How Could States Use Nuclear Weapons?

Four Models After the Bomb*

Alexandre Debs†

October 6, 2021


*For excellent comments and suggestions, I thank Jean-Francois Belanger, Tyler Bowen, Dale Copeland, Matthew Fuhrmann, Eliza Gheorghe, Alexander Lanoszka, Gerard Padro-i-Miquel, Robert Powell, Or Rabinowitz, Kenneth Scheve, Todd Sechser, Joshua Shifrinson. All remaining errors are my own.

†Department of Political Science, Rosenkranz Hall, Room 311, Yale University, PO Box 208301, New Haven, CT, 06520-8301. Email: alexandre.debs@yale.edu
Introduction

How could states use nuclear weapons? Do greater nuclear capabilities translating into greater coercive leverage?

These are some of the most important questions in nuclear politics. According to the Theory of the Nuclear Revolution (TNR), nuclear weapons are a force for peace (Brodie 1946; Bundy 1988; Glaser 1990; Jervis 1989; Powell 1990; Schelling 1966; Waltz 1981, 1990). They are better to deter than to compel. Greater resolve, not greater military capabilities, determines the outcome of crises.

TNR has dominated international relations scholarship for decades. It is currently under attack. Some claim that it is a “myth,” a “failed” revolution (Green 2020; Lieber and Press 2020). Washington came close to escaping the conditions of Mutually Assured Destruction (MAD) by the late Cold War (Green and Long 2017). It was able to track and target submarines
and mobile missiles (Long and Green 2015), even develop a Hard-Target Kill counterforce doctrine (Green and Long 2016). Its future nuclear dominance could be even more pronounced, as the computer revolution further eroded the survivability of nuclear forces, enabling the United States to shield itself from retaliatory strikes (Lieber and Press 2006, 2017). In fact, critics argue, Washington’s pursuit of nuclear superiority is rational. Superiority helps coerce enemies (Green 2020, Kroenig 2018, Lieber and Press 2020) and prevent proliferation among allies (Gavin 2020).

Yet despite their spirited attack, these critiques are very narrow in scope. Most concede that nuclear weapons have been a force for peace and that they are the ultimate tools of deterrence. They do not explain how nuclear compellence could succeed, and how greater nuclear capabilities bolsters compellence. This is surprising. TNR scholars themselves were tentative in arguing for the superiority of deterrence. For Thomas Schelling, deterrence is easier than compellence if countries know each other’s resolve. “The resulting world - the world without uncertainty - would discriminate in favor of passivity against initiative. It is easier to deter than to compel” (Schelling 1966, 100, emphasis in the original). But in the real world, uncertainty abounds, and countries do not know each other’s resolve. In that world, Schelling (1966, 101-103) hesitated to proclaim the superiority of deterrence. If a state triggered a crisis, couldn’t it signal its greater resolve and extract concessions? Robert Jervis wondered, “if the aggressor does take the initiative, is there any reason to think that it is easier to deter than to compel?”
Under some circumstances, he concludes, defending the status quo “is no easier than altering it” (Jervis 1989, 34). Thirty years later, we still have much to learn in understanding how compellence could overcome deterrence.

I address this problem in a game-theoretic model where states exchange compellent and deterrent threats, bargain over an issue in dispute, and enter a preemption game. I show in a baseline set-up that, consistent with TNR, compellence fails because the coercer would never trigger a nuclear disaster. Then I allow the compellent state to bolster the credibility of its threats through standard techniques: *burning bridges*, removing its ability to back down; *probabilistic threats* or *brinkmanship*, increasing the risk of disaster following non-compliance; and the *rationality of irrationality*, feigning a preference for disastrous war over dishonorable peace (Schelling 1966, 37, 43, 91). I show that compellence may then indeed succeed, and that its success is bolstered by greater military capabilities. However, I conclude that greater capabilities provide coercive leverage by increasing the risk of disaster, with first-strike capabilities being especially destabilizing. TNR was correct to highlight the dangers of pursuing nuclear superiority.

The paper proceeds as follows. Section 2 reviews these four recent critiques of TNR. Section 3 presents the game-theoretic argument. Section

---


2See also Bélanger and Bowen 2020.
concludes. Proof of the formal results and discussion of coding rules are contained in the Appendices.

2 All But Compellence and Deterrence:

The ABCD of TNR’s Critiques

According to TNR, nuclear weapons revolutionize international relations (Brodie 1946; Bundy 1988; Glaser 1990; Jervis 1989; Powell 1990; Schelling 1966; Waltz 1981; 1990). They are the ultimate instruments of deterrence, the weapon of the weak. In the pre-nuclear age, conventional superiority could produce military victory. In the nuclear age, the prospect of a nuclear attack acts as a powerful deterrent. According to Jervis (1989, Chapter 1), TNR generates the following predictions: 1. peace between great powers, 2. the preservation of the status quo, 3. the infrequency of crises, 4. states’ frequent attempts during crises to highlight the risk of nuclear disaster, so as to demonstrate their resolve and gain a competitive advantage, 5. a tenuous link between the military balance and political outcomes. A critique of TNR could question whether its implications are correct and whether it outlines the right mechanism generating the predictions.

Most recent critiques accept that nuclear weapons have encouraged peace between great powers. Lieber and Press (2020) concede: “To be clear, nuclear weapons have had a huge impact on international relations by helping
to prevent great power war.” Gavin (2020, 194, 198) admits that TNR’s “most important insight [...] is correct: few if any political political objectives are worth the extraordinary costs of a thermonuclear war.” The theory offers “a powerful lens to understand [...] the disappearance of great-power war.” Kroenig (2018, 29) argues that TNR “is correct that any rational leader should be incredibly unlikely to intentionally initiate a major war against a state with a secure, second-strike capability.” Green (2020, 1-2) is more skeptical, but he only goes so far as saying that “the prediction of peace” is “difficult to evaluate” because there could be alternative explanations.

Recent critiques also do little to challenge TNR’s prediction that nuclear weapons are the ultimate tools of deterrence and help enforce the status quo. If two nuclear weapons confront one another, one issuing a compellent threat, hoping to extract concessions, and the other issuing a deterrent threat, trying to enforce the status quo, then the deterrer should have the advantage, according to TNR. Compellence is costly. Countries have already accepted the status quo (Schelling 1966, 70-72, 82, Jervis 1989, 30-31). Yet TNR qualified its claim. If a state pursued compellence, despite its associated costs, couldn’t it signal greater resolve and therefore prevail? (Schelling 1966, 101-103; Jervis 1979, 297-298; Jervis 1989, 34). Could we always identify the “status quo”? Since it may be ambiguous, all sides may claim to defend it and expect to prevail (Jervis 1989, 32)3.

3On the implication of incompatible narratives for mutual optimism and war, see Debs (2020).
Despite TNR’s own misgivings, its critics do not explain how nuclear compellence could overcome deterrence. Lieber and Press (2020, 2, 14) concede: “Nuclear weapons are the most effective instruments of deterrence.” They “are uniquely deterring because they appear to make victory in war impossible. They are the ultimate tools of stalemate.” Kroenig (2018, 114), after presenting a theory of nuclear coercion, admits: “Compellence may be more difficult than deterrence, as Schelling and others have maintained.” Green and Gavin embrace TNR’s own reservations, but do not explain how compellence could prevail (Green 2020, 21-26; Gavin 2020, 63-64). Gavin ultimately accepts that deterrence is easier than compellence, an “important and obvious” insight from Sechser and Fuhrmann (2013), which “should not be surprising.” (Gavin 2020, 73).

It is not obvious that deterrence is easier than compellence. In the Cuban Missile Crisis, Schelling (1966, 82) concedes, a deterrent failed before a compellent threat succeeded. In September 1962, Kennedy declared that the “gravest issues would arise” if the Soviets deployed any “significant offensive capability” in Cuba (U.S. Department of State 1962). Khrushchev responded by expanding his deployment (Fursenko and Naftali 1997, 204-213). Kennedy then announced a blockade, a classic example of a probabilistic threat, and he succeeding in compelling the Soviets to withdraw from Cuba. Moreover, Sechser and Fuhrmann (2013) do not show that deterrence is easier than compellence. Their dataset contains no deterrent threat.

4See also Sechser and Fuhrmann (2017). Sechser (2011, 395-396) suggests
It is important to establish whether and how compellent threats could prevail. If deterrence is stable, then TNR’s main implications, on peace between great powers and the preservation of the status quo, will endure. We do not have a clear understanding of this question. Canonical game-theoretic work does not analyze the difference between deterrent and compellent threats (Fearon 1994; Powell 1990). Once states exchange these threats, they may worry about rapid escalation. Recent work has made progress in understanding crisis preemption games, but debates remain. Schelling (1960, 207) had wondered if the reciprocal fear of surprise attack could lead to war, even with “no “fundamental” basis for an attack by either side.” Powell (1990, Chapter 5) concluded that peace would prevail, if any country could end the crisis. Baliga and Sjöström (2004) showed instead that conflict would be inevitable, under some conditions, if countries could strictly prefer to attack no matter what their enemy chooses (see also Baliga and Sjöström 2020, N.d.). It is not clear, though, that there is then “no “fundamental” basis for an attack by either side,” or that this set-up captures the strategic outlook of states facing a nuclear disaster. Revisiting this question, Fearon (N.d.) adds that if countries receive a warning, even imperfect, of their enemy’s action, then peace prevails when first-strike advantages are small. None of these models allows that compellence may be slightly less effective than deterrence, as documented by other datasets. Yet he admits that the difference is small (41.4% vs. 50-59%) and the “conclusion is only tentative,” given concerns over selection effects and confounding factors.
for strategic selection, crisis bargaining, and the exchange of compellent and deterrent threats. I present such a model below.

For now, I observe that recent critiques of TNR are quite narrow. They concede TNR’s most important prediction - peace between great powers - and the main mechanism generating this prediction - nuclear weapons are the ultimate tools of deterrence. Their critique really centers on Jervis (1989, Chapter 1)’s fifth implication - that there is a link between the military balance and political outcomes. I evaluate each argument in turn.

### 2.1 Lieber and Press on a MAD Competition

Lieber and Press (2020, 2) argue that “the continuation of great power competition under the shadow of nuclear weapons” is “the central anomaly of the nuclear age.” TNR fails to account for the following features of the Cold War: the superpowers engaged in arms races and cared about “relative gains,” strategic territory, and alliances (Lieber and Press 2020, 16). TNR is wrong to attribute U.S. drive for nuclear superiority to “irrational” behavior, based on leaders’ misperceptions or the strength of military organizations (Lieber and Press 2020, 4-5). There is a strategic logic justifying such an effort, due the challenges of “creating strategic stalemate, maintaining stalemate, and practicing deterrence under stalemate” (Lieber and Press 2020, 5). Washington at various times built war plans for a disarming first strike and eventually acquired such capabilities (Lieber and Press 2020, Chapters 2
and 3). Aggressive nuclear postures, such as those built on “flexible, limited nuclear options,” can advance a state’s interests even under conditions of MAD, and they are more likely to adopted by conventionally weaker states (Lieber and Press 2020, Chapter 4, esp. p. 96).

This is a thought-provoking critique, but it focuses mainly on an extreme version of TNR (Waltz 1990). The fact that nuclear weapons are the ultimate tools of deterrence does not mean that strategic territory and alliances are worthless. Great powers continue to have interests even if they possess nuclear weapons. They wish to advance these interests if they can avoid a nuclear confrontation, and if they are engaged in a nuclear crisis, they may use aggressive tactics to bolster the credibility of their threats (Schelling 1966 and Jervis 1989’s fourth implication). The fact that nuclear weapons are weapons of the weak suggests that weaker states would rely on nuclear threats more aggressively, fearing that they may need to use them or lose them (e.g., Powell 2015). Flexible, limited options are compatible with TNR if they hope to coerce by raising the risk of nuclear disaster (Powell 2015). They were rejected by TNR scholars because they were presented as a way to control escalation and avoid nuclear disaster (Jervis 1979-1980).

Lieber and Press (2020)’s most powerful critique argues for a link between the military balance and political outcomes. But how do we know that such a link exists, that policymakers held such beliefs, and that this belief motivated their drive for nuclear superiority? Lieber and Press (2020) do not offer any evidence of such a link, and they offer very little evidence of policymakers’
strategic outlook. Their technological analyses may show that the United States was close to acquiring a first-strike capability, but this does not prove that U.S. policymakers intended to build such a capability, or that greater military capabilities improved their coercive leverage.

TNR actually provides valuable insights for understanding the U.S. drive for nuclear superiority. The basic problem, Jervis (1989, 186) explains, is that nuclear policy often eludes clear criteria for success: “In many cases it is hard to find objective indicators of whether a policy has succeeded or failed. Thus there is a great deal of room for states to influence the interpretation of the outcomes, and that may be more important than the outcomes themselves.” In particular, nuclear superiority “matters because others […] think it matters” (Jervis 1989, 196). As a result, the United States, Jervis concluded, had “become trapped in a world largely of its own making” (Jervis 1989, 212). The U.S. drive to superiority and “credible threats” was chosen to impress the Soviets, “especially,” but also “to reassure the Europeans” and even reassure “itself” (Jervis 1989, 196, 212). Trapped by previous discourse on the essential features of competent nuclear management, policymakers may choose policies that appear sensible, but are in fact inefficient, even dangerous.° As Jervis (1989) did not spell out how this insight captures the evolution of U.S. nuclear strategy, it may easily be discounted, but it is actually powerful.

---

°On the pitfalls of transparency, see, e.g., Canes-Wrone, Herron and Shotts (2001); Debs and Weiss (2016).
With the New Look, the Eisenhower administration argued that nuclear weapons would serve as an affordable and effective deterrent against the conventionally superior Soviet army. Soon after, the French invoked this logic to justify their own force de frappe. The problem for Washington is that it risked entrapment by a small, autonomous nuclear arsenal, as Moscow developed its ICBM capabilities.

Flexible response could reduce the risk of entrapment, pressing for the centralization of nuclear decisions, while shifting some of the burden for conventional defense onto U.S. allies. But how could the United States convince its allies to spend more on conventional forces, after arguing that Soviet forces were overwhelmingly superior? Washington first had to debunk the claim of Soviet conventional superiority (Enthoven and Smith 2005, Chapter 4). Then it argued that at least it possessed nuclear superiority, as Deputy Secretary of Defense Roswell L. Gilpatric did in October 1961 (see, Bundy 1988 381-382, 418-419; Cameron 2017 96). Kennedy had campaigned on the dangers of a missile gap, believing at the time that the Soviets were ahead. After learning that the reverse was true, the administration could now boast of its superiority.

Yet discussions of nuclear superiority could again trap U.S. officials. It would be expensive, perhaps impossible to maintain. In January 1962, Kennedy’s top national security advisers supported an arms buildup because “[s]entiment for more missiles and more nuclear weapons” in Congress was “pretty strong.” At the same time, they encouraged him to question whether
“this sentiment could be rationally defended” and state that he did “not subscribe to the doctrine of long-term ‘nuclear superiority’” and believed that “in the long run” Washington was “headed for a nuclear stalemate” (quoted in Cameron 2017, 28). U.S. officials continued to speak of nuclear superiority, but downplayed its importance and devised a criterion for effective deterrence - assured destruction - which did not rely on the nuclear balance. “In terms of numbers of separately targetable, survivable, accurate, reliable warheads, U.S. strategic forces have remained consistently superior to those of the Soviet Union. However, the relationship of this “superiority” to U.S. military and political objectives is unclear. [...] Once each side has enough nuclear forces virtually to eliminate the other’s urban society in a second strike, the utility of extra nuclear forces is dubious at best. In this context, notions of nuclear “superiority” are devoid of significant meaning” (Enthoven and Smith 2005, 183).

By the time Nixon took office, the end of U.S. nuclear superiority appeared imminent. In his first press conference in January 1969, Nixon promised instead “sufficient military power” for the protection of U.S. interests and commitments (Nixon 1969). This objective was vague, seemingly consistent with assured destruction while allowing future flexibility. NSC staffer Morton Halperin admitted: “we will call whatever option we choose sufficiency.” In the short run, though, Washington’s policies would be measured against previous standards. According to Kissinger, European publics “would be amazed if they heard that we were not vastly superior to the Soviet Union”
Ultimately, the SALT accords enshrined Soviet quantitative superiority. Promoting them in 1972, Nixon doubled down on “sufficiency,” suggesting that U.S. forces were “without question sufficient for the maintenance of our security and the protection of our vital interests” ... and also that they were superior to Soviet forces: “No power on earth is stronger than the United States of America today. And none will be stronger than the United States of America in the future” (Nixon 1972). This had to refer to qualitative superiority, but how could Washington establish such superiority? Certainly, it had innovated with MIRV technology, but what would happen once the Soviets developed their own? In August 1973, Chairman of the Joint Chiefs of Staff Admiral Moorer assured Kissinger that the United States still had “technical superiority.” Exasperated, Kissinger retorted: “I keep hearing about our technical superiority. We haven’t developed a new missile for 15 years and they have three new ones. Where is our technical superiority?” Moving away from superiority could undermine U.S. assurances. As State Department official Seymour Weiss admitted: “We told them [our allies] we were qualitatively superior. We can’t now say that that doesn’t make any difference” (quoted in Green and Long 2016 42).

The elusive evaluation of nuclear strategy also complicated U.S. relations with the Soviet Union. The Johnson administration had called for arms control agreements as early as January 1964, but Washington’s insistence on nuclear superiority undermined its argument that Moscow should be content
with assured destruction (Cameron 2017, 85-94). After SALT I, the Soviet arsenal was the largest in the world. Brezhnev could now placate Washington’s preference for “sufficiency,” admitting to Kissinger in March 1974: “Let me just say, I would be willing, before your very eyes, to destroy 100 launching sites. Would that change anything? Nothing. President Nixon was right in saying in our first meeting that the Soviet Union could destroy the United States seven times over, and the United States could destroy the Soviet Union seven times over” (U.S. Department of State 1974, 730). In this context, Washington’s subsequent move to counterforce didn’t square with a commitment to sufficiency. In early 1983, Soviet General Secretary Yuri Andropov warned Warsaw Pact leaders that “the new round of the arms race, imposed by the USA, has major, qualitative differences ... in creating modified missile systems, [the Americans] do not hide the that they are really intended for a future war. From here spring the doctrines of ‘rational’ and ‘limited’ nuclear war” (quoted in Green and Long 2017, 618).

In sum, TNR does offer a compelling logic for understanding the U.S. drive for nuclear superiority. Rational policymakers, eager to convince audiences of their competent nuclear management, may be trapped by previous discourse and changing circumstances, even if nuclear superiority does not necessarily advance U.S. security. Whether it does is of course the crux of the matter. This question remains unanswered.
2.2 Green on MAD Fragility

Green (2020, 247) argues that the nuclear revolution failed, “at least as applied to peacetime nuclear competition,” because it did not convince U.S. policymakers in the late Cold War “to adopt its preferred stabilizing policies.”

According to the standard narrative, the logic of TNR dominated U.S. policymaking until the turn to counterforce in the mid-1970s. President Nixon had adopted the logic of TNR in pursuing arms control agreements with the Soviet Union. In June 1972, he praised the SALT accords for “check[ing] the wasteful and dangerous spiral of nuclear arms,” based on “the recognition that two nuclear nations, each of which has the power to destroy humanity, have no alternative but to coexist peacefully because in a nuclear war there would be no winners - only losers” (Nixon 1972). Kissinger explained that “to the extent that balance of power means constant jockeying for marginal advantages over an opponent, it no longer applies [...] now both we and the Soviet Union have begun to find that each increment of power does not necessarily represent an increment of usable political strength” (U.S. Department of State 1972, 402).

This account is wrong, Green argues. Throughout the 1969-1979 period, U.S. policymakers did not believe in TNR, and competed aggressively with the Soviet Union in their nuclear acquisition, employment, and arms control policy.

According to Green, a state may desire to engage in arms races because of
its many benefits: diverting enemy resources, bolstering deterrence and coercion, and strengthening alliances (Green 2020, 50-51). Whether and how it does so depends on its “comparative constitutional fitness” on competitive and cooperative dimensions. The competitive dimensions are resource production, which favors intense competition; resource extraction, which favors high quantitative targets; resource direction, which permits technological innovation. The cooperative dimensions are efficiency at distributional bargaining, which bolsters diplomatic negotiations; enforcement, which helps ensure the enemy’s compliance. When a state has an advantage in all competitive tasks, it does not pursue an arms control agreement. When it has an advantage in all cooperative tasks, it pursues an arms control agreement favoring qualitative stability. When it has an advantage in some or no cooperative tasks, it pursues a competitive arms control agreement, channeling competition to areas of strength (Green 2020, 57).

This general theory explains U.S. nuclear policy between 1969 and 1979, Green (2020) contends, when it agreed to quantitative limitations while pursuing qualitative advantages. Washington had an advantage in some competitive tasks and a disadvantage in both cooperative tasks (Green 2020, 59). Its great economic potential, innovative industries, and productive civil-military relations supported resource production and direction. Its political institutions and ideology undermined distributional bargaining and enforcement. Electoral swings and congressional oversight complicated planning. Transparent institutions prevented cheating (Green 2020, 108-9). Liberal ideology
bound the country to international agreements, irrespective of its terms or
the enemy’s compliance.

Green’s theory is complex and the evidence is very rich, analyzing well
over 20,000 pages of declassified documents. This close engagement with
the qualitative evidence generates two important contributions.

First, Green (2020) illustrates in stark terms how U.S. policymakers’ pri-

tive deliberations clashed with their public discourse (see also Cameron 2017;
Maurer 2019; Petrelli and Pulcini 2018). TNR, as I argue above, can explain
how policymakers may feel trapped by previous discourse on competent nu-
clear management. Still, it is helpful to document how they justified their
policy privately.

Second, Green (2020) highlights the fluidity of policymakers’ assessment
of the nuclear balance. Even with large nuclear arsenals, states may worry
about losing their retaliatory capabilities. Between 1969 and 1974, Wash-
ington went from confidence that neither side could obtain a first strike, to
concern that Moscow may have such a capability, to confidence that it was
alone in possessing it.

---

6 For a critique, see, e.g., Sagan (2021).
7 Such views are quite complex, and don’t amount to a wholesale rejection
of the nuclear stalemate. As Green (2020) admits, “the American docu-
mentary record is replete with examples of leaders bluntly confirming the
existence of nuclear stalemate” (Green 2020, 44).
Indeed, in February 1969, Packard asserted that “[n]either side can get a first strike capability.” Still, he worried about future developments: “Uncertainty is if they have new idea, we don’t, ball game could change” (U.S. Department of State 1969a, 2, 4). The following year, Washington worried that Minuteman missiles were vulnerable to Soviet ICBMs, given their improved accuracy and possible MIRV capability, and bombers were vulnerable to Soviet SLBMs, if these adopted a depressed trajectory (Green 2020, 92-93). Packard thus explained in March 1970, “we must then rely increasingly on the Polaris. We have 20 Polaris subs on station with 16 missiles per sub. This means 320 warheads. If we assume 80% reliability, then we have 256 which are reliable. If they have ABM then that would be bad news for us” (quoted in Green 2020 94-95). Yet by 1974, Washington exuded confidence. Kissinger declared in August: “We are the only ones who would gain in a first strike because most of their force is land-based,” and hence highly vulnerable (quoted in Green 2020 136). NSC Staff member Phil Odeen added a month later that “a [Soviet] decision to strike first, based largely on vulnerability calculations, is simply unreal” (quoted in Green 2020 134).

Despite these important contributions, the book fares better in explaining how a nuclear power would compete than whether and why it does. To begin with, the key variables of his theory are, by his own admission, “slow changing” (Green 2020 59). If they correctly predict the arms control agreements of 1969-1979, they fail to predict the absence of such agreements in 1959-1969, when the U.S. arsenal more than doubled. Moreover,
Green’s theory does not spell out how state preferences aggregate into arms control agreements. If a state has an advantage in cooperative tasks, the other does not. The first may want an ambitious agreement; the other would not. A state may also need to trade restrictions in areas of strength to obtain restrictions in areas of weakness. For example, U.S. policymakers had the advantage in ABM technology and cruise missiles. They traded concessions on ABM technology for limits on quantitative competition, but they pressed - unsuccessfully - for an agreement condoning competition on cruise missiles. Green’s theory cannot explain these different outcomes.

TNR, in fact, offers a more compelling explanation of the U.S. drive for nuclear superiority. Green (2020, 20), like Lieber and Press (2020), downplays TNR’s explanation as focused on “irrational” or “parochial” interests. Yet TNR explains how rational policymakers would pursue superiority, even if it does not in fact serve the U.S. national security interest. Whether superiority does serve U.S. national security is the crux of the matter, which Green (2020)’s evidence does not elucidate.

Put differently, scholars already knew that Washington rejected TNR. It is certainly helpful to document that U.S. policymakers turned against it sooner. However, a “better test” of TNR, Sagan (2021, 138) concludes, “would focus not on whether U.S. government leaders believed in MAD, but rather on whether or not the U.S. pursuit of counterforce capabilities during the Cold War increased the risk of accidental war by creating dangers of mistaken preemption, or what Thomas Schelling famously called “the reciprocal
fear of surprise attack” and “the dynamics of mutual alarm.”” I provide such an analysis below.

2.3 Gavin on the Coherence of U.S. Grand Strategy

Gavin (2020, 194), for his part, argues that TNR “has failed to predict and explain critical aspects of U.S. nuclear policies, including nuclear strategy and nonproliferation.” The nuclear forces built by Washington “made little sense if the United States had fully embraced the consequences of mutual vulnerability” (Gavin 2020, 200). If nuclear weapons are weapons of peace, then “more may be better,” as Waltz (1981) famously argued, and Washington should have been sanguine about proliferation (Gavin 2020, 195-196). To better explain U.S. nuclear policies, Gavin (2020, 127) proposes to move “beyond deterrence” in a new theory, Multiple and Interactive Deterrence, Assurance, and Reassurance (MIDAR). Nuclear primacy was just one of Washington’s “strategies of inhibition,” intended to prevent proliferation among allies and neutrals, a goal which “has rarely been understood as a core, long-standing, and driving goal of U.S. grand strategy” (Gavin 2020, 76).

Gavin’s critique, like Lieber and Press (2020)’s, is mainly based on an extreme version of TNR (Waltz 1990). As we argued above, TNR has a compelling explanation for the U.S. drive for nuclear superiority, in which, it should be noted, concerns for reassuring allies play a central role.8 The fact that nuclear weapons are the ultimate tools of deterrence does not mean

\[\text{Gavin (2020, 118)’s historiography is also confusing.} \]

\[\text{Gavin (2020, 118)}\]
that the United States should view nuclear proliferation favorably. If nuclear weapons are the weapons of the weak, then the weak would really want to acquire nuclear weapons, but other states would be very concerned about their nuclear acquisition (Debs and Monteiro 2017; Monteiro and Debs 2014). The idea that Washington has opposed proliferation among allies and neutrals has been widely known, even before Gavin proposed the expression “strategies of inhibition” to describe U.S. nonproliferation efforts (see, e.g., Trachtenberg 1999, 251-256, 305; Solingen 2007; Maddock 2010; Monteiro and Debs 2014). Suggesting that nonproliferation has always been a first priority would be incorrect. Washington actually varied in its commitment to nonproliferation, as Gavin acknowledges (Gavin 2020, 96). The key challenge is to understand the extent of U.S. concerns, and the effectiveness of different nonproliferation tools, questions that the existing literature already addresses (see, e.g., Gheorghe 2019; Lanoszka 2019; Monteiro and Debs 2014; Whitlark 2017).

Gavin (2020)’s broad generalizations obscure the role that history would play in informing debates on nuclear politics. In his earlier work, Gavin criticizes that the impact of flexible response has been “overdrawn.” Yet that impact has long been recognized as limited, among scholars and policymakers alike, including McNamara himself (Freedman 1981, 243-246; McNamara 1983, 65; Gaddis 1982, 219-220; Trachtenberg 1999, 287.)

9 On the possible effects of nuclear proliferation, see, e.g., Anderson, Debs and Monteiro (2019); Bell (2021).
icized scholars for failing to engage the “treasure trove” of available archival evidence (Kroenig 2013; Sagan and Waltz 2012; Sechser and Fuhrmann 2013; Gavin 2012, 576-577; Gavin 2014, 29-31). In response, the authors encouraged Gavin to explain how to extract general lessons from history (Sagan and Waltz 2013, 147; Sagan 2014, 7; Fuhrmann, Kroenig and Sechser 2014, 40-42). Unfortunately, Gavin refrains from doing so, instead copying his prior reviews, published some six and eight years earlier.

In his own substantive claims, Gavin (2020) does not explain how he has engaged with the “treasure trove” of archival evidence, and what standard of evidence he set for reaching his conclusions. When presenting MIDAR theory, Gavin does not discuss any primary evidence, calling the theory “a heuristic framework” which “does not fully capture the nuance and context of the history of U.S. nuclear statecraft. It is easy to think of many U.S. behaviors or policies that are not explained by or that even contradict this analysis” (Gavin 2020, 143). When arguing that NATO’s goal was less to deter the Soviet Union than to prevent proliferation, a goal which has been “unstated and largely unrecognized” (Gavin 2020, 112), Gavin discusses only one primary document in any meaningful way, MC-48, a 1954 report by the Military Committee to NATO arguing that nuclear superiority is essential in deterring a Soviet attack (Gavin 2020, 114-115). Gavin (2020) again apologizes for the thin evidentiary basis for his claim: “this narrative [...] is more speculative than we would like. We do not have the quality or quantity of documents for the 1970s and 1980s that we do for earlier periods” (Gavin
2020, 122). For these earlier periods, Gavin borrows a narrative “laid out by several scholars,” which he believes is “worth repeating” (Gavin 2020, 112) (see, e.g., Trachtenberg 1999, 305).

In sum, TNR and Gavin actually agree on the importance of allied considerations in the U.S. drive for nuclear superiority. Whether and how military capabilities truly coerce enemies, however, remains an open question.

2.4 Kroenig and the Ill Logic of Nuclear Superiority

Kroenig (2018) represents the most direct attack on the irrelevance of the military balance for political outcomes. In his view, U.S. officials believed in nuclear superiority, which did offer coercive benefits.

During key crises, he claims, U.S. officials “explicitly linked the nuclear balance of power to policy recommendations about crisis escalation” (Kroenig 2018, 88). In October 1962, General Maxwell Taylor, Chairman of the Joint Chiefs of Staff, wrote to McNamara: “We have the strategic advantage in our general war capabilities... This is no time to run scared.” McNamara himself, while advocating for a “restrained approach [...] took into account the nuclear balance of power, arguing “if we had to fight a war with the Soviet Union, we’d have fewer casualties today than if we had to do it later.”” Ultimately, Secretary of the Treasury C. Douglas Dillon attributed U.S. success to “[o]ur nuclear preponderance [...] That’s what made the Russians back off” (quoted in Kroenig 2018, 88). U.S. officials were correct. Simulations show that U.S.
casualties increase with the size of the enemy’s nuclear arsenal and decreases with the size of its own arsenal (see https://nuclearsecrecy.com/nukemap/). Statistical tests show that larger arsenals are associated with more favorable crisis outcomes. They provide bargaining leverage, according to Kroenig’s “superiority-brinkmanship” theory. For Kroenig, “the logic of the argument is simple: in a game of chicken we might expect the smaller car to swerve first even if a crash would be disastrous for both” (Kroenig 2018, 4).

Kroenig (2018)’s quantitative work is characteristically innovative. But the fact that larger arsenals inflict greater damage is not surprising. What remains to be shown is whether the resulting differences are politically and diplomatically meaningful.

For example, Kroenig (2018, 49) concludes that if the U.S. arsenal underwent deep cuts, casualties from a Russian first strike would increase from 70 to 125 million. With fewer U.S. nuclear weapons to destroy, Russia could hurl more missiles at U.S. cities. Does that mean that U.S. officials would accept a death toll of 70 million, and run the risk of nuclear war? If we calculated Russian casualties, and they turned out to be higher, could Washington coerce Moscow to back down because of its “low” death toll of 70 million?

In a conflict with North Korea, Kroenig (2018, 48) argues, the continental United States would be spared a nuclear disaster. Would Washington initiate war against Pyongyang? It has not, even when Pyongyang was nonnuclear. In his 1999 report to Congress, former Secretary of Defense William J. Perry
declared that “deterrence of war on the Korean peninsula is stable on both sides [...] It is likely that hundreds of thousands of persons - U.S., ROK, and DPRK - military and civilian - would perish, and millions of refugees would be created. ... [T]he prospect of such a destructive war is a powerful deterrent to precipitous U.S. or allied action” (Perry October 12, 1999, 3). Even Republican officials balked at the prospects of war. Condoleezza Rice, Secretary of State during the DPRK’s first nuclear test in July 2006, acknowledged in her memoirs: “The military option against Pyongyang was not a good one; it was fraught with unintended consequences and the near-certainty of significant damage to Seoul” (Rice 2011, 712). And while Trump boasted about the size of his nuclear button, his adviser Steve Bannon confided: “There’s no military solution [to North Korea’s nuclear threats], forget it. Until somebody solves the part of the equation that shows me that 10 million people in Seoul don’t die in the first 30 minutes from conventional weapons, I don’t know what you’re talking about, there’s no military solution here, they got us” (Kuttner 2017).

Assessing his own work, Kroenig highlights not its quantitative evidence but its qualitative evidence and theoretical argument. The statements from the Cuban Missile Crisis “are as close to “smoking gun” evidence as one can reasonably expect to find in qualitative, social science research” (Kroenig 2018, 88). Researching the book, Kroenig “could not call to mind a single, clear explanation for why a strategic nuclear advantage might translate into

---

10See also Debs and Monteiro (2018).
a geopolitical advantage” (Kroenig 2018, ix). The theory was the “biggest contribution” of a 2013 article, a precursor to the book (Kroenig 2014, 63).

The correlation between the nuclear balance and crisis outcomes should indeed be taken with caution, given the small number of crises, and the difficulty of establishing causation. But it is difficult to read much from the qualitative evidence. We already know that U.S. policymakers seemed to believe in nuclear superiority in their turn to counterforce in the mid-1970s, and Green (2020)’s evidence suggests that the transition may have begun sooner. Qualitative evidence that U.S. policymakers believed in nuclear superiority in actual crises, and that such beliefs explained their policy choices, would be valuable. Is this an accurate description of the Cuban Missile Crisis? Unfortunately, Kroenig misrepresents the views of key officials, does not engage with evidence contradicting his view, and does not provide any new evidence to adjudicate between competing interpretations.

Key U.S. officials, including McNamara and President Kennedy himself, did not attribute their success in the crisis to their nuclear superiority. Instead, they highlighted their local conventional superiority and the simple risk of nuclear disaster (see, e.g., Blight and Welch 1990). McNamara did not “take into account” the nuclear balance of power and “explicitly link it” to policy recommendations about crisis escalation. He recommended restraint, based solely on the risk of nuclear disaster. His statement, quoted

\footnote{For critical reviews, see, e.g., Glaser (2019); Sechser (2019).}
above, describes the inescapable mathematical fact that casualties increase with the number of nuclear weapons detonated. Yet McNamara did not believe that superiority should instruct policy choices. Dillon himself recognized as much, in the very statement quoted by Kroenig. “I agree totally with Nitze,” Dillon stated, referring to Paul Nitze, Assistant Secretary of Defense for International Security Affairs, “and think the McNamara thesis that our nuclear superiority made little or no difference is dead wrong. Our nuclear preponderance was essential. That’s what made the Russians back off, plus the fact of our total conventional superiority in the region” (Blight and Welch 1990, 152-153). McNamara indeed declared that “nuclear weapons serve no military purpose whatsoever. They are totally useless, except to deter one’s opponent from using them. This is my view today. It was my view in the early 1960s” (McNamara 1983, 79, emphasis in the original).

The Secretary of Defense, along with other key decisionmakers, analyzed the crisis as such: “American nuclear superiority was not in our view a critical factor, for the fundamental and controlling reason that nuclear war, already in 1962, would have been an unexampled catastrophe [...] No one of us ever reviewed the nuclear balance for comfort in those hard weeks. The Cuban missile crisis illustrates not the significance but the insignificance of nuclear superiority in the face of survivable thermonuclear retaliatory forces. It also shows the crucial role of rapidly available conventional strength” (Rusk et al. 1982, 85, emphasis in the original).

Which group was most influential in shaping President Kennedy’s out-
look, and how did he approach the crisis? Dillon himself acknowledged McNamara’s “great influence on the President” (Blight and Welch 1990, 154). McNamara declared that Kennedy “accepted my recommendation” to “never initiate, under any circumstances, the use of nuclear weapons” (McNamara 1983, 79). Prior to the crisis, Kennedy had expressed the view that “there will not be winners in the next nuclear war, if there is one, and this country and other countries will suffer very heavy blows” (Kennedy 1962). Early in the crisis, hawks pressed for an attack on Cuba, convinced that U.S. nuclear superiority would deter Soviet escalation. Kennedy refused, telling the congressional leadership on October 22: “If we go into Cuba we have to all realize that we are taking a chance that these missiles, which are ready to fire, won’t be fired. So that’s a gamble we should take. In any case we are preparing to take it. I think that is one hell of a gamble” (May and Zelikow 2001, 176). Looking back at the crisis in his 1963 State of the Union address, Kennedy declared that “a line of destroyers in a quarantine, or a division of well-equipped men on a border, may be more useful to our real security than the multiplication of awesome weapons beyond all rational need” (quoted in Gaddis 1982, 216). Ultimately, President Kennedy’s handling of the crisis appears more consistent with TNR than with superiority-brinkmanship theory.

Turning to the theory, it is surprising that Kroenig “could not call to mind a single, clear explanation for why a strategic nuclear advantage might translate into a geopolitical advantage.” There is a long tradition making this
very argument (see, e.g. Kahn 1960, 291-292; Kahn 1965, 9-13; Gray 1976, 79; Gray and Payne 1980, 19; Nitze 1984, 160). If Kroenig takes exception with the characterization of his argument as “simply standard bargaining theory” (Glaser 2019, 6; Kroenig 2019, 25-26), earlier proponents embraced it. For example, Nitze argued that it was “a copybook principle in strategy that, in actual war, advantage tends to go to the side in a better position to raise the stakes by expanding the scope, duration or destructive intensity of the conflict. By the same token, at junctures of high contention short of war, the side better able to cope with the potential consequences of raising the stakes has the advantage. The other side is the one under greater pressure to scramble for a peaceful way out” (Nitze 1984, 160).

Kroenig’s argument does go beyond these earlier works, by presenting a sophisticated model, but the model was developed by Powell (1990, Chapter 3). Kroenig’s contribution is to reinterpret the variables of Powell’s model, assume that military capabilities factor into a country’s resolve, so that they affect diplomatic outcomes.

Kroenig’s argument is simple but confusing. In the game of Chicken, the stronger power does not necessarily prevail. Indeed, consider the representation of the game below, where countries I and II choose to ‘Stand Firm’ or ‘Submit’ (Powell 1990, 35). If both stand firm, disaster ensues, country i receives payoff \( d_i \). If country i stands firm and country j submits, country i wins, country j loses. They get payoffs \( w_i, l_j \), respectively. If both submit, they get a compromise payoff \( c_i \), where \( w_i > c_i > s_i > d_i \) \( \forall i \in \{I, II\} \).
There are two pure strategy Nash equilibria, where either power prevails (\((\text{Stand Firm, Submit})\) and \((\text{Submit, Stand Firm})\)), and a mixed-strategy equilibrium, which produces disaster with positive probability.

---Figure 1 about here---

The game’s multiple equilibria explain its complicated legacy in nuclear politics. It was first used by Bertrand Russell (Russell 1959), who argued not for superiority but for disarmament, worried that nuclear crises would lead to disaster (Russell 1959, 30). Herman Kahn objected, outlining a standard plea for superiority: “If one is unwilling to risk global war, while the other side is willing to risk it, the side which is willing to run the risk will be victorious in all negotiations and will ultimately reduce the other side to complete impotence” (Kahn 1960, 291-292; see also Kahn 1965, 9-13).

To argue that the stronger party should prevail, Kroenig could appeal to equilibrium selection criteria, say risk dominance (Harsanyi and Selten 1988). A pure-strategy equilibrium is risk dominant if players are willing to play the posited strategies for a larger set of beliefs about their opponent’s play or, equivalently, if the product of deviation losses at that equilibrium is larger (Harsanyi and Selten 1988, 87). The equilibrium (Stand Firm, Submit) risk dominates (Submit, Stand Firm) if and only if

\[
\frac{s_I - d_I}{w_I - c_I} < \frac{s_{II} - d_{II}}{w_{II} - c_{II}}
\]

(1)

\[12\] Another criterion, payoff dominance, is undiscriminating here.
Assume, following Kroenig (2018), that as the nuclear balance becomes more favorable to country I, country I’s disaster payoff $d_I$ increases and country II’s disaster payoff $d_{II}$ decreases. Then, condition (1) is easier to satisfy. A more favorable nuclear balance reduces a country’s concerns for standing firm. It is more likely to stand firm in the risk-dominant equilibrium.

This result could justify Kroenig’s conclusion. Yet this model misses an important feature of nuclear crises, highlighted by Kroenig’s own superiority-brinkmanship theory: disaster may strike inadvertently. This risk of inadvertent war, it turns out, may allow the weaker power to prevail in the risk dominant equilibrium.

Assume that disaster strikes with probability $f > 0$ if only one country stands firm. The game otherwise remains the same.

Figure 2 about here

(Stand Firm, Submit) risk dominates (Submit, Stand Firm) if and only if

$$\frac{s_I - d_I}{w_I - c_I - f(w_I - s_I)} < \frac{s_{II} - d_{II}}{w_{II} - c_{II} - f(w_{II} - s_{II})}$$ (2)

As before, this condition is satisfied more easily as the nuclear balance becomes more favorable to country I. However, the weaker country may now prevail in the risk dominant equilibrium, even if it would not absent any inadvertent war.

Let $fd_i + (1 - f)w_i > c_i$, so that this remains a game of Chicken.

---

13Let $fd_i + (1 - f)w_i > c_i$, so that this remains a game of Chicken.
exogenous risk of disaster.\footnote{Let I be stronger than II, with $w_i = 1$, $s_i = 0$, $d_I = -1$, $d_{II} = -4$, $c_I = 3/4$, $c_{II} = 1/4$, $f = 1/10$.}

Here, the weaker nuclear power does fear the harsh consequences of a nuclear exchange. Yet if disaster may follow compliance, the weaker power may rely aggressively on the threat of nuclear weapons, fearing that it needs to “use them or lose them.” Put differently, nuclear threats are effective only if assurances are credible \citep[74-75]{Schelling1966}. Such is a key feature of brinkmanship crises. It escapes Kroenig’s conceptualization.

In short, while \citet{Kroenig2018} is thought provoking, it misses important nuances of nuclear dynamics and does little to explain whether and how nuclear superiority promotes national security.

\section{2.5 TNR is Dead! Long Live TNR!}

In sum, critics argue that TNR should be abandoned, yet they consent that nuclear weapons have been a force for peace, the ultimate tools of deterrence. They do not establish how compellence could overcome deterrence, or how military capabilities correlate with diplomatic success.

Table 1 and 2 takes a stab at comparing the effectiveness of deterrence and compellence. They rely on the qualitative evidence from \cite{Sechser2019}.

\footnote{See also \cite{Cebul2021, Kydd2017, Pauly2019}.}
Fuhrmann (2017), distinguishing between two types of threats, those triggering a crisis to change the status quo and those emitted during a crisis to enforce the status quo. Call the former “compellent” and the latter “deterrent.”

The tables suggest that, according to this definition, compellence is indeed harder than deterrence. Nuclear threats never succeeded when triggering a crisis and often succeeded in enforcing the status quo.

This evidence could support TNR’s claim about the effectiveness of deterrence. Yet the universe of possible cases is small. TNR itself would caution that the status quo may be ambiguous, that countries could successfully compel, by attempting to do so despite its cost, given uncertainty about countries’ resolve, and strategic selection into crises.

I study this problem using game-theoretic tools. States exchange nuclear threats before a preemption game. Country 1 decides whether to issue a compellent threat, requesting revisions to the status quo. Country 2 decides whether to accept country 1’s demand or to issue a deterrent threat, hoping to enforce the status quo. Country 1 decides whether to back down, accepting the status quo, or to stand firm, triggering a preemption game.

\[16\] It would be even smaller if we limit it to the set of crises between two nuclear-weapons states, or between two nuclear-weapons states with survivable second strikes, though the conclusion would be unchanged.
The baseline model concludes that compellence is indeed ineffective because it is incredible. Then the paper lets country 1 use any of the standard techniques bolstering the credibility of threats. If country 1 burns bridges, i.e. if it commits to attack after country 2’s rejection, then compellence is more effective than deterrence. By construction, deterrence would fail. Country 1 extracts concessions when country 2 lacks resolve. If country 1 uses a probabilistic threat, i.e. if it commits to attack with some probability after country 2’s rejection, then compellence may succeed, depending on how closely probabilistic threats resemble burning bridges. If country 1 feigns irrationality, i.e. if it mimics the behavior of types preferring war to the status quo, then compellence may fully succeed. Under some conditions, all types pool on the same, relatively generous offer. Country 2, uncertain of country 1’s rationality, accepts its offer.

Next, I conclude that a more favorable nuclear balance improves a country’s peaceful terms without increasing the risk of disaster. Country 1 simply exploits its improved bargaining leverage. Greater first-strike capabilities, however, may improve peaceful terms, but only under a heightened risk of nuclear disaster. Only the most resolved types of country 1 consider a first strike. Improved first-strike capabilities emboldens them. They make even more aggressive offers, tolerating a greater risk of war. I develop this argument below.
3 Game-Theoretic Argument

Consider a game between two countries, 1 and 2, illustrated in Figure 3. They exchange compellent and deterrent before entering a crisis preemption game, with reciprocal fear of surprise attack.

Country 1 chooses whether to trigger a crisis to revise the status quo. If it does not, the game ends. Country $i$ gets its peace payoff $\pi_i$, $i \in \{0, 1\}$, where $\pi_1 + \pi_2 = 1$. If country 1 triggers a crisis, it proposes to give $x_i$ to country $i$, $x_1 + x_2 = 1$.

Country 2 chooses whether to accept country 1’s demands. If it does, country 1’s proposal is implemented. If it does not, country 1 can back down, accepting the status quo after all, giving country $i$ a payoff of $\pi_i$, or stand firm, triggering a preemption game.

In the crisis preemption game, both countries simultaneously choose whether to attack. If neither attacks, the status quo prevails. Country $i$’s payoff is $\pi_i$. If at least one country attacks, war ensues. If both attack, country $i$ gets a payoff of $p_i - c_i + h_i$, where $p_i$ is the probability that country $i$ wins the conflict, $p_1 + p_2 = 1$, $c_i$ is country $i$’s cost of war, and $h_i$ is a ‘hostility’ parameter, or preference for attacking (Baliga and Sjöström N.d.). If only country $i$ attacks, then country $i$’s payoff is $p_i - c_i + h_i + f_i$ and country $j$’s payoff is $p_j - c_j - f_i$, where $f_i > 0$ is a first-strike parameter, improving country $i$’s payoff at the expense of country $j$. Countries are uncertain of
each other’s hostility, which we may call their type or “resolve.” $h_i$ is taken from a distribution with cdf $F_i$ over support $[h_i, \overline{h_i}]$.

Under mutually assured destruction (MAD), “there is no “fundamental” basis for an attack by either side” (Schelling 1960, 207). First-strike advantages do not make war preferable to peace. Still, countries fear being the target of a surprise attack, and may prefer to attack if they expect their enemy to do so. Technically, this is a game of strategic complements:

$$\pi_i - (p_i - c_i) > f_i - f_j \quad \forall i, j \neq i$$  \hspace{1cm} (3)

Call country $i$ a coordination type if it prefers to attack if and only if $j$ attacks, a dominant strategy dove if it prefers not to attack for any strategy of $j$, and a dominant strategy hawk if it prefers to attack for any strategy of $j$ (Baliga and Sjöström N.d., 6). Under MAD, there are some coordination types and dominant strategy doves:

$$h_i < -f_j < \overline{h_i} \quad \forall i, j \neq i$$  \hspace{1cm} (4)

but no dominant strategy hawks:

$$\pi_i - (p_i - c_i) > h_i + f_i \quad \forall h_i \forall i$$  \hspace{1cm} (5)

Dominant strategy hawks have “irrational” preferences.

We solve this game for a Perfect Bayesian Equilibrium, where strategies
are optimal given beliefs and the strategies of other players, and beliefs are consistent with Bayes’ rule whenever possible. In the preemption game, we say that an equilibrium is in cut-off strategies if country \( i \) chooses to attack if and only if \( h_i > h_i' \), for some \( h_i' \in [\underline{h}_i, \bar{h}_i] \).

We evaluate the relative effectiveness of “compellent” and “deterrent” threats. A compellent threat is issued when country 1 triggers a crisis to revise the status quo. It succeeds if country 2 accepts country 1’s offer. A deterrent threat is issued when country 2 rejects country 1’s offer and attempts to enforce the status quo. It succeeds if country 1 backs down.

We also evaluate the effect of military capabilities on the terms of peace and the odds of war, through their effect on the nuclear balance and first-strike advantages. Let \( b \) represent the nuclear balance. Greater values representing a balance more favorable to country 1, increasing country 1’s war payoff at the expense of country 2:

\[
\frac{\partial(p_1 - c_1)}{\partial b} = -\frac{\partial(p_2 - c_2)}{\partial b} > 0
\]  

(6)

### 3.1 Analysis

#### 3.1.1 Baseline Model

First, we construct an equilibrium where, under MAD, compellence fails and deterrence succeeds, because country 1 would never trigger a nuclear disaster:
Lemma 1 The following forms an equilibrium: Country 1 chooses the status quo if \( h_1 < h \) and otherwise triggers a crisis, offering \( x_2 < \pi_2 \), for some \( h \leq \tilde{h}_1 \); Country 2 accepts country 1’s offer if and only if \( x_2 \geq \pi_2 \); Country 1 backs down; Country \( i \) does not attack (\( i \in \{1, 2\} \)).

More generally, we can show the following:

Lemma 2 (i) If there were a peaceful revision of the status quo to \( x'_2 \neq \pi_2 \) in equilibrium, then country 2’s rejection of \( x'_2 \) would lead to war with positive probability. (ii) It is not sequentially rational for country 1 to stand firm if the preemption game produces war with positive probability; (iii) There is no equilibrium where a peaceful revision of the status quo occurs with positive probability.

In other words, a peaceful revision of the status quo requires a credible compellent threat (i), but under MAD, compellent threats are incredible (ii), hence they must fail (iii).

In addition, we can show that the presence of dominant strategy doves convinces all types not to attack, if the support of hostility parameters is sufficiently wide:

Lemma 3 If \( F_i \) is the uniform distribution and

\[
\left( \frac{\pi_1 - (p_1 - c_1) - (f_1 - f_2)}{\bar{h}_1 - \tilde{h}_1} \right) \left( \frac{\pi_2 - (p_2 - c_2) - (f_2 - f_1)}{\bar{h}_2 - \tilde{h}_2} \right) < 1 \quad (7)
\]
then there is a unique cut-off equilibrium of the preemption game, where country \( i \) does not to attack for any \( h_i \) (\( i \in \{1, 2\} \)).

These results confirm TNR’s intuition about the challenges of compellence. They also expose the conditions for its success. Country 1 could successfully coerce if it lacked the option of backing down after country 2’s rejection or if it convinced country 2 that it prefers war to a dishonorable peace. These are the standard tools bolstering the credibility of nuclear threats, highlighted by Schelling (1966). I consider each technique in turn.

3.1.2 Burning Bridges

Let country 1 “burn bridges,” making a commitment to stand firm and attack after country 2’s rejection. We conclude:

**Lemma 4** If country 1 burns bridges and \( F_i \) is the uniform distribution, the following forms the unique equilibrium: Country 1 triggers a crisis and offers:

\[
x_2^* = \begin{cases} 
  p_2 - c_2 + h_2 & \text{if } h_1 \leq h' \\
  p_2 - c_2 + \frac{1-(p_1-c_1)-(p_2-c_2)-h_1+h_2}{2} & \text{if } h_1 \in (h', h'') \\
  p_2 - c_2 - f_1 & \text{if } h_1 \geq h'' 
\end{cases}
\]
for some \( h' < h'' \). Country 2 accepts \( x_2 \) if and only if

\[
x_2 \geq \begin{cases} 
  p_2 - c_2 - f_1 & \text{if } h_2 < -f_1 \\
  p_2 - c_2 + h_2 & \text{if } h_2 \geq -f_1 
\end{cases}
\]

(9)

In the preemption game, country 2 attacks if and only if \( h_2 \geq -f_1 \).

By burning bridges, country 1 ensures that any rejection by country 2 triggers a nuclear disaster. Country 1 is “turning the tables” by “relinquishing the initiative” (Schelling 1966, 43-49). Country 1 then faces a traditional risk-return tradeoff. A more aggressive offer generates a more favorable peace but it is more likely to be rejected. Consistent with TNR, the more country 1 is willing to run the risk of nuclear disaster, the more favorable are the terms of peace.

In contrast with TNR, however, the more resolved a state is, the least likely it is to succeed. The reason is twofold. First, a greater display of resolve could coerce only if it corresponded with a greater risk of war after rejection, but by burning bridges, country 1 ensures that war follows any rejection. Second, country 1 chooses the terms of peace. The more resolved a state is, the more willing it is to run the risk of war. It chooses a more aggressive offer, which is more likely to be rejected.

Turning to the effect of country 1’s military capabilities on diplomatic outcomes, we conclude:
Result 1 If country 1 burns bridges: (i) as the nuclear balance becomes more favorable to country 1, it enjoys more favorable peaceful terms and the probability of war remains unchanged; (ii) as country 1’s first-strike capabilities improve, it may enjoy more favorable peaceful terms, under an increased probability of war.

A more favorable nuclear balance increases country 1’s war payoff at the expense of country 2. Country 1’s emboldenment is exactly offset by country 2’s timidity. Only the terms of peace change, not the probability of war. Country 1’s first-strike capabilities, for their part, matter only if country 2 would not attack, knowing that country 1 would, or only if country 2 is a strategy dominant dove. These types accept the lowest offers, made by the most hostile types of country 1. Greater first capabilities embolden these types of country 1. They make more aggressive offers, tolerating a greater risk of war.

The strategy of burning bridges may be rarely implemented\textsuperscript{17} Nevertheless, it serves as a useful benchmark for probabilistic threats, which we turn to next.

3.1.3 Probabilistic Threats

Let country 1 issue “probabilistic threats,” which may trigger a nuclear disaster. Formally, if country 2 rejects the offer, then with probability $1 - \epsilon$, the

\textsuperscript{17}For discussion of the Soviet Perimeter program, see, e.g., Hoffman (2009).
status quo prevails, and with probability \( \epsilon \), country 1 is committed to stand firm and attack \( (\epsilon \in (0, 1)) \). This model stands as a middle ground between the previous two, approximating the baseline model as \( \epsilon \) approaches 0 and the model of burning bridges as \( \epsilon \) approaches 1.

This model has broad applicability. Whenever policymakers deploy military power, officers assess proximate threats and decide whether to take actions that could quickly escalate. In the Cuban Missile Crisis, the shooting of a U-2 spy plane on October 27 could have triggered a quick escalation, and so could have U.S. efforts to hold the quarantine line. We conclude:

**Lemma 5** If country 1 issues probabilistic threats and \( F_1 \) is the uniform distribution, the following forms the unique equilibrium: Country 1 triggers a crisis and offers:

\[
x_2^* = \begin{cases} 
(1 - \epsilon)\pi_2 + \epsilon(p_2 - c_2 + h_2) & \text{if } h_1 \leq h' \\
(1 - \epsilon)\pi_2 + \epsilon \left[ p_2 - c_2 + \frac{(p_1 - c_1) - (p_2 - c_2) - h_1 + h_2}{2} \right] & \text{if } h_1 \in (h', h'') \\
(1 - \epsilon)\pi_2 + \epsilon(p_2 - c_2 - f_1) & \text{if } h_1 \geq h'' 
\end{cases}
\]

for some \( h' < h'' \). Country 2 accepts \( x_2 \) if and only if

\[
x_2 \geq \begin{cases} 
(1 - \epsilon)\pi_2 + \epsilon(p_2 - c_2 - f_1) & \text{if } h_2 < -f_1 \\
(1 - \epsilon)\pi_2 + \epsilon(p_2 - c_2 + h_2) & \text{if } h_2 \geq -f_1 
\end{cases}
\]

In the preemption game, country 2 attacks if and only if \( h_2 \geq -f_1 \).
By issuing probabilistic threats, country 1 engages in a competition in risk taking (Schelling 1966, 91). The greater is the risk of disaster (the closer $\epsilon$ is to 1), the more probabilistic threats resemble burning bridges, making compellence more effective, and increasing the concessions obtained by country 1. Probabilistic threats, therefore, exhibit the same comparative statics as burning bridges with regard to the nuclear balance and first-strike capabilities (see the Appendix).

3.1.4 The Rationality of Irrationality

Finally, let country 1 control whether to stand firm, but assume that it may be “irrational,” preferring war to the status quo, even without a first strike: $\bar{h}_1 > \pi_1 - (p_1 - c_1)$. Let $\epsilon = 1 - F_1(\pi_1 - (p_1 - c_1))$ be the mass of country 1 with these preferences. Otherwise, assume that condition (5) holds for all $h_2$, and conditions (3) and (4) hold. We ask whether a “rational” type, which would not trigger nuclear war, can extract concessions by appearing irrational. This strategy may be rare, but it has been considered, for example in Nixon’s initial approach to the resolution of the Vietnam War (Kimball and Burt 2015). We conclude:

**Lemma 6** If country 1 may be “irrational” and $F_i$ is the uniform distribution, there are critical values $\bar{h}_1'$, $\bar{h}_1''$, $h_2'$, such that if $\bar{h}_1 \in \left(\bar{h}_1', \bar{h}_1''\right)$, $\bar{h}_2 < \bar{h}_2'$, then the following forms an equilibrium, with appropriate off-the-
equilibrium-path beliefs: Country 1 triggers a crisis and offers

\[ x_2^* = \bar{x}_2 \equiv (1 - \epsilon)\pi_2 + \epsilon(p_2 - c_2 + \bar{h}_2) \]  

(12)

for all \( h_1 \). Country 2 accepts \( x_2 \) if and only if

\[
x_2 \geq \begin{cases} 
\bar{x}_2 \equiv (1 - \epsilon)\pi_2 + \epsilon(p_2 - c_2 - f_1) & \text{if } h_2 < -f_1 \\
(1 - \epsilon)\pi_2 + \epsilon(p_2 - c_2 + h_2) & \text{if } h_2 \geq -f_1
\end{cases}
\]  

(13)

Country 1 stands firm after \( x_2 \geq x_2^* \), if and only if \( h_1 > \pi_1 - (p_1 - c_1) \) and after \( x_2 < x_2^* \) if and only if \( h_1 > \pi_1 - (p_1 - c_1) - F_2(-f_1)f_1 \).

In the preemption game, country 2 attacks if and only if \( h_2 \geq -f_1 \). After \( x_2 \geq x_2^* \), country 1 attacks if and only if \( h_1 > -f_2 \). After \( x_2 < x_2^* \), country 1 attacks if and only if \( h_1 > (1 - F_2(-f_1))(-f_2) + F_2(-f_1)(\pi_1 - (p_1 - c_1) - f_1) \).

In this equilibrium, all types pool on the same offer, which is accepted by country 2. Compellence succeeds, not because it signals anything about country 1’s willingness to fight, but precisely because it does not. Country 2 cannot discern whether it faces a bluffing rational type or an irrational type, which would prefer to trigger a nuclear war. To sustain this equilibrium, country 2 cannot be too bellicose (\( \bar{h}_2 < \bar{h}_2' \)), so that it accepts significant concessions, and country 1’s most bellicose preferences must be intermediate (\( \bar{h}_1 \in (\bar{h}_1', \bar{h}_1'') \)), so that the threat of its irrationality is effective, but it is willing to accept peaceful concessions.
Turning to the effect of military capabilities on diplomatic outcomes:

**Result 2** If country 1 may be irrational: (i) as the nuclear balance becomes more favorable to country 1, it enjoys more favorable peaceful terms and the probability of war remains unchanged; (ii) as country 1’s first-strike capabilities improve, it affects neither the terms of peace nor the probability of war.

As before, the nuclear balance may safely improve country 1’s diplomatic outcome. Here, country 1’s first strike capabilities have no effect, because its offer targets the most resolved type of country 2, which would fight in the preemption game, denying country 1 the possibility of striking first.

4 Conclusion

The Theory of the Nuclear Revolution (TNR) is under attack, but calls for its rejection do not seem warranted. Recent critiques are quite narrow, generally accepting that nuclear weapons have been a force for peace and that they are the ultimate tools of deterrence. They mainly focus on the U.S. drive for nuclear superiority, arguing that it was rational and there is a link between military capabilities and political outcomes. Yet they do little to explain how nuclear weapons compel, and how greater military capabilities provide coercive leverage. The fact that U.S. policymakers pursued nuclear superiority does not mean that it served U.S. national security interests. In
fact, TNR offered a compelling explanation for the U.S. pursuit for nuclear superiority, explaining how policymakers may be trapped by previous discourse on competent nuclear management.

This paper shows that compellence may overcome deterrence once we account for strategic selection into crises, and allow coercers to use standard techniques to bolster the credibility of their threats. At the same time, it shows that improved coercion may come at the cost of an increased risk of nuclear disaster. First-strike capabilities prove especially destabilizing. If TNR scholars were too sanguine about the stability of MAD, they were right to raise the alarm in the late 1970s and 1980s, when Washington came close to escaping the nuclear stalemate. Future nuclear policy choices should weigh any potential coercive benefit with the increased risk of disaster.

References


Wasn’t There: Soviet Reactions to the Late Cold War Nuclear Balance.”


Pauly, Reid B. C. 2019. “Stop or I’ll Shoot, Comply and I Won’t”: Coercive Assurance in International Politics PhD thesis Massachusetts Institute of Technology. [https://www.reidpauly.com/home-1](https://www.reidpauly.com/home-1).


Appendix I: Proof of the Formal Results

Proof. (Proof of Lemma 1). It is straightforward to show, using backward induction, that the strategies for an equilibrium. In the preemption game, strategies form a mutual best response given condition \((5)\). ■

Proof. (Proof of Lemma 2). For (i), let peace prevail if country 2 rejects \(x'_2\). Country 2 accepts if and only if \(x'_2 \geq \pi_2\). Country 1 triggers the crisis, offering \(x'_2\) if and only if \(x'_2 \leq \pi_2\). These imply \(x'_2 = \pi_2\). For (ii), if the preemption game produces war with positive probability, then it offers country 1 strictly less than \(\pi_1\), given conditions \((5)\) and \((4)\). Country 1 strictly prefers to back down. For (iii), note that it follows from (i) and (ii). ■

Proof. (Proof of Lemma 3). In a cut-off equilibrium, country \(i\)'s net benefit of attacking is \(h_i + f_j - F_j(h'_j)[\pi_i - (p_i - c_i) - (f_i - f_j)]\). It is zero at the cut-off \(h'_i\) so that, using \(F_i\) as the uniform distribution:

\[
\frac{h'_i + f_j}{\pi_i - (p_i - c_i) - (f_i - f_j)} = \frac{h'_i - h_j}{h_j - h_j} \tag{14}
\]

Simplifying this system of equations (for \(i \in \{1, 2\}\)), we have:

\[
h'_i = h_i - \alpha_i[h_i + f_j] - \alpha_i \beta_i[h_j + f_i] \tag{15}
\]

where \(\alpha_i = \frac{1}{1 - \frac{\pi_i - (p_i - c_i) - (f_i - f_j)}{\pi_j - (p_j - c_j) - (f_j - f_i)}} \frac{\pi_i - (p_i - c_i) - (f_i - f_j)}{h_j - h_j}, \; \beta_i = \frac{\pi_i - (p_i - c_i) - (f_i - f_j)}{h_j - h_j} \)
Using condition (7), $h_i' - h_i \geq 0$ if and only if

$$
\left(\frac{\pi_i - (p_i - c_i) - (f_i - f_j)}{h_i + f_j}\right) \left(\frac{\pi_j - (p_j - c_j) - (h_j + f_j)}{h_j - h_i}\right) \geq 1 \quad (16)
$$

which follows from conditions (3) to (4).

**Proof.** (Proof of Lemma 4). Proceed by backward induction. Country 2’s decision to attack, and its response to $x_2$, are straightforward.

Moving up, consider country 1’s optimal offer. Clearly, country 1 prefers $x_2 = p_2 - c_2 + h_2$ to any higher offer. It also prefers $x_2 = p_2 - c_2 - f_1$, accepted by $h_2 \leq -f_1$, to any lower $x_2$, rejected by them, given conditions (3) and (5).

Country 1’s problem is to identify $h_2' = x_2 - (p_2 - c_2)$ to solve:

$$
max_{h_2' \in [-f_1, h_2]} F_2(h_2')[1 - (p_2 - c_2) - h_2'] + (1 - F_2(h_2'))(p_1 - c_1 + h_1) \quad (17)
$$

This is a convex problem. The interior solution is given by

$$
1 - (p_1 - c_1) - (p_2 - c_2) - h_1 + h_2 - 2h_2' = 0 \quad (18)
$$

Consider the following conditions:

$$
1 - (p_1 - c_1) - (p_2 - c_2) \leq \overline{h}_1 - h_2 + 2f_1 \quad (19)
$$

$$
1 - (p_1 - c_1) - (p_2 - c_2) \geq \underline{h}_1 - h_2 + 2\overline{h}_2 \quad (20)
$$
If condition (19) holds, then for any \( h_1 \), country 1 chooses the lower bound \( h'_2 = -f_1 \). The optimal offer is given by equation (8), where \( h' < h_1 = h'' \).

If condition (20) holds, then for any \( h_1 \), country 1 chooses the upper bound \( h'_2 = \overline{h}_2 \). The optimal offer is given by equation (8), where \( h' = \overline{h}_1 < h'' \).

If conditions (19) and (20) fail, then for some \( h_1 \), the optimal offer is the interior solution. The optimal offer can be summarized by equation (8), where \( h_1 \leq h' < h'' \leq \overline{h}_1 \). If the upper and lower bounds are offered by some type, then \( h' \) and \( h'' \) are defined by

\[
1 - (p_1 - c_1) - (p_2 - c_2) = h' - \overline{h}_2 + 2 \overline{h}_2 \tag{21}
\]

\[
1 - (p_1 - c_1) - (p_2 - c_2) = h'' - \overline{h}_2 - 2f_1 \tag{22}
\]

Finally, country 1 prefers to trigger a crisis, since it can offer \( x_2 = p_2 - c_2 + \overline{h}_2 \), accepted by any \( h_2 \), and get more than \( \pi_1 \), by condition (5).

**Proof.** (Proof of Result 1.(i)) Given condition (6), \( \frac{\partial x_2^*}{\partial b} = \frac{\partial p_2 - c_2}{\partial b} < 0 \) and \( \frac{\partial h'}{\partial b} = \frac{\partial h''}{\partial b} = 0 \), using equations (21) and (22). Similarly, \( \frac{\partial \text{prob(war)}}{\partial b} = 0 \) since

\[
\text{prob(war)} = \begin{cases} 
0 & \text{if } h_1 \leq h' \\
1 - F_2 \left( \frac{1-(p_1-c_1)-(p_2-c_2)-h_1+h_2)}{2} \right) & \text{if } h_1 \in (h', h'') \\
1 - F_2(-f_1) & \text{if } h_1 > h'' 
\end{cases} \tag{23}
\]
For (ii), note using equation (8) that $\frac{\partial x^*}{\partial f_1} = 0$ for $h_1 < h''$ and $\frac{\partial x^*}{\partial f_1} < 0$ for $h_1 \geq h''$. Also note using equations (21) and (22) that $\frac{\partial h'}{\partial f_1} = 0$, $\frac{\partial h''}{\partial f_1} > 0$. Finally, note using equation (23) that $\frac{\partial \text{prob(\text{war})}}{\partial f_1} = 0$ for $h_1 < h''$ and $\frac{\partial \text{prob(\text{war})}}{\partial f_1} > 0$ for $h_1 \geq h''$. 

**Proof.** (Proof of Lemma 5). Proceed by backward induction. Country 2’s decision to attack, and its response to $x_2$, are straightforward.

Moving up, consider country 1’s optimal offer. Clearly, it prefers $x_2 = (1-\epsilon)\pi_2 + \epsilon(p_2 - c_2 + \bar{h}_2)$ to any higher offer. Also, it prefers $x_2 = (1-\epsilon)\pi_2 + \epsilon(p_2 - c_2 - f_1)$, accepted by $h_2 \leq -f_1$, to any lower $x_2$, rejected by them, given conditions (3) and (5), and $\epsilon > 0$. Let $h'_2$ be the type $h_2$ indifferent between accepting and rejecting $x_2$, $h'_2 = \frac{x_2-(1-\epsilon)\pi_2}{\epsilon} - (p_2 - c_2)$. Country 1’s problem is to identify $h'_2$ to solve

$$\max_{h'_2 \in [-f_1, \bar{h}_2]} F_2(h'_2)[(1-\epsilon)\pi_1 + \epsilon[1-(p_2 - c_2) - h'_2]] + (1-F_2(h'_2))((1-\epsilon)\pi_1 + \epsilon(p_1 - c_1 - h_1))$$

which reduces to (17). The rest of the proof follows the above logic. 

**Result 3** If country 1 issues probabilistic threats: (i) as the nuclear balance becomes more favorable to country 1, it enjoys more favorable peaceful terms and the probability of war remains unchanged; (ii) as country 1’s first-strike capabilities improve, it may enjoy more favorable peaceful terms, under an increased probability of war.

**Proof.** The proof follows the logic of the proof of Result 1. 

63
Proof. (Proof of Lemma 6). First specify countries’ beliefs (cut-off values $\bar{h}_1$, $\bar{h}_1''$, and $\bar{h}_2$ are discussed below):

- After $x_2 \geq \underline{x}$, country 2 believes that $h_1 \sim U[h_1', \bar{h}_1]$. After $x_2 < \underline{x}$, country 2 believes that $h_1 = \bar{h}_1$.

- After $x_2 < \underline{x}$ is rejected, country 1 believes that $h_2 \sim U[h_2', \bar{h}_2]$. After $x_2 \in [\underline{x}, \bar{x}]$ is rejected, country 1 believes that $h_2 \sim U[\underline{x}-(1-\epsilon)x_2-(p_2-c_2), \bar{h}_2]$. After $x_2 > \bar{x}$ is rejected, country 1 believes that $h_2 = \bar{h}_2$.

- After $x_2 \geq \underline{x}$ is rejected and country 1 stands firm, country 2 believes that country 1’s type $h_1 \sim U[\pi_1-(p_1-c_1), \bar{h}_1]$. After $x_2 < \underline{x}$ is rejected and country 1 stands firm, country 2 believes that $h_1 \sim U[\pi_1-(p_1-c_1)-F_2(-f_1)f_1, \bar{h}_1]$.

To prove that the above forms a PBE, proceed by backward induction.

After any $x_2$, country 2 expects country 1 to attack if $\pi_1-(p_1-c_1) > -f_2$, which follows condition (3). Thus, country 2 attacks if and only if $h_2 \geq -f_1$.

After $x_2 \geq \underline{x}$, country 1 expects country 2 to attack if $\underline{x}-\frac{(1-\epsilon)x_2}{\epsilon}-(p_2-c_2) \geq -f_1$, which is true since $x_2 \geq \underline{x}$. Therefore, country 1 attacks if and only if $h_1 \geq -f_2$. After $x_2 < \underline{x}$, country 1 expects country 2 to attack with probability $1 - F_2(-f_1)$. Therefore, country 1 attacks if and only if $h_1 > (1 - F_2(-f_1))(-f_2) + F_2(-f_1)(\pi_1-(p_1-c_1)-f_1)$.

Now consider country 1’s decision to back down or stand firm. After $x_2 \geq \underline{x}$, country 1 expects country 2 to attack. Any $h_1 < -f_2$ would not
attack and prefers to back down if and only if \( \pi_1 > p_1 - c_1 - f_2 \), which holds by condition (3). Any \( h_1 \geq -f_2 \) would attack and prefers to stand firm if and only if \( h_1 > \pi_1 - (p_1 - c_1) \). After \( x_2 < x_2' \), country 1 expects country 2 to attack with probability \( 1 - F_2(-f_1) \). Any \( h_1 < (1 - F_2(-f_1))(-f_2) + F_2(-f_1)(\pi_1 - (p_1 - c_1) - f_1) \) would not attack and prefers to back down if and only if \( \pi_1 > (1 - F_2(-f_1))(p_1 - c_1 - f_2) + F_2(-f_1)(\pi_1 - (p_1 - c_1) - f_1) \), which holds by condition (3). Any \( h_1 \geq (1 - F_2(-f_1))(-f_2) + F_2(-f_1)(\pi_1 - (p_1 - c_1) - f_1) \) would attack and prefers to stand firm if and only if \( h_1 > \pi_1 - (p_1 - c_1) - F_2(-f_1)f_1 \).

Moving up, consider country 2’s evaluation of \( x_2 \). If \( x_2 \geq x_2' \), country 2 expects country 1 to stand firm and attack with probability \( \epsilon \) and to back down with probability \( 1 - \epsilon \). It proceeds as in condition (13). If \( x_2 < x_2' \), country 2 expects country 1 to back down. It rejects since \( x_2 < x_2' < \pi_2 \). Thus, it proceeds as in condition (13).

Moving up, consider country 1’s optimal offer. Clearly, \( x_2 = x_2' \) is preferable to \( x_2 > x_2' \). Choosing \( x_2 \in [x_2, x_2'] \) reduces to (17). Given the proof of Lemmas 4 and 5 all types \( h_1 \) strictly prefer \( x_2 = x_2' \) to any \( x_2 \in [x_2, x_2'] \) if condition (20) holds. Condition (20) specifies an upper bound \( \overline{h}_2' \) for \( \overline{h}_2 \).

Next evaluate country 1’s preference between \( x_2 = x_2' \) and any \( x_2 < x_2' \). Any \( h_1 \leq \pi_1 - (p_1 - c_1) - F_2(-f_1)f_1 \) prefers \( x_2 = x_2' \) if and only if \( 1 - x_2' > \pi_1 \), which holds given condition (5) for \( i = 2 \). Any \( h_1 > \pi_1 - (p_1 - c_1) - F_2(-f_1)f_1 \) prefers \( x_2 = x_2' \) if and only if \( 1 - x_2' > p_1 - c_1 + \overline{h}_1 + F_2(-f_1)f_1 \) or

\[
\pi_1 - (p_1 - c_1) - \overline{h}_1 - F_2(-f_1)f_1 + \epsilon(\pi_2 - (p_2 - c_2) - \overline{h}_2) > 0 \tag{25}
\]
which holds if and only if $h_1 \in (h_1', h_1'')$ for some values $h_1', h_1''$. Indeed, the left-hand side of the above condition is concave in $h_1$, since $\frac{\partial^2 \epsilon}{\partial h_1^2} < 0$. The condition fails when $h_1 = \pi_1 - (p_1 - c_1)$ and when $h_1$ tends to infinity. It can be satisfied, along with all others, in the following example: $\pi_i = p_i = 1/2$, $c_1 = 1$, $c_2 = 4$, $f_i = \frac{1}{4} \forall i$, $h_i = -1 \forall i$, $h_1 = 2$, $h_2 = 0$.

Moving up, country 1 prefers to trigger a crisis, since offering $x_2^*$ generates strict concessions ($1 - x_2^* > \pi_1$) and is accepted with probability 1. ■

**Proof.** (Proof of Result 2) (i) and (ii) are immediate from equation (12) and the fact that peace prevails. ■
Appendix II: Coding Rules and Decisions for Tables 1 and 2

Cases are taken from Sechser and Fuhrmann (2017, Table III.1, p. 128), excluding catalytic threats between allies or neutral states. Whether threats trigger a crisis is based on the ICB project and Sechser and Fuhrmann (2017), unless otherwise noted. The nuclear coercer was successful if it achieved “victory” peacefully, according to the ICB project, unless otherwise noted.

Korean War, 1953

*Summary of the crisis:* After PRC-DPRK attacks on UN troops on April 16, 1953, Washington approved contingency nuclear war plans on May 19. The next day, Secretary of State John Foster Dulles told Indian Prime Minister Jawaharlal Nehru of U.S. intentions to “make a stronger rather than a lesser military exertion,” a threat to be relayed to Beijing (quoted in Sechser and Fuhrmann 2017, p. 176). Four days later, Washington performed a nuclear test. A month later, Beijing agreed to an armistice, making key concessions on the issue of prisoners of war.

*Did nuclear threats trigger the crisis?* No. The crisis began on April 16. The ICB dataset identifies three crises during the war: June-September 1950, September 1950-July 1951, and April-July 1953. Sechser and Fuhrmann (2017, p. 128) identify a single crisis for the entire war, but adopt a narrower time frame in other wars, i.e. the Indochina War and the Vietnam War. I do so here.

*Was the nuclear coercer successful?* Yes. Hostilities ended, with key concessions from the PRC.

Indochina War, 1954

*Summary of the crisis:* On March 13, 1954, French positions were attacked at Dien Bien Phu. In March, Dulles told the press that the conflict might easily escalate to massive attacks on China itself. In April, he declared that NATO should use nuclear weapons “whenever or wherever it would be of advantage to do so, taking account of all relevant factors.” Dien Bien Phu fell on May 7. Later in May, news leaked that the Chairman of the Joint Chiefs of Staff, Admiral Arthur W. Radford, had told the House Foreign Affairs Committee that “any United States intervention in Indochina should be on an all-out basis, including use of atomic weapons” (quoted in Sechser and Fuhrmann 2017, p. 183). A peace agreement was signed on July 21.

*Did nuclear threats trigger the crisis?* No. The crisis began with the attack on French forces.

*Was the nuclear coercer successful?* Yes. Terms of peace improved after U.S. nuclear coercion (Trachtenberg 2013, p. 22). Though the ICB project codes the US as defeated, Sechser and Fuhrmann (2017, p. 128) codes US nuclear coercion efforts as “apparently successful.”
First Taiwan Strait crisis, 1954-5

*Summary of the crisis:* In August 1954, the creation of SEATO, to include the ROC, was being discussed. The PRC began shelling Quemoy and Matsu, controlled by the ROC, on September 3. Nine days later, Washington sent the Seventh Fleet to the region. In March 1955, it publicly suggested using tactical nuclear weapons. The PRC proposed negotiations. The crisis abated.

*Did nuclear threats trigger the crisis?* No. The crisis began with negotiations over SEATO (per the ICB project) or the shelling of Quemoy and Matsu (per Sechser and Fuhrmann 2017, p. 189).

*Was the nuclear coercer successful?* Yes. The status quo endured. The ICB project codes the outcome as a stalemate. Sechser and Fuhrmann (2017, p. 128) codes US efforts as “apparently successful.”

**Suez crisis, 1956**

*Summary of the crisis:* On July 26, 1956, Egypt nationalized the canal from an Anglo-French corporation. Israel invaded Sinai on October 29, later joined by the UK and France. On November 5, Soviet Prime Minister Nikolai Bulganin asked British Prime Minister Anthony Eden: “In what situations would Britain find herself if she was attacked by stronger states, possessing all types of modern destructive weapons?” (quoted in Sechser and Fuhrmann 2017, p. 225). He reminded Eisenhower that the US and USSR “are two great powers possessing all contemporary forms of armaments, including atom and hydrogen weapons,” and warned: “If this war is not stopped, it is fraught with danger and can grow into third world war” (U.S. Department of State, 1956, pp. 993-994). The UK and France agreed to cease fire and withdraw their forces. Israel later withdrew its forces.

*Did nuclear threats trigger the crisis?* No. The crisis began with the nationalization of the canal.

*Was the nuclear coercer successful?* Yes. Britain, France, and Israel withdrew from the area.

**Second Taiwan Strait crisis, 1958**

*Summary of the crisis:* The PRC began shelling Quemoy and Matsu on August 23. Two days later, Washington deployed military forces to the region, invoking the nuclear option. On September 4, Dulles said that “acquiescence” to the PRC’s aggression “would threaten peace everywhere.” Three days later, Soviet premier Nikita Khrushchev complained that Washington was “trying [...] to resort to atomic blackmail against China.” On September 17, Washington deployed three nuclear-capable artillery guns on Quemoy. In late September, Air Force Secretary James Douglas declared that the U.S.’s “most modern fighters are on the spot” and “are as capable of using high-explosive bombs as more powerful weapons” (quoted in Sechser and Fuhrmann 2017, pp. 195-196). Talks between the US and PRC led to a ceasefire on October 6, later extended. The crisis abated. On October 23, the United States and the ROC issued a joint communique, reaffirming their alliance while stressing its defensive nature.
Did nuclear threats trigger the crisis? No. The crisis began either in July, with the deployment of PRC forces (according to the ICB dataset), or on August 23, with the shelling of Quemoy and Matsu (according to Sechser and Fuhrmann 2017).

Was the nuclear coercer successful? Yes. The ICB dataset codes all actors (PRC, ROC, and the US) as victorious. The PRC could see the joint communique as a victory. Yet the status quo endured. U.S. effort are seen as “apparently successful” by Sechser and Fuhrmann (2017, p. 128).

Berlin ultimatum I, 1958-9

Summary of the crisis: On November 27, 1958, Khrushchev gave six months for West Berlin to be demilitarized and declared a “Free City,” threatening to sign a peace treaty with East Germany, which would control access routes to West Berlin, and come to its defense against the West. On December 1, he threatened to enforce the agreement, telling Hubert Humphrey: “We have rockets... We mean business,” invoking a recent hydrogen-bomb test (quoted in Sechser and Fuhrmann 2017, p. 134). In May 1959, Washington redeployed aircraft carriers, armed with nuclear weapons, to the Mediterranean. Khrushchev revoked his ultimatum.

Did nuclear threats trigger the crisis? Yes. The crisis began with the Soviet ultimatum.

Was the nuclear coercer successful? No. The status quo endured. The ICB dataset codes the outcome as a stalemate. Sechser and Fuhrmann (2017, p. 128) code Soviet efforts as a “failure.”

Berlin ultimatum II, 1961

Crisis summary: Khrushchev renewed his ultimatum in June 1961, threatening “all the consequences” after any violation of East German sovereignty (U.S. Department of State, 1961, pp. 90-91). On August 13, he built the Berlin Wall. In October, he dropped his ultimatum.

Did nuclear threats trigger the crisis? Yes. Though the ICB codes the crisis as beginning in August 1961, triggered by a non-state actor - refugees moving to West Berlin - Sechser and Fuhrmann (2017, pp. 136-137) begin their narrative with Khrushchev’s ultimatum, a conventional choice.

Was the nuclear coercer successful? No. Though the ICB dataset codes the Soviet Union as achieving “victory” - it did restrict defections to the West, the status quo otherwise prevailed. Sechser and Fuhrmann (2017, p. 128) codes this Soviet nuclear coercion effort as a “failure.”

Cuban Missile Crisis, 1962

Summary of the crisis: On October 16, Kennedy learned of the Soviet deployment to Cuba and announced a quarantine on October 22 to remove them. Two days later, the US raised its alert state to DEFCON 2. On October 28, the US and USSR agreed on the removal of Soviet missiles.
Did nuclear threats trigger the crisis? No. The crisis began on October 16.

Was the nuclear coercer successful? Yes. The USSR removed its missiles from Cuba.

**Seizure of USS Pueblo, 1968**

*Summary of the crisis:* On January 21, 1968, a DPRK commando unit attempted to assassinate the ROK president in his residence. On January 22, the DPRK seized the USS Pueblo, a U.S. intelligence vessel. President Johnson demanded the return of the ship and its crew, threatening a military response, deploying forces, equipped with nuclear weapons, to the area. Washington eventually backed down. The DPRK returned the crew but not the ship itself.

Did nuclear threats trigger the crisis? No. The crisis began either with the assassination attempt (per the ICB dataset) or the seizure of the Pueblo (per Sechser and Fuhrmann 2017, p. 166).

Was the nuclear coercer successful? No. North Korea returned the crew but not its ship.

**Sino-Soviet border crisis, 1969**

*Summary of the crisis:* On March 2, 1969, Chinese forces ambushed Soviet guards on Zhenbao Island in the Ussuri River. Moscow struck back on March 15, later placing their Strategic Rocket Forces on high alert, reminding Beijing in radio broadcasts of its nuclear superiority. High-level negotiations began on October 20, reducing tensions. By early February 1970, the Soviets had ceded Zhenbao and other disputed islands, and adopted a less aggressive patrolling posture.

Did nuclear threats trigger the crisis? No. The crisis began on March 2.

Was the nuclear coercer successful? Yes. The status quo was enforced. The ICB dataset, which ends the crisis in October 1969, codes the outcome as a “stalemate.” Sechser and Fuhrmann (2017, p. 128) describes Soviet efforts as “apparently successful.”

**Vietnam War, 1969**


Did nuclear threats trigger the crisis? Yes. The ICB dataset does not record a crisis for this specific episode. Sechser and Fuhrmann (2017, p. 142)’s narrative begins with U.S. threats.

Was the nuclear coercer successful? No. The United States did not extract any concessions.
Bangladesh War, 1971

Summary of the crisis: In 1970, the Awami League won the right to form the government of Pakistan. The national assembly was suspended. On March 25, West Pakistani forces attacked Dacca University students. The following day, East Pakistani politicians declared independence. As East Pakistan, supported by India, was prevailing, Washington deployed a task force to the region, equipped with nuclear weapons. The following day, India agreed to a ceasefire.

Did nuclear threats trigger the crisis? No. The crisis began on March 25.

Was the nuclear coercer successful? Yes. Nuclear threats helped end the conflict. The ICB does code Pakistan as defeated, but Sechser and Fuhrmann (2017, p. 128) describes U.S. coercion as “apparently successful”.

Yom Kippur war, 1973


Did nuclear threats trigger the crisis? No. The crisis began on October 5.

Was the nuclear coercer successful? Yes. All parties agreed to a ceasefire. Sechser and Fuhrmann (2017) identifies nuclear threats by “Israel/United States” against “United States/Egypt/Soviet Union,” which were “apparently successful.” I ignore Israel’s attempt to catalyze U.S. actions and focus on U.S. nuclear coercion against Egypt and USSR.

Falklands War, 1982

Summary of the crisis: On March 31, 1982, the UK learned of an imminent invasion by Argentina of the Falklands islands. Argentina took control of the islands on April 2. The UK deployed troops to area, carrying nuclear weapons, and attacked on June 3, winning on June 14.

Did nuclear threats trigger the crisis? No. The crisis began on March 31.

Was the nuclear coercer successful? No. The British did not peacefully prevail.

Kashmir crisis, 1990

told his counterpart that “war clouds would hover over the subcontinent if timely action was not taken,” who later stated that he “comprehended the nuclear connotations of Yakub’s warning and reported immediately to Prime Minister Singh” (quoted in Ganguly and Hagerty 2005, p. 89). Both countries deployed forces to the border. Pakistan allegedly readied its nuclear arsenal, hoping to catalyze U.S. intervention. In May, U.S. Deputy National Security Advisor Robert Gates traveled to the region. The crisis soon abated.

*Did nuclear threats trigger the crisis?* Yes. The ICB project says that it began when Pakistan formulated its policy on January 14, reacting a January 13 massacre. The dates appear incorrect. Ganguly and Hagerty (2005, p. 88) call the January 20 massacre “the first of many spasms of mass violence”. Pakistan chose its policy on January 21, issuing its nuclear threats.

*Was the nuclear coercer successful?* No. The status quo endured. Sechser and Fuhrmann (2017, p. 128) codes Pakistan as coercing “India/United States” in an “apparently successful” effort, yet their narrative discusses Pakistani success in catalyzing U.S. intervention, and acknowledge that India could not detect Pakistan’s nuclear preparations (Sechser and Fuhrmann, p. 230).

**Third Taiwan Strait, 1995-6**

*Summary of the crisis:* On May 22, 1995, Washington allowed ROC president Lee Teng-Hui to visit his alma mater, Cornell University, in June. The PRC interpreted this move as promoting Taiwanese independence, and conducted military exercises in the East China Sea in July. In October, PRC Lieutenant General Xiong Guangkai told U.S. diplomat Charles Freeman that unlike in the 1950s, when the United “threatened nuclear strikes on us,” the PRC “can hit back. So you will not make those threats. In the end you care more about Los Angeles than you do about Taipei” (quoted in Sechser and Fuhrmann 2017, p. 170). In March 1996, the PRC tested nuclear-capable M-9 missiles near Taiwan, hoping to hurt Lee’s chances in the upcoming elections. These were held as scheduled on March 23. The PRC ended its military exercises.

*Did nuclear threats trigger the crisis?* No. The crisis begins on May 22.

*Was the nuclear coercer successful?* No. The PRC did not obtain any concessions. Sechser and Fuhrmann (2017, p. 128) codes nuclear threats by “China/US” against “Taiwan/China/US,” but only discuss those issued by the PRC (pp. 170-171), my focus of analysis as well.

**Kargil War, 1999**

*Summary of the crisis:* On May 9, 1999, rockets were launched against Kargil, Kashmir. India blamed Pakistan, retaliating against guerrilla positions on May 26. The next day, Prime Minister Nawaz Sharif claimed that Pakistan could respond to India on “equal terms,” later claiming: “If there is war, or if the present confrontation continues on the borders, it will bring so much devastation, the damage of which will never be repaired.” (quoted in Sechser and Fuhrmann 2017, pp. 147-148). Pakistan allegedly readied its arsenal, detected by U.S. and Indian
intelligence. On July 4, Sharif met with President Bill Clinton, who insisted on Pakistan’s withdrawal. Sharif agreed, ending the war.

*Did nuclear threats trigger the crisis?* No. The crisis began on May 9.

*Was the nuclear coercer successful?* No. Pakistan failed to obtain concessions and lost the war (on whether nuclear threats deterred India from crossing the Line of Control, see Sechser and Fuhrmann 2017, pp. 151-152).

**Indo-Pakistani border crisis, 2001-2**

*Summary of the crisis:* On December 13, 2001, Pakistani-backed militants attacked the Indian parliament. The next day, Indian Prime Minister Atal Bihari Vajpayee vowed to fight “a war to the finish” and launched Operation Parakram, later adding that “no weapon would be spared in self-defence”. On January 12, 2002, Pakistani President Parvaz Musharraf agreed to crack down on terrorism. On May 14, terrorists attacked an Indian army base. In early June, Indian Defense Secretary Yogendra Narain said that a Pakistani nuclear attack would lead to “mutual destruction” (quoted in Sechser and Fuhrmann 2017, pp. 155-158). On June 6, Musharraf agreed to “permanently” crack down on terrorism. Operation Parakram ended in October.


*Was the nuclear coercer successful?* No. Pakistan’s fight against terrorism was limited. The ICB project codes India as achieving “victory” in the first crisis and a “stalemate” in the second, but Sechser and Fuhrmann (2017, pp. 128, 156) describes its efforts as a “failure.”

**Korean crisis, 2013**

*Summary of the crisis:* On February 12, the DPRK conducted a nuclear test and threatened to hit the ROK and the US with “lighter and smaller nukes.” Arguably, it hoped for a loosening of sanctions, recognition as a nuclear power, and an end to joint US-ROK exercises. Instead, these were held the next day. On March 7, the U.N. condemned the nuclear test, imposing additional sanctions. On March 11, the DPRK nullified the 1953 armistice and threatened to use “the right to a pre-emptive nuclear attack to destroy the strongholds of the aggressors” (quoted in Sechser and Fuhrmann 2017, pp. 162-163). In June, it proposed a resumption of six-party talks.

*Did nuclear threats trigger the crisis?* Yes.

*Was the nuclear coercer successful?* No. The DPRK did not obtain any concessions.
II

<table>
<thead>
<tr>
<th></th>
<th>Stand Firm</th>
<th>Submit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand Firm</td>
<td>$d_I, d_{II}$</td>
<td>$w_I, s_{II}$</td>
</tr>
<tr>
<td>Submit</td>
<td>$s_I, w_{II}$</td>
<td>$c_I, c_{II}$</td>
</tr>
</tbody>
</table>

Figure 1: Nuclear Crisis as a Game of Chicken

II

<table>
<thead>
<tr>
<th></th>
<th>Stand Firm</th>
<th>Submit</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand Firm</td>
<td>$d_I, d_{II}$</td>
<td>$fd_I + (1-f)w_I, fd_{II} + (1-f)s_{II}$</td>
<td></td>
</tr>
<tr>
<td>Submit</td>
<td>$fd_I + (1-f)s_I, fd_{II} + (1-f)w_{II}$</td>
<td>$c_I, c_{II}$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Nuclear Crisis as a Game of Chicken with Exogenous Risk
Figure 3: Extensive Form

Country 1

Status Quo

\( \pi_1, \pi_2 \)

Trigger

Country 2

Accept

\( x_1, x_2 \)

Reject

Back Down

Country 1

Stand Firm

Country 2

\[ p_1 - c_1 + h_1, \ p_2 - c_2 + h_2 \]
\[ p_1 - c_1 - f_2, \ p_2 - c_2 + h_2 + f_2 \]

\[ p_1 - c_1 + h_1 + f_1, \ p_2 - c_2 - f_1 \]
\[ \pi_1, \pi_2 \]
Table 1: Nuclear Threat and Crisis Outcome
– Nuclear Threat Triggers Crisis

<table>
<thead>
<tr>
<th>Crisis Name</th>
<th>Years</th>
<th>Challenger</th>
<th>Target</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin ultimatum I</td>
<td>1958-9</td>
<td>USSR</td>
<td>US</td>
<td>Failure</td>
</tr>
<tr>
<td>Berlin ultimatum II</td>
<td>1961</td>
<td>USSR</td>
<td>US</td>
<td>Failure</td>
</tr>
<tr>
<td>Vietnam War</td>
<td>1969</td>
<td>US</td>
<td>USSR, North Vietnam</td>
<td>Failure</td>
</tr>
<tr>
<td>Kashmir crisis</td>
<td>1990</td>
<td>Pakistan</td>
<td>India</td>
<td>Failure</td>
</tr>
<tr>
<td>Korean crisis</td>
<td>2013</td>
<td>North Korea</td>
<td>South Korea, US</td>
<td>Failure</td>
</tr>
</tbody>
</table>

Table 2: Nuclear Threat and Crisis Outcome
– Nuclear Threat Does Not Trigger Crisis

<table>
<thead>
<tr>
<th>Crisis Name</th>
<th>Years</th>
<th>Challenger</th>
<th>Target</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean War</td>
<td>1953</td>
<td>US</td>
<td>China, North Korea</td>
<td>Success</td>
</tr>
<tr>
<td>Indochina War</td>
<td>1954</td>
<td>US</td>
<td>China, USSR, Viet Minh</td>
<td>Success</td>
</tr>
<tr>
<td>First Taiwan Strait crisis</td>
<td>1954-5</td>
<td>US</td>
<td>China</td>
<td>Success</td>
</tr>
<tr>
<td>Suez crisis</td>
<td>1956</td>
<td>USSR</td>
<td>France, UK</td>
<td>Success</td>
</tr>
<tr>
<td>Second Taiwan Strait crisis</td>
<td>1958</td>
<td>US</td>
<td>China</td>
<td>Success</td>
</tr>
<tr>
<td>Cuban missile crisis</td>
<td>1962</td>
<td>US</td>
<td>USSR</td>
<td>Success</td>
</tr>
<tr>
<td>Seizure of USS Pueblo</td>
<td>1968</td>
<td>US</td>
<td>North Korea</td>
<td>Failure</td>
</tr>
<tr>
<td>Sino-Soviet border crisis</td>
<td>1969</td>
<td>USSR</td>
<td>China</td>
<td>Success</td>
</tr>
<tr>
<td>Bangladesh War</td>
<td>1971</td>
<td>US</td>
<td>India</td>
<td>Success</td>
</tr>
<tr>
<td>Yom Kippur War</td>
<td>1973</td>
<td>US</td>
<td>Egypt, USSR</td>
<td>Success</td>
</tr>
<tr>
<td>Falklands War</td>
<td>1982</td>
<td>UK</td>
<td>Argentina</td>
<td>Failure</td>
</tr>
<tr>
<td>Third Taiwan Strait crisis</td>
<td>1995-6</td>
<td>China</td>
<td>Taiwan, US</td>
<td>Failure</td>
</tr>
<tr>
<td>Kargil War</td>
<td>1999</td>
<td>Pakistan</td>
<td>India</td>
<td>Failure</td>
</tr>
<tr>
<td>Indo-Pakistani border crisis</td>
<td>2001-2</td>
<td>India</td>
<td>Pakistan</td>
<td>Failure</td>
</tr>
</tbody>
</table>