion. In the coercion literature, it is an understudied type of commitment problem, and the relationship between credible commitments and credible threats has been undertheorized. This is in part because assurance is assumed to be automatic in the bargaining model of war, which the field extended to coercion. Punishments are costly to carry out, the logic goes. Coercers should not pay those costs if their target already backed down and they got what they wanted, so targets should not fear that they will be punished after they comply. In reality, targets can fear unavoidable pain and thus defy credible threats.

Threat-centric biases also inform how practitioners approach coercion. US National Security Strategy documents, which since 1987 have communicated signals of US interests and intent to allies, adversaries, and the public, eschew the logic of coercive assurance.¹⁰ The documents contain an average of thirteen times more threats than coercive assurances (see figure 0.1).¹¹ Scholars have observed a similar rejection of coercion theory in general within the US military. “Military professionals are not entirely comfortable with violence as a bargaining process,” writes Tami Biddle. “One does not, they believe, ‘bargain’ with one’s enemies—one fights them.”¹²

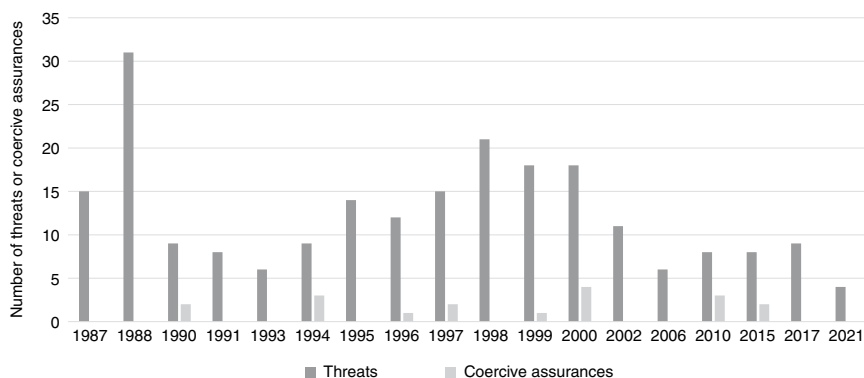


Figure 0.1. Frequency of threats and coercive assurances in US National Security Strategy documents.

A common analogy in statecraft to the game of chess also effectively scopes out the problem of coercive assurance. As Secretary of State John Foster Dulles explained in a 1957 National Security Council meeting on US nuclear strategy, “in a chess game you wouldn’t normally ever go so far as to take your opponent’s king; you checkmate the king and don’t play out the rest.”¹³ In Dulles’s view, effective nuclear counterforce was a checkmate that would lead to bargaining advantages. In the real world, cornered opponents can still defy or lash out.

The US government faces a particularly acute assurance dilemma. It has trouble coercing weak countries because of its strength, not in spite of it. In Pyongyang US threats are perceived as a ploy, in Tehran as a trap, in Caracas as a catch-22. Yet viewing the history of coercive diplomacy through the lens of coercive assurance reveals that even Washington’s targets are more likely to concede when leaders come to acknowledge and address the assurance dilemma. The credibility of coercive assurance will matter a great deal to the achievement of important US foreign policy aims, from deterring a revanchist Russia, to restraining the North Korean nuclear arsenal, to avoiding a war with China over Taiwan.

This book, therefore, seeks to improve the study and practice of coercive diplomacy. When threats put targets in a “damned if you do, damned if you don’t” position, we should expect defiance.

Studying Coercion in the History of Nuclear Nonproliferation

The spread of nuclear weapons is a consequential challenge of our time, one that many leaders have devoted significant energy and resources to combat. Proliferators, often motivated in their pursuit by security fears, are

loath to give up acquisition of a security asset, especially when faced with coercive threats from the very enemies whose uncertain intentions they feared in the first place. Findings on which signals help to overcome the assurance dilemma in the high-stakes domain of counterproliferation are more likely to travel outside of it to coercive bargaining over other, lower stakes. The historical record also permits process tracing to reveal the intentions and perceptions of coercers and targets over time.

On at least twenty-one occasions, states have issued threats of sanctions or force to enforce the nonproliferation regime. This book delves into four cases of coercive bargaining between nonallies over nuclear weapons programs in South Africa, Iraq, Libya, and Iran. Exploiting within-case variation, I explain not only the occurrence but also the timing of nonproliferation bargains using primary documents from US government archives, the South African apartheid-era government archives, tape recordings of high-level meetings from Saddam Hussein's Iraq, and the International Atomic Energy Agency (IAEA) archives. I supplement these documents with the memoirs, recollections, and writings of target state policymakers, military leaders, and nuclear scientists. I also conducted interviews with policymakers who participated in the coercive bargaining processes.

In each chapter, the lens of the assurance dilemma reveals novel insights that revise conventional wisdom. Each case offers strong evidence that target perceptions of noncredible assurance hinder coercers, even when those targets perceive credible threats. Increases in the perception of assurance credibility, not threat credibility, are most proximately associated with acquiescence. Coercive assurance can tip the balance one way or another.

Plan of the Book

The book proceeds as follows. In chapter 1, I critique the threat-centric coercion paradigm and derive theory on the assurance dilemma as an impediment to successful coercion. I explain why targets fear that insincere coercers may choose to punish unconditionally and how even sincere coercers can lose control over the decision to apply or withhold punishment. The chapter concludes with case selection and a rationale of the book's methods and case studies.

Chapter 2 concerns South Africa, which succeeded in building a secret nuclear arsenal despite international pressure to sign the Nuclear Nonproliferation Treaty (NPT). Conventional explanations for this outcome point to South Africa's fears of Soviet domination and US indifference to anti-communist proliferation. Instead, internal regime documents show how counterproliferation pressure on South Africa failed in the 1970s and 1980s because of entangled demands. Owing to intense coercive demands to end its brutal practice of apartheid, Pretoria's leaders perceived that

acquiescence on the nuclear issue would provide no relief from the pain of economic sanctions. The Ronald Reagan administration attempted to disentangle the issues in the mid-1980s until Congress passed comprehensive antiapartheid sanctions legislation. The assurance dilemma thus proves a powerful lens to explain the failure of coercion in South Africa. Later, South Africa's dismantlement of its nuclear arsenal is sufficiently explained by its racist motivation not to hand over nuclear weapons to a postapartheid Black-majority government. Primary documents reveal as well how from 1989 to 1993 South African leaders planned and executed a strategy to conceal their past weaponization from the IAEA.

Substantial documentary records of the South African nuclear weapons program are now available to scholars, and many South African policy-makers and nuclear scientists have written firsthand accounts of their nuclear decision-making. The case is ripe for reexamination and theory testing. South Africa's nuclear journey was more interactive with external coercers than traditional accounts of its internal political machinations about nuclear weapons suggest.

Chapter 3 examines the case of Iraq, where a mercurial dictator in Baghdad sought nuclear weapons and came close to acquiring the capability. Saddam Hussein also came under intense counterproliferation pressure, up to and including noncoercive brute force. Conventional wisdom paints a tragedy of a duplicitous and quixotic dictator hiding his nuclear pursuits until exasperated great powers chose to invade. The reality is even more frustrating. After the Gulf War, Washington attempted to coerce Saddam into abandoning a weapons of mass destruction (WMD) program and at first to great success. Iraq destroyed its program and admitted inspectors, but it would not come totally clean. Its concessions consistently matched what it thought its coercers already knew about its past WMD programs. Yet it was never deemed enough. As the 1990s proceeded, Iraqi elites came to conclude that no amount of compliance would lift sanctions. In the midst of this coercive effort, Saddam framed his choices to his advisers: Iraq could either "have sanctions with inspectors or sanctions without inspectors."¹⁴ He perceived no escape and so confidently defied.

Iraqi documents and recordings, captured and shipped out of the country by US forces after the 2003 invasion, offer a rare glimpse into the perspective of a target of coercion. In meetings with his advisers, Saddam evaluated his evolving position, the intentions of his enemies, and options available for compliance or defiance.

Chapter 4 explains how, in bargaining with Libya, the United States and Britain succeeded in coercing a rogue proliferator into abandoning its nuclear pursuits. The timing of the final bargain after the 2003 invasion Iraq has led to some shallow lessons drawn from the case about the importance of demonstrating highly credible threats—that is, invading another country in the region over similar stakes. The reality of the strategic interaction was

far more consistent with the assurance dilemma. During negotiations, Muammar Qaddafi was continually suspicious that his coercers intended to disarm him and attack anyway. The invasion of Iraq and capture of Saddam Hussein fed these fears even more. Washington and London overcame this perception slowly, taking pains to assure Qaddafi. They froze out spoilers and privately shared with Libya intelligence gleaned from their penetration of the black-market A. Q. Khan network to assure him that his conceding would not reveal more than they already knew. Washington and London worked hard to convince Libya that it would not be the next Iraq. But eight years after striking a nonproliferation bargain, US and British airpower did help to topple Qaddafi. Despite reneging on their bargain, the evidence reveals that their signals of coercive assurance were critical to closing the 2003 deal and convincing Qaddafi to accept the risks of conceding.

Fewer primary records are available in the Libyan case. The fate of the official government archives after the 2011 civil war is unknown. Nevertheless, US and British participants have reflected on the case since and are willing to discuss the 2003 deal in writings and interviews because the Qaddafi government is gone. The case shows that even the most credible and severe threats alone cannot succeed without coercive assurance. Spikes in threat credibility—after the invasion of Iraq and the capture of Saddam—resulted not in acquiescence by Qaddafi but demands for assurance.

Chapter 5 examines Iran and its primary coercer, the United States. In 2015 Iran struck a bargain with the P5+1 coalition (the US, the UK, France, Russia, China, and Germany) to accept enhanced verification and limits on its nuclear program. It is puzzling why Iran was willing to strike a nonproliferation bargain so soon after the demise of Libya's Qaddafi. Conventional explanations for the Iran deal emphasize leadership changes and a window for bargaining—the elections of Barack Obama and Hassan Rouhani—as well as the inclusion of significant carrots in the deal for Iran. The lens of the assurance dilemma highlights instead Iran's fears of its coercers' duplicity and how the strategies of coercers changed to overcome these suspicions. For instance, another adversary, Israel, was not party to the negotiations and was a potential spoiler. I show that Tehran came to the table only after the window of a credible Israeli threat closed—a puzzle for threat-centric theory. The Obama administration also disentangled its demands and punishments to make clear that it was negotiating over Iran's nuclear program and not its missiles or foreign policy. Iran sought and acquired the specific separation of entangled sanctions designations during negotiations. Moreover, the Obama administration crafted domestic legislation to bound congressional oversight over sanctions relief. And documents from Iran's "atomic archive" compared with IAEA reports reveal how much coercers already knew about the "possible military dimensions" of Iran's past nuclear programs before the 2015 nuclear deal.

Earlier, in 2003, Iran had scaled back a nuclear weaponization program because of highly credible and severe military threats bolstered by the US invasion of Iraq. Iran curtailed its nuclear program in the absence of attempts to assure it. But Iran's concessions were only partial. The nuclear program transitioned underground and to more deniable pursuits. Iran's leaders then put out feelers for coercive assurance. The 2003 episode shows how assurance is necessary but that its credibility can vary with the severity of threatened punishments.

Research on the Iran case is important and possible because of the successful conclusion and subsequent US withdrawal from the 2015 Joint Comprehensive Plan of Action (JCPOA). Interviewees were motivated to speak to either defend or criticize the deal. Moreover, IAEA reports and a controversial tranche of Iranian documents stolen by Israeli intelligence in 2018 shed new light on technical aspects of the former Iranian nuclear weapons program.

The book concludes with lessons for theory and policy and examines extensions of the logic of coercive assurance outside of the nuclear domain—from the origins of the Pacific War to the Cuban missile crisis, ransomware, and beyond.