2023

The Spread of Nuclear Technology: The Difficulty of Arms Control and the Consequences of Nuclear Proliferation

Bryce Allen Beschorner
University of Kentucky, bryce.beschorner16@gmail.com
Author ORCID Identifier:
https://orcid.org/0009-0007-9167-4738
Digital Object Identifier: https://doi.org/10.13023/etd.2023.180

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Bryce Allen Beschorner, Student

Dr. Jesse Johnson, Major Professor

Dr. Michael Zillis, Director of Graduate Studies

THE SPREAD OF NUCLEAR TECHNOLOGY: THE DIFFICULTY OF ARMS CONTROL AND THE CONSEQUENCES OF NUCLEAR TECHNOLOGY PROLIFERATION

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Arts and Sciences
at the University of Kentucky

By
Bryce Allen Beschorner
Lexington, Kentucky
Director: Dr. Jesse Johnson, Associate Professor of Political Science
Lexington, Kentucky
2023

https://orcid.org/0009-0007-9167-4738

ABSTRACT OF DISSERTATION

THE SPREAD OF NUCLEAR TECHNOLOGY: THE DIFFICULTY OF ARMS CONTROL AND THE CONSEQUENCE OF NUCLEAR TECHNOLOGY PROLIFERATION

The potential impact of nuclear technology on international relations raises a number of important questions for scholars and policymakers. This dissertation focuses on different aspects of nuclear-related cooperation and competition and attempts to answer some of these questions. In this dissertation, I address three main puzzles related to nuclear politics using a mixed methods approach. First, I examine the consequences of the spread of nuclear technology on the outcomes of international crises, specifically focusing on the impact of nuclear weapons tests on crisis outcomes. Using data on nuclear weapons tests, I argue and find support for the notion that nuclear powers are hesitant to ban nuclear weapons tests because they use nuclear weapons tests as part of a brinkmanship strategy during international crises, with the winner being determined by which state has greater levels or resolve.

Next, I begin to examine why it is difficult for the international community to control the spread of nuclear technology. Focusing on international arms control efforts, I ask what prevents states from designing strong arms control agreements. Using both statistical analysis and case studies, I find that different sources of threat lead to different design choices. This is especially the case when the threats originate from states that are not part of the negotiation process, states choose to design flexible, less constraining agreements. Finally, I again examine on why it is difficult to control the spread of nuclear technology, but this time focus on the efforts of the United States. Previous research demonstrates that the United States uses a mix of positive and negative inducements in order to get states to reverse or slow down their nuclear programs, with the most effective approach being a dual-track approach that uses both positive and negative inducements. What is unclear is why the United States does not always pursue the successful dual-track approach to nuclear counterproliferation. I argue and present evidence that the United States' counterproliferation policy is constrained by the interaction of domestic politics and the broader international strategic environment.

KEYWORDS: Nuclear Politics, Arms Control, Foreign Policy, International Conflict

Bryce Allen Beschorner

(Name of Student)

04/21/2023

Date

THE SPREAD OF NUCLEAR TECHNOLOGY: THE DIFFICULTY OF ARMS CONTROL AND THE CONSEQUENCE OF NUCLEAR TECHNOLOGY PROLIFERATION

By
Bryce Allen Beschorner

Jesse Johnson
Director of Dissertation

Michael Zillis
Director of Graduate Studies

04/21/2023
Date

ACKNOWLEDGMENTS

The following dissertation, while an individual work, benefited from the insights and direction of several people. First, my Dissertation Chair, Jesse Johnson, exemplifies the high-quality scholarship to which I aspire. In addition, Jesse provided timely and instructive comments and evaluation at every stage of the dissertation process, allowing me to complete this project. I could not have asked for a better mentor and advisor over the last six years and am going to miss our weekly interactions. Next, I wish to thank the complete Dissertation Committee, and outside reader, respectively: Jill Haglund, Dan Morey, Robyn Brown, and J.S. Butler. Each individual provided insights and taught me the skills that I need to complete this project.

In addition to the technical and instrumental assistance above, I received equally important assistance from family and friends. My girlfriend, Tabitha Hawk, provided on-going support throughout the dissertation process, that was crucial for me in finishing this project. My good friend, Pete Lynch, helped make living in Lexington fun the past few years and helped me make it through graduate school. Finally, I wish to thank my parents, my sister, and my best friends. Without you guys, none of this would be possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER1. INTRODUCTION

Nuclear technology has the potential to change the world or to destroy it. Widespread cooperation between states on nuclear technology may be one of the most important weapons the world has in the battle against climate change, as this technology can be used to generate electricity in a process that is better for the environment than the burning of fossil fuels. At the same time, the spread of nuclear technology can potentially lead to widespread conflict and destruction as some states may seek to use nuclear technology to develop their own weapons, while other states seek to stop them through military action. Even with nuclear technology playing a potentially dangerous role in international politics, the international community has struggled to effectively control the spread of nuclear technology and limit states' ability to acquire it. While most states who have acquired nuclear technology have not proceeded to build nuclear weapons, notable exceptions, such as North Korea, illustrate the need to better understand the consequences of the spread of nuclear technology on international politics and why the international community and individual states, like the United States, struggle to control this technology.

The spread of nuclear technology raises a number of interesting and important questions regarding the relationship between nuclear technology and international conflict and cooperation for both scholars and policymakers alike. Answers to questions such as "what are the consequences of nuclear proliferation?" and "what prevents states from designing strong arms control agreements?", or "why in some instances does the United States struggle to control the spread of nuclear weapons?" are important for policymakers interested in creating a safer world through better control over the spread of nuclear technology. At the same time, answers to these questions contribute to scholars' under-

standing of nuclear politics and could lead to additional research programs. In the section below, I'll discuss how the rest of the dissertation is organized in more detail.

1.1 Controlling the spread of nuclear technology

Since the United States tested the world's first nuclear weapon in the "Trinity" test, a total of 2081 nuclear weapon tests have been conducted between 1945 and 2017 (Cha 2013; Johnston 2009; Nagdy and Roser 2013; Nitkin 2021; U.S. Department of Energy 2015). The states who have successfully detonated nuclear devices include all five permanent members of the United Nations Security Council, as well as regional powers like India, Pakistan, South Africa, and Israel and most recently, the rogue regime of North Korea. As the pace of testing nuclear devices began to accelerate during the early years of the Cold War, international activists, non-nuclear states, and some policymakers in the nuclear powers began to call for an international ban on the testing of nuclear devices, citing concerns ranging from the impact on the environment, to negative physical health effects (Gerl 2014), to the political effects of "nuclear blackmail," and above all else, the fear of nuclear war.

In 1958, in response to wide-scale nuclear weapons testing, the world's three leading nuclear powers at the time (United States, United Kingdom, and the Soviet Union) began negotiations over a treaty that would limit their ability to test nuclear devices. After five years of negotiation, the three states agreed to one of the first successfully-negotiated arms control agreements of the Cold War-era, with the treaty entering into force on October 10th, 1963. The resulting treaty, the Treaty Banning Nuclear Weapons Tests in the Atmosphere, in Outer Space, and Under Water (also known as the Limited Test Ban Treaty), banned the testing of nuclear devices in the atmosphere, outer space,

and underwater (LTBT Treaty). Unfortunately, the LTBT treaty did not end all nuclear tests. The signatories of the treaty could continue to test nuclear devices underground, which the United States, United Kingdom, and Soviet Union all continued to do, while France and China, opted not to sign the treaty and continued to conduct atmospheric nuclear tests until 1974 and 1980, respectively (Johnston 2009; Roser and Nagdy 2013).

The Limited Test Ban Treaty and other subsequent arms control agreements, such as the Nuclear Nonproliferation Treaty (NPT), did not prevent other states, such as India, Israel, or North Korea from testing nuclear weapons and eventually joining the nuclear club. These failures to prevent other states from joining the nuclear club created a much more uncertain and dangerous international environment with increased levels of risk for accidents, misunderstandings, and miscommunication that could lead to a nuclear conflict. Nor did intense counterproliferation efforts by the United States prevent these states from developing their own nuclear weapons. The United States has historically been the leading state in the international counterproliferation regime because of its' preeminent economic and military position in the international system. The United States has the most resources to use in the fight against nuclear proliferation, and arguably the most to lose when other states acquire nuclear technology, as it can erode their military and economic advantages. Understanding the determinants of U.S. counterproliferation policy is an important part in understanding why controlling nuclear technology is so difficult.

These developments led to a flurry of scholarship in the academic and policy making communities during the Cold War, with two broad camps developing: nuclear optimists and nuclear pessimists (Gartzke and Jo 2009). Building off deterrence theory, nuclear optimists see nuclear weapons as a stabilizing force in the international system

(Mearsheimer 1984; Waltz 1990), while nuclear pessimists view nuclear weapons as destabilizing forces that incite fear, conflict, and misperceptions (Jervis 1984; Sagan 1989). A lot of the post-Cold War scholarship has often staked out a middle ground, arguing that elements of both the optimist and pessimists' positions are both right and wrong. This dissertation falls in that middle ground, as nuclear proliferation does not automatically lead to nuclear war, but demonstrating that there are some escalatory dangers and consequences to nuclear proliferation.

Overall, the struggle that both the international community and powerful states, such as the United States, face in trying to effectively control the spread of such a pivotal and potentially dangerous technology illustrates how difficult it is to control the proliferation of nuclear technology. The historical process that was described above, from unconstrained nuclear weapons testing in the 1950s to the attempts by both the international community and the United States to design arms control agreements and counterproliferation policies to better control the spread of nuclear technology in the following decades, serves as a motivating template for this research program. In this dissertation, I use each empirical chapter to address a puzzle related to nuclear politics and the historical process laid out above.

1.2 Dissertation Outline

In chapter 2 I answer "what are the consequences of nuclear proliferation?" One of the consequences of nuclear proliferation that I identify is that states who acquire nuclear weapons tend to test these weapons a lot, including during crises with other states. This widescale testing is puzzling because historical evidence indicates that states have tested more weapons than would be necessary if they were only concerned with improv-

ing or maintaining their current nuclear capabilities. Additionally, there has been more than one major arms control agreement that has attempted to regulate or outright ban nuclear weapons tests. It is clear there has been a longstanding international appetite for stopping the testing of nuclear weapons, yet states have conducted over 2000 tests since 1945. To address this puzzle, I use data on state nuclear weapons tests that has not featured in political science research before as well as a case study of the 1969 Soviet-Chinese Border Crisis. I argue and find support for the notion that nuclear powers have conducted so many nuclear weapons tests and are hesitant to ban these tests because they use nuclear weapons tests as part of a brinkmanship strategy during international crises, with the winner being determined by which state has greater levels or resolve. States use nuclear weapons tests to reveal their levels of resolve to their crisis adversaries, and so, the state who conducts more nuclear weapons tests is more resolved and more likely to win the crisis. This research builds off the existing theoretical and empirical brinkmanship literatures (Schelling 1966; Powell 2003; Kroenig 2018), and captures the dynamic, provocative actions that are at the heart of a brinkmanship strategy.

Chapter 3 examines what prevents states from designing strong arms control agreements. This is puzzling because states spend so much time and resources negotiating these agreements, only for some of them to feature weak designs that will struggle to effectively control the spread of nuclear and other related weapons technologies. To address this puzzle, I build off some of the existing arms control and international institutional design literature (Kreps 2018; Koremenos 2005). The design aspect of the agreements is important because the design of these agreements determines how effective they will be, with stronger, more constraining, and less flexible agreements most likely doing

a better job of regulating state behavior when it comes to nuclear technology. Using both statistical analyses and two case studies, I find that different sources of threat lead to different design choices, especially when the threats originate from states that are not part of the negotiation process. When threats originate from states that are not part of the arms control agreement negotiation process, states choose to design weaker, more flexible, and less constraining agreements. By addressing threats originating from states that are not part of the negotiation process, policymakers will be able to form more effective arms control agreements and thus, create a safer world.

In chapter 4, I once again examine why it is difficult to control the spread of nuclear technology but this time focus on the policies of the leading state in the international counterproliferation regime, the United States. I build off previous research that demonstrates that the most effective counterproliferation strategy the United States has in convincing potential proliferators to forego nuclear weapons development, is using a mix of positive (civilian nuclear cooperation, economic aid, military aid, and political aid) and negative (sanctions and the threat of military force) inducements in what is called the dual-track approach (Mehta 2020). What is unclear is why the United States does not always pursue a dual-track approach to nuclear counterproliferation, as this effective approach is rarely used.

I argue that United States' counterproliferation policy is constrained by the interaction of domestic politics and the broader international strategic environment. At the domestic level, the president's ability to overcome congressional barriers is key to getting a dual-track approach implemented, with the level of legislative concurrence between the president and Congress being a key indicator of this ability. At the international level,

presidents are constrained by the potential proliferator's relationship with the United States. U.S. presidents are unlikely to want to use the negative inducements that are included in a dual-track approach against a potential proliferator with whom the U.S. has a friendly relationship with. Using a statistical analysis, I find that the dual-track approach is most likely to be implemented when there are high levels of concurrence between the president and Congress and the potential proliferator has had a MID with the United States in the past few years. Understanding under what conditions the United States is able to implement a dual-track approach will allow policymakers to better overcome the domestic and international obstacles that prevent this approach from being used more often, and thus, increase the probability of nuclear reversal.

Finally, chapter 5 is used for concluding remarks. I argue that this dissertation contributes to scholarly understanding of nuclear politics, arms control, U.S. foreign policy, and international conflict and cooperation, more broadly. Additionally, I discuss how future research could build off the findings in this project. Finally, understanding the consequences of failing to control the spread of nuclear technology and why it is so difficult to control its' spread, will provide policymakers with the information they need to design more effective arms control agreements and counterproliferation policies, in order to better control nuclear technology's spread and create a safer world.

CHAPTER 2. TESTING FOR VICTORY: NUCLEAR WEAPONS TESTS, BRINKMANSHIP, AND NUCLEAR CRISIS OUTCOMES

2.1 Introduction

On 30 October 1961, the Soviet Union detonated the most powerful weapon in the history of mankind. The nuclear weapon later nicknamed the "Tsar Bomba," had an explosive yield of fifty megatons[1] (equivalent to fifty million tons of TNT) and immediately generated intense global backlash against the Soviet Union and the testing of nuclear weapons (Wellerstein 2021). While the explosive yield of the Tsar Bomba makes it unique among nuclear weapons tests, the international community's response to the test and the Soviet rationale for conducting the test was not unique. An obvious reason why the Soviets decided to test a nuclear weapon so powerful, or any nuclear weapons for that matter, is to gather data to maintain and improve existing weapon systems or develop new ones. However, historical evidence indicates that the global number of nuclear weapons tests exceeds the number of weapons tests states would have conducted if they only cared about using these tests to gather data. This raises the following question: why have states conducted so many nuclear weapons tests?

I argue that nuclear weapons tests serve a dual-purpose for the state carrying out the tests. Nuclear weapons tests allow states to learn about and develop new capabilities but can also be a part of a brinkmanship strategy for crisis participants. In brinkmanship, states exert pressure on their adversaries by taking steps and carrying out provocative actions that raise the risk that events will accidentally spiral out of control and lead to an exchange of nuclear weapons (Schelling 1966). According to brinkmanship then, the

[1] The Tsar Bomba had a potential yield of 100 megatons but was only tested at 50 megatons because of concerns about the number of radioactive particles released in the atmosphere that would result from a 100-megaton detonation (Wellerstein 2021).

state most willing to take provocative actions that "leave something to chance" and run the risk of nuclear war is most likely to emerge victorious from a nuclear crisis, with the level of risk-acceptance of each crisis participant directly determined by their level of resolve (Powell 2003; Schelling 1966). I argue that during a nuclear crisis, states use nuclear weapons tests to reveal the balance of resolve between themselves and their adversary, as well as manipulate the shared risk of war by pushing crisis opponents closer to war, with the winner of the crisis being determined by which state conducts more nuclear weapons tests. The Tsar Bomba test, for example, was one of several tests that the Soviet Union conducted during the Berlin Wall Crisis, a crisis in which they conducted more nuclear weapons tests than the United States and won.

While the brinkmanship model has existed for almost six decades and received intense theoretical examination through formal models and qualitative studies (e.g., Jervis 1989; Powell 1988; Powell 1990; Powell 2003; Schelling 1966), the empirical literature concerning brinkmanship is lacking for such an important model.[2] This project expands the existing empirical literature of brinkmanship, presents a novel theoretical and empirical analysis of nuclear weapons tests, while also offering a more direct empirical test of the brinkmanship model than past studies. Additionally, this project adds to the study of crisis outcomes by examining how a previously overlooked factor, nuclear weapons tests, influences crisis outcomes.

The paper proceeds as follows: I begin with a general discussion of nuclear weapons tests before introducing a theory that fits nuclear weapons tests within the brinkmanship model. I then test my theory on all crisis dyads between 1945 and 2016 where both states in the dyad possess nuclear weapons. I find that states that conduct more nuclear

[2] For a notable exception see Kroenig (2013).

weapons tests than their opponent during a crisis are more likely to win the crisis, while also showing that the possession of nuclear weapons alone is not enough to achieve victory in a crisis. I conclude with a discussion of the important implications these results have for policy-making and future academic research.

2.2 The basics of nuclear weapons tests

Since the United States tested the world's first nuclear weapon in the "Trinity" test, a total of 2081 nuclear weapon tests have been conducted between 1945 and 2017 (Cha 2013; Johnston 2009; Nagdy and Roser 2013; Nitkin 2021; U.S. Department of Energy 2015). This includes 2052 tests conducted by single states, as well as 28 joint United States and United Kingdom tests and 1 unconfirmed test that is thought to be a joint South African-Israeli test.[3] According to the 1974 Threshold Test Ban Treaty (TTBT), a nuclear weapons test is defined as "either a single underground nuclear explosion conducted at a test site, or two or more underground nuclear explosions conducted at a test site within an area delineated by a circle having a diameter of two kilometers and conducted within a total period of time of 0.1 second" (Threshold Test Ban Treaty 1974: section 1.2).

Two parts of this definition need clarification. First, this definition only mentions underground nuclear weapons tests because atmospheric testing was banned with the signing of the Partial Test Ban Treaty (PTBT) in 1963. Not all nuclear powers immediately signed onto the PTBT, and some, like France and China, continued to test nuclear weapons in the atmosphere until 1974 and 1980, respectively (Johnston 2009).

[3] This is the infamous *Vela Incident*, which is thought to have most likely been a joint South African-Israeli nuclear weapons test in the south Atlantic Ocean in 1979 (Krzyzaniak 2019).

The count of total nuclear weapons tests presented here includes any and all nuclear weapons tests, including tests in the atmosphere, underground, and underwater. The second part of the nuclear weapons test defintion that needs clarification is the second clause describing "two or more underground nuclear explosions" only counting as one test. This was adopted because by 1974, the United States and Soviet Union had increasingly begun to use salvo shot testing. This is a form of nuclear weapons test in which the testing state detonates multiple nuclear weapons devices simultaneously at the same testing site, but in separate testing holes. Even though multiple nuclear devices were being detonated, both states decided that salvo shots would only count as a single test. This part of the definition is reflected in the total count of nuclear weapons tests, with salvo shot tests being counted as a single test. The following states have conducted nuclear weapons tests, with the year of the first test in parentheses: United States (1945), Soviet Union/Russia (1949), United Kingdom (1952), France (1960), China (1964), India (1974), Israel[4] (1979), South Africa[5] (1979), Pakistan (1998), and North Korea (2006). All of these states still possess nuclear weapons, with the exception of South Africa, who gave their weapons up in 1990 (de Villiers, Jardine, and Reiss 1993). Table 2.1 (below) gives the total number of tests for each testing state.

2.2.1 The dual-purpose of nuclear weapons tests

The primary reason states test nuclear weapons is to gather data in order to maintain and improve weapons systems or develop new capabilities. However, as previously stated, states have conducted more nuclear weapons tests than would be necessary if they were solely interested in using the tests to gather data. For example,

[4] This would be the *Vela Incident*.
[5] This would be the *Vela Incident*.

before the Tsar Bomba test, the lead Soviet scientist on the project, Andrei Sakharaov, stated that he believed that if the test was successful then the Soviet Union would not need to conduct any more tests in the atmoshpere (Wellerstein 2021). He reiterated this belief after the test was successfully conducted, but the Soviet Union went on to conduct over 100 more atmospheric nuclear weapons tests in the next few months before the PTBT entered into force. These additional tests, as well as other historical evidence (Carr 2020; Wellerstein 2021), demonstrates that there is a secondary reason why states test nuclear weapons: to signal to their adversaries as part of a brinkmanship strategy during a crisis. These two reasons for testing nuclear weapons are not in competition with each other, as a single test (like the Tsar Bomba) can serve both purposes. Additionally, since other states cannot easily discern the purpose of a single nuclear weapons test, we must treat each test as being theoretically and empirically the same.

Table 1[6] (below) lists the total number of nuclear weapons tests conducted by each testing state between 1945 and 2017 and also how many of those tests took place within international crises[7] in which the testing state was a participant. 21.7%[8] of all nuclear weapons tests[9], excluding joint tests, took place during crises. This ranges from 6.7% of all Chinese nuclear weapons tests to 100% of Pakistan's tests. Since over a fifth of all nulear tests taking place during crises, this suggests that testing may be important for crisis participants and their goals. As I will explain in later sections of the paper, nuclear weapons tests can be an effective part of a state's brinkmanship strategy, with

[6] The United States has conducted half of all single-state nuclear weapons tests, but the United States does not drive the results presented in the main analysis. See Table 12 in Appendix 1.
[7] Data on international crises comes from the International Crisis Behavior (ICB) project (Hewitt 2003).
[8] This decreases to 9.9% of all tests when focusing on crises in which both participants are nuclear powers.
[9] I only count tests that take place between the day after the crisis officially starts and the day before the crisis officially ends. This ensures that each test that is counted takes place clearly within the crisis window.

states that conduct more tests than their adversary having a much greater chance of victory. I now briefly discuss brinkmanship.

Table 1: Global nuclear weapons tests, 1945-2017

State	Number of nuclear weapons tests	Number of tests during all crises
United States	1026	216 (21.1%)
Soviet Union/Russia	715	186 (26.0%)
France	210	29 (13.8%)
China	45	3 (6.7%)
United Kingdom	45	5 (11.1%)
North Korea	6	3 (50%)
India	3	1 (33.3%)
Pakistan	2	2 (100%)
Joint U.S.-UK	(28)	--
Joint South Africa-Israel	(1)	--
Total	2052 (2081)	445 (21.7%)

Sources: Cha (2013); Johnston (2009); Nagdy and Roser (2013); Nitkin (2021); U.S. Department of Energy (2015).

2.3 Brinkmanship and crisis outcomes

The dawn of the nuclear age in the aftermath of World War II dramatically changed the nature of interstate disputes, especially for the few states who possessed nuclear weapons. Scholars have argued that due to the increased costs of war brought by nuclear weapons, states are deterred from engaging in direct military combat with one another, leading to a change in the form of interstate disputes among nuclear powers from interstate wars to international crises (Hoffman 1965). Additionally, as the form of

interstate disputes among nuclear powers changed from war to crises, the nature of crises have shifted along with it, moving from competitions in military capabilities to "competitions in risk-taking" (Schelling 1966). This means that the outcome of interstate disputes, when both states possess nuclear weapons, is not going to be determined by which state has greater military capabilities but instead, by which state is the most resolved and risk-accepting when it comes to the likelihood of the dispute escalating to nuclear war. States reveal their levels of resolve and take part in this "competition in risk-taking" by using a strategy of brinkmanship.

Brinkmanship is a strategy that attempts to gain an advantageous outcome for a state by exerting coercive pressure on its' adversary through taking steps and carrying out provocative actions that raise the risk that events will accidentally spiral out of control and lead to a nuclear exchange (Powell 2003; Schelling 1966). More simply put, brinkmanship is about manipulating the shared risk of nuclear war to figure out which of the crisis participants is more resolved (Schelling 1966: 99). According to the brinkmanship model then, there are 2 ways to win a crisis: first, states can take actions that reveal their relative levels of resolve to their adversary and remove the uncertainty over resolve that is underlying the crisis; or, states can take actions that bring both states closer to war, with the more risk-accepting (and resolved) state having the advantage as they are more likely to let events spiral out of control and get closest to nuclear war. As I will discuss in the following sections, some actions, like the testing of nuclear weapons, accomplish both of these tasks and can be part of an effective brinkmanship strategy.

2.3.1 Brinkmanship evidence

Numerous scholars have examined how different actions or nuclear capabilities can influence the balance of resolve and fit into a brinkmanship model.[10] Powell (2003) uses formal theory to examine how a national missile defense system would influence a state's level of resolve and their likelihood of victory in a nuclear crisis. Powell argues that an effective national missile defense system would give the United States more freedom of action to raise the level of risk in order to get an adversary to back down in a crisis because an effective national missile defense system will prevent some or all of the adversary's missiles from striking the United States, reducing the expected costs the United States would pay if the crisis escalates to nuclear war. Other scholars have used systematic qualitative analyses to evaluate brinkmaship. In a series of case studies, Betts (1987) examines explicit nuclear threats made by the United States and Soviet Union during the Cold War and how those threats influenced policymakers' decisionmaking and crisis outcomes. The results are inconclusive, as Betts concludes that neither the balance of resolve nor the nuclear balance between crisis participants provides a satisfactory explanation for crisis outcomes.

Kroenig (2013, 2018[11]) offers one of the first systematic empircial examinations of brinkmanship.[12] Kroenig argues and finds support for the notion that nuclear superior[13] states have a greater probability of victory in crises because of the effect that the balance

[10] Not all scholars accept the brinkmanship model and some even question the relevance of nuclear weapons in international politics. See Tannenwald's (1999) discussion of the nuclear taboo for an example. For evidence against the existence of the nuclear taboo see Sagan and Valentino (2017) or Horschig (2022).

[11] In Kroenig (2018) the author expands on the argument from Kroenig (2013) and renames it the superiority-brinkmanship synthesis theory.

[12] For a systematic empirical analysis that questions the validity of brinkmanship see Sechser and Fuhrmann (2013, 2017).

[13] Kroenig (2018) defines nuclear superiority as a military nuclear advantage over an opponent. It is operationalized according to a state's expected cost of nuclear war and considers the balance of nuclear warheads between adversaries.

of nuclear warheads has on the balance of resolve. States with more warheads than their adversary are thought to have a greater level of resolve because they will have lower expected costs if nuclear war does occur. Nuclear superior states have lower expected costs of war than their opponents because they are better able to engage in a counterforce targeting strategy that limits their opponents' ability to inflict damage on them. These lowered expected costs of nuclear war increase the nuclear superior state's level of resolve in a crisis and allows them to get closer to the "brink" in a nuclear crisis, giving them greater bargaining leverage over their crisis opponent.This gives the state with a larger nuclear arsenal the ability to be more risk-accepting than their adversary and gives them a greater probability of crisis victory.

By focusing on nuclear weapons tests states during a crisis, this study probes the brinkmanship argument further. A focus on nuclear weapons tests is an important next step in analyzing the efficacy of brinkmanship because nuclear weapons tests capture dynamic provocative actions that take place during a crisis and are at the heart of brinkmanship strategies, as opposed to more static features of the crisis actors that do not provdie new information to their opponents. To motivate the analysis, in the sect section I decribe how nuclear weapons tests can form an effective part of states' brinkmanship strategy and help them prevail in crises.

2.4 Nuclear weapons tests as brinkmanship

States have used nuclear weapons tests in order to coerce and intimidate other states into backing down during a crisis. As stated earlier, the Soviet Union's Tsar Bomba test had a dual-purpose; not only did it serve as a test for high-yield nuclear weapons, but it was also part of the Soviet Union's strategy for prevailing in the Berlin Wall Crisis of

1961. Officials who were around Soviet Premier Nikita Khruschev during the discussions surrounding the decision to test the Tsar Bomba noted that some of the motivation for the timing of the test was the Berlin Wall Crisis, which was reaching its' zenith when Tsar Bomba was detonated, and Khruschev's wish to look more resolved to the United States (Wellerstein 2021). American officials reacted harshly to the test, calling it a "diplomatic bomb" whose purpose was "intimdiation, fear and blackmail" (United Press International 1961). Clearly, officials in both the Soviet Union and the United States viewed the Tsar Bomba test as having a purpose beyond just gathering data, with both sides seeing it as part of the Soviet strategy to get the U.S. and its' allies to back down in Berlin. Additionally, the International Crisis Behavior (ICB) Project considers the Berlin Wall Crisis to be a victory for the Soviet Union (Brecher et al. 2021).

How do nuclear weapons test fit into our understanding of brinkmanship? I argue that nuclear weapons tests can form part of a state's brinkmanship strategy during a crisis in two complementary ways, each of which I will discuss in the following paragraphs. First, nuclear weapons tests reveal the balance of resolve between crisis adversaries. Second, nuclear weapons tests during a crisis between two nuclear-armed states bring both sides closer to war.

2.4.1 Nuclear weapons tests and the balance of resolve

I argue that through nuclear weapons tests states update each other about their respective levels of resolve, which helps remove uncertainty over the overall balance of resolve. Crises only occur when there is uncertainty about the balance of resolve between two adversaries (Powell 2003). Uncertainty over the balance of resolve causes crises because if states knew where the balance of resolve stood before the crisis begins, then

the less resolved state could back down and avoid the crisis entirely. Nuclear weapons tests during a crisis are informative to the testing state's opponents, because each time a state tests a nuclear weapon during a crisis, they raise the level of risk they and their opponent must pay and push the states closer to nuclear war. Additionally, because of the costs associated with nuclear weapons testing, nuclear weapons tests during a crisis serve as a form of sunk cost signalling that reveals the testing state's level of resolve. Costs associated with nuclear weapons tests include economic costs[14], politcal and reputational costs[15], health costs[16], and environmental costs.[17] The costs of nuclear weapons testing are paid ex ante and are informative in that it is assumed that states invest less in signalling when their resolve is lower because they do not value whatever the "good" or policy in dispute is enough to invest enough resources to acquire it (Fearon 1997; Gartzke et al. 2017).

During a crisis, after State A conducts a nuclear weapons test, State B needs to decide whether to raise the level of risk through their own test or back down and admit defeat. Since State B now knows that State A is somewhat resolved in the crisis because of the nuclear weapons test, some of the uncertainty around the balance of resolve between the two states is reduced. If State B decides to conduct its' own nuclear weapons test after State A's test, then both sides know that the other side has some level of resolve and they need to continue to test their nuclear weapons to remove more of the uncertainty

[14] The United States spent about $2.3 billion dollars (in 2020 dollars) during the Crossroads test series in 1946 (Schwartz 1998).

[15] After India and Pakistan exchanged nuclear weapons tests in 1998, the United States, Japan, and other states enacted sanctions on both India and Pakistan and cut off other forms of cooperation and aid (CNN 2001).

[16] A 1991 study by the International Physicians for the Prevention of Nuclear War predicted that eventually 2.4 million people could die from cancer because of atmospheric nuclear weapons testing (IPPNW 1991).

[17] Former inhabitants of Bikini Atoll, a U.S. testing site in the 1950s, and their descendants have been unable to move back to the islands because the groundwater is contaminated due to nuclear weapons tests (Scott 2017).

surrounding the balance of resolve. If State B backs down after State A's test, then State A is clearly more resolved and wins the crisis. Since states incur costs for each subsuquent nuclear weapons test, during the duration of a crisis each additional nuclear weapons test reveals more information about the testing state's resolve to their adversary. The more a state tests during a crisis, the more resolved they appear to be to their opponent because of the greater costs they must pay. This is useful because the difference in the number of nuclear weapons tests conducted by crisis opponents reveals which side is more resolved, allowing the less resolved state to back down short of nuclear war. Eventually, a state will bid up the amount of risk and incur enough costs through nuclear weapons testing to reveal that they are more resolved than their opponent, thus removing the underlying uncertainty over the balance of resolve that caused the crisis to develop in the first place. Since a state's resolve, along with uncertainty about their adversary's level of resolve, directly determines a state's level of risk-acceptance; and, since the balance of resolve is revealed through nuclear weapons testing during a crisis, the state that conducts the most nuclear weapons tests during a crisis is most likely to win.

2.4.2 Nuclear weapons tests and manipulating the risk of war

Nuclear weapons tests fulfill one of the key elements of brinkmanship in that they reveal the balance of resolve between crisis adversaries. The second way that nuclear weapons tests conducted during a crisis fit into the brinkmanship model is that they manipulate the risk of war, moving opponents closer to nuclear war. The steps-to-war framework is useful for explaining how testing nuclear weapons during a crisis moves crisis participants closer to war. The steps-to-war framework argues that war is "chosen" step-by-step, with each step increasing the probability that the next step will be taken

(Sample 2016). The core of the steps-to-war framework is the notion that territorial disputes are more likely to escalate to war when compared to other types of disputes (Senese and Vasquez 2005; Vasquez 1993). This is especially the case when states use realpolitik measures, such as forming alliances and arms racing, to attempt to win the territorial dispute by demonstrating their capability and will. The reason why these realpolitik measures make war more likely is that each subsequent step produces a security dilemma that leads each state to feel more threatened and more hostile towards the other state and causes each state to take additional steps aimed at their opponents in order to increase their own security, creating a hostile spiral (Senese and Vasquez 2005). Additionally, this vicious cycle of hostile steps solidifies the position of hardliners in domestic politics and tends to increase the number of hardliners and their influence, pushing leaders to take more realpolitik measures to deal with the threat from the opposing state (Senese and Vasquez 2005).

How does the steps-to-war approach explain why nuclear weapons tests during a crisis move states closer to nuclear war? First, in both territorial disputes and nuclear crises, neither side are completely in control of events, and there is a chance of hostilities spiraling out of control (Schelling 1966; Vasquez 1993). Not only are control of events out of the hands of participants in both territorial disputes and nuclear crises, but both situations are characterized by an overt tension and uneasiness that increase the chances of an accident or misinterpretation that leads to war.

Second, and most importantly, nuclear weapons tests can be considered a realpolitik measure. Nuclear weapons tests serve a dual-purpose that includes helping states develop new types of nuclear capabilities. Thus, it is easy to see nuclear weapons tests as a

realpolitik measure that either increase the testing state's nuclear capability, like how forming an alliance or investing in an arms race increases a state's traditional military capabilities or demonstrates a previously unrevealed nuclear capability. Like other realpolitik measures then, nuclear weapons tests create a security dilemma where the crisis opponent of the testing state feels more threatened after a nuclear weapons test because of the testing state's perceived increase in nuclear capabilities, and so, undertakes their own nuclear test to improve their nuclear capabilities and overall security. This creates a hostile spiral that can escalate to war with each subsequent nuclear weapons test then as another "step" closer to war. A state that tests can push the crisis disputants closer to nuclear war which can encourage the less resolved participant to back down short of war, leading to crisis victory for the testing superior and more resolved state.

2.4.3 Detecting nuclear weapons tests

Finally, nuclear weapons tests would not be an effective part of a brinkmanship strategy if crisis adversaries could not detect and confirm each other's tests. The ability to detect other states' nuclear weapons tests are necessary because leaders have an incentive to misrepresent their resolve in a crisis in order to try and get the other state to back down. This means leaders may purposely lie about conducting nuclear weapons tests during a crisis in order to appear more resolved. These detection methods allow other states to confirm that their crisis opponents have actually conducted nuclear weapons tests.

There is a fine line with how much information states want to reveal with their nuclear weapons tests. On the one hand, the tests will not be an effective brinkmanship signal if other states cannot detect them. On the other hand, states are wary about

revealing classified information about their nuclear programs, such as specific weapon design, that may give their crisis adversaries an advantage. States sometimes make it easy for their adversaries by announcing tests ahead of time, like with the Tsar Bomba test (Wellerstein 2021), or immediately confirm their nuclear weapons tests after they have been conducted, like with the series of Indian nuclear weapons tests in 1998. If the testing state's leaders do not publicly confirm that they have conducted a nuclear weapons tests, states have numerous methods of detection, including espionage, "sniffer" planes[18], national technical means[19], and the Comprehensive Test Ban Treaty Organization's (CTBTO) international monitoring system.[20]

Overall, nuclear weapons tests can be an effective part of a state's brinkmanship strategy during a crisis. Nuclear weapons tests fulfill the two main processes associated with brinkmanship: revealing the balance of resolve and removing the uncertainty surrounding it and moving crisis opponents closer to war. This leads to the following hypothesis:

Nuclear Testing Superiority hypothesis: During a crisis, states that conduct more nuclear weapons tests than their adversaries will be more likely to win that crisis.

2.5 Research design

To test my theory of nuclear crisis outcomes, I use the International Crisis Behavior Project's (ICB) dyadic-level data set (Hewitt 2003). In combination with the ICB's system-level and actor-level data sets, the ICB dyadic-level data set provides plenty of information on crisis outcomes, the trigger and termination dates of crises, and state-level

[18] See Rhodes (1995: 368-372) for a description.
[19] National technical means include spy satellites and spy planes.
[20] The CTBTO's monitoring system includes seismic, hydroacoustic, infrasound, and radionuclide monitors all around the world and can detect tests that take place underground, underwater, in outer space, or in the atmosphere (CTBTO Preparatory Commission 2012).

characteristics related to crisis outcomes. The ICB defines a crisis as an interstate dispute that threatens at least one state's values, has a heightened probability of military escalation, and has a finite time frame for resolution (Brecher and Wilkenfield 2000). Additionally, according to Schelling, "it is the essence of a crisis that the participants are not fully in control of events" (Schelling 1966). As previously mentioned, a nuclear crisis can occur whether nuclear weapons are used, are threatened, or are the disputed issue in the crisis (Kroenig 2013).

The temporal domain for this analysis is from 1945 to 2016. The unit of analysis is the dyad crisis, as the main independent variable, nuclear testing superiority, can only be measured at the dyadic-level. I employ directed-dyads because the dependent variable, crisis outcomes, varies for each state in the dyad. Another difference between my analysis and the existing literature, is that I only include two-sided crises in my sample, while other authors (Kroenig 2013) include both one-sided and two-sided crises.[21] This leaves me with a sample of 62 crisis-dyads across 24 different crises[22] in which both states possess nuclear weapons[23] at the time of the crisis.

2.5.1 Dependent variable: Crisis outcomes

The dependent variable in this analysis is a binary variable called *Victory. Victory* measures whether a country achieves victory in a crisis and is drawn from the variable *OUTCOM* from the ICB's actor-level data set. The original ICB variable is an ordinal

[21] This is necessary because the ICB actor-level data set, from which I derive my dependent variable and many of the controls, only includes information on states who are considered crisis actors. If I included one-sided crises in my sample, I would be left with missing data for at least one state in each of one-sided crises.

[22] For a list of the nuclear crises analyzed and the victors of each crisis see Table 9 in Appendix 1.

[23] To determine when each state first possesses nuclear weapons, I use ICB's variable *Nuclear*. I code states as possessing nuclear weapons if they score a '3' or above, except for Israel who I recode as possessing nuclear weapons from 1970 on (Nuclear Threat Initiative 2014).

indicator that codes whether a state achieves a victory, a compromise, a stalemate, or is defeated in a crisis. Following past research (Beardsley and Asal 2009, Kroenig 2013), I dichotomize this variable to code whether a state achieves a victory in a crisis or not. *Victory* equals 1 when a state wins a crisis and is coded as a 0 for all other types of outcomes. The ICB defines a victory as a crisis in which an actor achieves its "basic goals" (Brecher et al. 2021). There can be multiple victors in a crisis, as well as cases where no clear victor is produced. Of the 62 crisis-dyads in the sample, a clear victory is achieved in 17 of them. Since the dependent variable is binary, I use probit models throughout. I also employ robust standard errors that are adjusted for clustering by crisis dyad to account for the interdependence of observations.

2.5.2 Main independent variables: nuclear testing superiority and nuclear testing ratio

The two main independent variables are *Test Superiority* and *Test Ratio.* Both *Test Superiority* and *Test Ratio* are used to evaluate the nuclear testing superiority hypothesis. *Test Superiority* is a binary variable that measures whether State A in each crisis-dyad conducted more nuclear weapons tests during the duration of the crisis than the other state in the dyad. To begin to construct this variable, I first counted the number of nuclear weapons tests that each state in the dyad conducted, beginning with the day after the crisis begins and ending with the day before the crisis ends. This ensures that each nuclear weapons test that is counted clearly occurs after the crisis begins and before it is resolved. I do not include nuclear weapons tests in which the sources disagree about whether it took place or not[24] (denoted by an "X" in Johnston's Archive). I only count

[24] The exception to this rule is France. Johnston's Archive questions the July 21st, 1966 French test codenamed "Ganymede", but other sources, including the U.S. Department of Energy (2015) and Kimball

tests that were conducted by a single state, since joint tests make it difficult to attribute the test, and the subsequent signals it sends to a single state. As previously discussed, I also count salvo shot tests as a single test, even though multiple nuclear devices are detonated during these tests. There are 14 instances (out of 62) where State A conducts more nuclear weapons tests than State B during a crisis. *Test ratio* is used to examine whether states that enjoy greater levels of nuclear weapons testing superiority are more likely to win crises. *Test ratio* is calculated as the number of nuclear weapons tests conducted by State A during the crisis divided by the total number of tests conducted by both State A and State B during the crisis. This variable can theoretically range from 0 to 1.

Data on nuclear weapons tests comes from multiple sources. For the United States, I rely on a list of tests provided by the U.S. Department of Energy (DOE) (2015). Unfortunately, similar government reports are not readily available for the other testing states. For the other testing states, I mostly rely on Johnston's Archive (Johnston 2009). To verify the accuracy of Johnston's Archive, I compare the list of United States' tests provided by the U.S. DOE and Johnston's Archive. The only differences between the two sources are that Johnston's Archive counts salvo shot tests as multiple tests instead of a single test and the DOE's list includes the 28 joint U.S.-UK tests. Once these differences are accounted for, Johnston's Archive lists the exact same number of tests, under the same codenames, as the DOE's list for the United States. This gives me confidence that Johnston's Archive is an accurate source for all nuclear weapons tests. Additionally, once salvo shot tests are accounted for, Johnston's Archive lists the exact same number of total tests for every testing state (through 2009) as other sources like the U.S. DOE (2015) re-

(2020) provide evidence that France conducted 210 nuclear weapons tests and so I count this test as being part of the French testing regime.

port and Kimball (2020). Johnston's Archive uses other sources to compile their nuclear weapons testing data.[25] I also used information from the Center for Strategic and International Studies (Cha 2013) and Congressional Research Service (Nitkin 2021) to fill in missing data related to North Korea's post-2009 tests since Johnston's Archive ends in 2009.

2.5.3 Control variables

I control for several additional factors that may influence crisis outcomes, following previous studies (Beardsley and Asal 2009; Kroenig 2013). I control for whether a state enjoys nuclear warhead superiority over the other state in the crisis-dyad, as Kroenig (2013) finds that states that possess more nuclear warheads than their adversary are more likely to prevail in crises. I first gathered data on each states' total number of nuclear warheads each year from Our World in Data (Roser and Nagdy 2013). I then created a binary variable, *Warhead Superiority*[26], which equals 1 when State A possesses more nuclear warheads than State B during a crisis.

Next, I control for whether violence was used by either side during a crisis, as high levels of violence during a crisis may be more likely to produce a clear winner. I use the *Violence* (called *VIOL* by the ICB) variable from the ICB system-level data set. *Violence* is a four-point ordinal variable that ranges from 1 (No violence) to 4 (Full-scale war).

I also include two variables that control for the stakes of the crisis. The first, *Gravity*, measures the relative gravity of the crisis for the actors involved. The graver the threat a crisis is to an actor, the more likely they are to push harder for victory in the cri-

[25] A list of sources used by Johnston's Archive can be found on Johnston's Archive's website.
[26] The correlation between *Nuclear Superiority* and *Test Superiority* is 0.0946 in the main sample.

sis when compared to other actors who may not be facing as grave a threat. *Gravity* uses data from the ICB actor-level data set that codes the severity of the crisis for each actor involved, from 0 (Economic threat) to 6 (Threat to existence).[27] *Gravity* equals 1 when State A is facing a graver threat than State B. The second control variable I include to control for the stakes of the crisis is the binary variable *Proximity. Proximity* measures whether the crisis took place closer to State A or State B. The closer to home a crisis is for an actor, the greater the stake the actor has in the crisis and the outcome, and the harder they are likely to push in the crisis. *Proximity* uses data from the ICB actor-level data set that codes the location of a crisis relative to each actor and ranges from 1 (crisis took place in the actor's home territory) to 4 (elsewhere).[28] *Proximity* equals 1 when the crisis takes place closer to State A than State B.

Additionally, I control for the traditional military capabilities in the crisis dyad, using the Correlates of War's (COW) National Material Capabilities data set (Singer et al. 1972). It is possible that states with greater traditional military capabilities are more likely to win crises than their weaker foes. I create a variable *Capabilities* that equals the capabilities of State A divided by the total combined capabilities of both State A and B. Each state's military capabilities are assessed using CINC scores.

I also include the variable, *Population*, that measures State A's population in a given year. States with larger populations may be better able to absorb a nuclear attack and remain intact than states with smaller populations, and thus, may be more likely to

[27] Called *GRAVTY* in the ICB actor-level data set. I recode all cases where the ICB data indicates the gravest threat to be "other" (GRAVTY=7) to equal -1, with the assumption being that if there is not a single, clear threat to a crisis actor, then the crisis is probably less grave for them than other crisis actors facing threats coded from 0-6.

[28] Called *CRACTLOC* in the ICB actor-level data set.

push harder in a crisis. I take the natural log of the population data from COW's National Material Capabilities data set to create *Population* (Singer et al. 1972).

The domestic regime type of each actor in a crisis may play a role in influencing crisis outcomes, as it has been argued that democracies are more likely to win crises because they select into crises that they are more likely to win (Gelpi and Griesdorf 2001). I create the variable *Democracy,* that measures whether State A is a democracy. Data on domestic regime types comes from ICB's actor-level data set, which contains an ordinal variable[29] that categorizes each crisis actor's domestic regime during the crisis. *Democracy* equals 1 when State A is a democracy during a crisis.

I also control for whether crisis participants possess second-strike nuclear capabilities. A state that possesses second-strike capabilities may be able to better resist nuclear coercion, as they could theoretically absorb a first strike and still have enough surviving nuclear forces for a retaliatory strike. Drawing once again on the ICB actor-level data set, which includes a variable measuring crisis actors' nuclear capability that ranges from 1 (no foreseeable nuclear capability) to 4 (developed nuclear capability with second strike capability)[30], I create *Second Strike*, which equals 1 when State A in the dyad possesses second-strike capabilities during a crisis.

Finally, I control for a state's security environment. A competitive security environment may influence both a state's likelihood to develop and possess nuclear weapons (and thus conduct nuclear weapons tests) and crisis outcomes. I create *Security*, which measures the average number of crises a state experiences each year. *Security* is the total number of crises a state was involved in during the years preceding the current crisis, di-

[29] See *REGIME* in the ICB actor-level data set.
[30] See *NUCLEAR* in the ICB actor-level data set.

vided by the number of years it has been since 1918 (the first year in the ICB data) or since a state has entered the international system, if later than 1918.[31]

2.6 Results

To begin evaluating my hypothesis I use a series of probit[32] models with robust standard errors that are clustered by crisis dyad. I predicted that states that conduct more nuclear weapons tests than their adversary during a crisis are more likely to win that crisis. Table 2 presents the results from the main analysis, with model 1 featuring *Test Superiority* and model 2 testing *Test Ratio*. The results across both models 1 and 2 lend support for the *Nuclear Testing Superiority hypothesis*. In model 1, *Test Superiority* has a positive coefficient and is statistically significant.[33] In model 2, *Test Ratio* has a positive coefficient and is statistically significant, as well.[34] Taken together, these results indicate that states that conduct more nuclear weapons tests that their opponents during a crisis are more likely to win that crisis. These results are supported by robustness checks in Appendix 1 (Tables 11 and 12), which expand the sample to include all crisis dyads where at least one state possesses nuclear weapons, controls for whether State A in the dyad initiates the crisis and removes the United States from the sample. I will discuss the substantive effects of *Test Superiority* and *Test Ratio* in models 1 and 2 below, but first, will briefly discuss the results for some of the control variables.

---

[31] See Beardsley and Asal (2009) for another explanation of how to construct this variable.

[32] Using maximum likelihood estimation with a small sample size biases toward finding results (Long 1997). Long (1997) states that the minimum is at least 100 observations, with 500 or more being "adequate". I recognize these issues, which is why I attempt to provide additional evidence for my hypotheses by enlarging the sample size in Table 3 and by brining in a case study.

[33] The p-value for *Test Superiority* in model 1 equals 0.010.

[34] The p-value for *Test Ratio* in model 2 equals 0.052.

Table 2: Probit models with only nuclear dyads

	Model 1	Model 2
Test Superiority	1.93**	
	(0.75)	
Test Ratio		1.53*
		(0.79)
Warhead Superior	0.13	0.01
	(0.52)	(0.63)
Violence	0.68**	0.61**
	(0.21)	(0.17)
Gravity	-1.47**	-1.38**
	(0.73)	(0.67)
Proximity	2.16**	2.00**
	(0.62)	(0.59)
Capabilities	2.45	2.37
	(1.67)	(1.54)
Population	-0.13	-0.16
	(0.22)	(0.21)
Democracy	1.56**	1.25**
	(0.60)	(0.61)
Security	2.87**	2.42**
	(1.04)	(0.87)
Second Strike	-0.94*	-0.63
	(0.48)	(0.42)
Constant	-4.95*	-3.97
	(2.63)	(2.65)
Observations	62	62

Standard errors in parentheses, * p<.10, ** p<.05

Interestingly, the variable *Gravity* has a negative coefficient and is statistically significant across both models 1 and 2. This is surprising because it indicates that states that are facing a graver threat than their adversaries, because of the crisis, are less likely to win the crisis. These results make more sense when the results from *Proximity* are also considered. The results from *Proximity* in both models 1 and 2 indicate that states that are geographically closer to the crisis location than their crisis adversary are more likely to

win the crisis, confirming prior expectations. Altogether, the results from *Gravity* and *Proximity* indicate that states facing more severe threats than their crisis adversary are less likely to win that crisis, but that this could change depending on the location of the crisis.

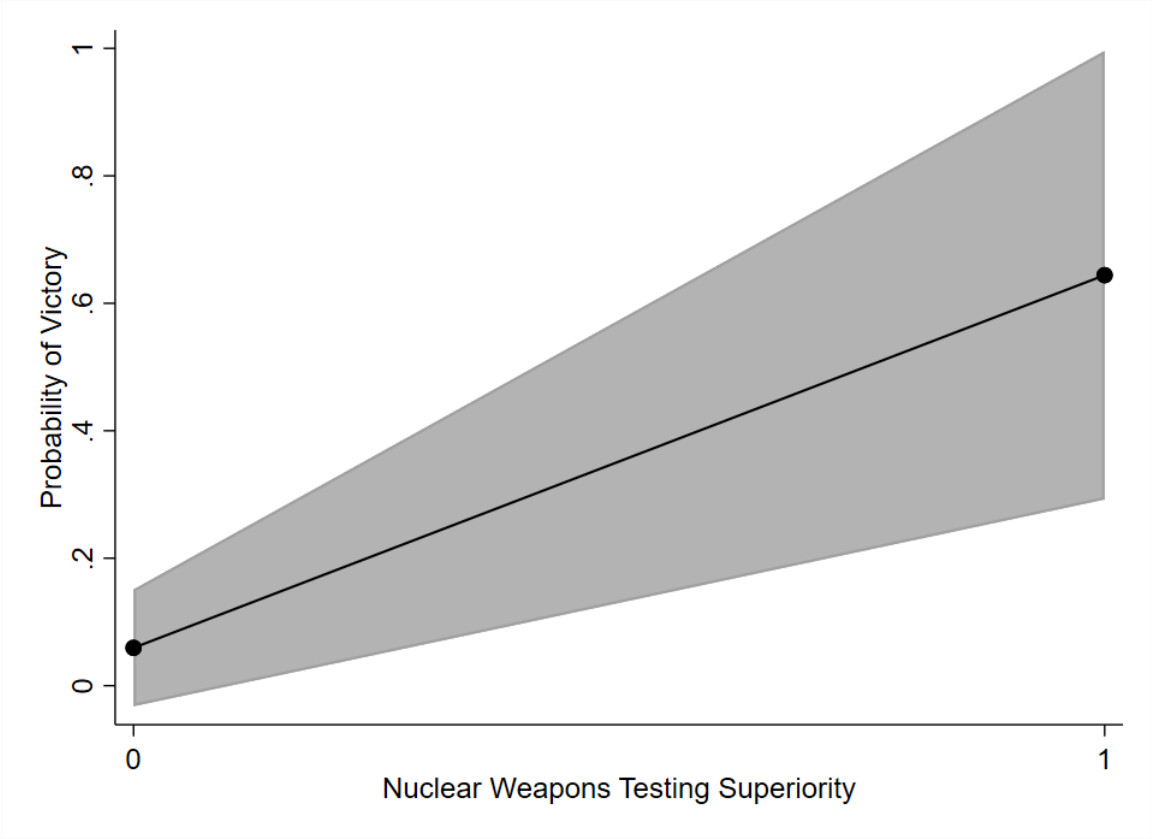
The other interesting results are from *Democracy* and *Security.* The coefficient for *Democracy,* across both models, is positive and statistically significant. The results indicate that democracies are more likely to win crises than non-democracies. This lends additional support to the notion that democracies are more likely to be victorious in crises because they select into crises that they are more likely to win (Gelpi and Griesdorf 2001). Finally, *Security* is positive and statistically significant in both models 1 and 2. *Security* measures the average number of crises that states participated in, in the years preceding whatever the current crisis is. The results for *Security* indicate that as the average number of crises a state is involved preceding the current crisis increases, the likelihood of victory in the current crisis also increases. This likely reflects some underlying and unmeasured factor, such as great power status, that causes states to be participants in more crises but also more likely to prevail in those crises. I now move towards interpreting the substantive effects from models 1 and 2.

2.6.1 Substantive effects for *Test Superiority* and *Test Ratio*

To determine the substantive effects of both *Test Superiority* and *Test Ratio*, I graph a state's predicted probability of victory in a crisis based on the state's *Test Superiority* and *Test Ratio*, while holding the control variables at their means. Figure 2.1 illustrates the substantive effects for *Test Superiority,* while Figure 1B shows the substantive effects for *Test Ratio*. With Figure 1A, you can see that moving from the lowest value of

Test Superiority to the highest value dramatically increases a state's probability of victory in a crisis. When states conduct less nuclear weapons tests than their crisis adversary (*Test Superiority*=0), their probability of victory in that crisis is about 5.9%.[35] As states engage in a strategy of brinkmanship and conduct more nuclear weapons tests than their crisis adversary (*Test Superiority*=1), their probability of victory in that crisis increases to about 64.4%. The difference in the probability of victory between being nuclear weapons testing inferior and nuclear weapons testing superior is 58.5%.

Figure 2.1: Substantive effects for *Test Superiority*



This is a percent change of about 991.5%, a massive substantive effect. It appears that nuclear-armed states that conduct more nuclear weapons tests than their adversaries

[35] I round all percentages to the nearest tenth.

during a crisis are in a much better position to secure victory, when compared to states who choose not to test their nuclear weapons during a crisis.

Figure 2.2: Substantive effects for *Test Ratio*

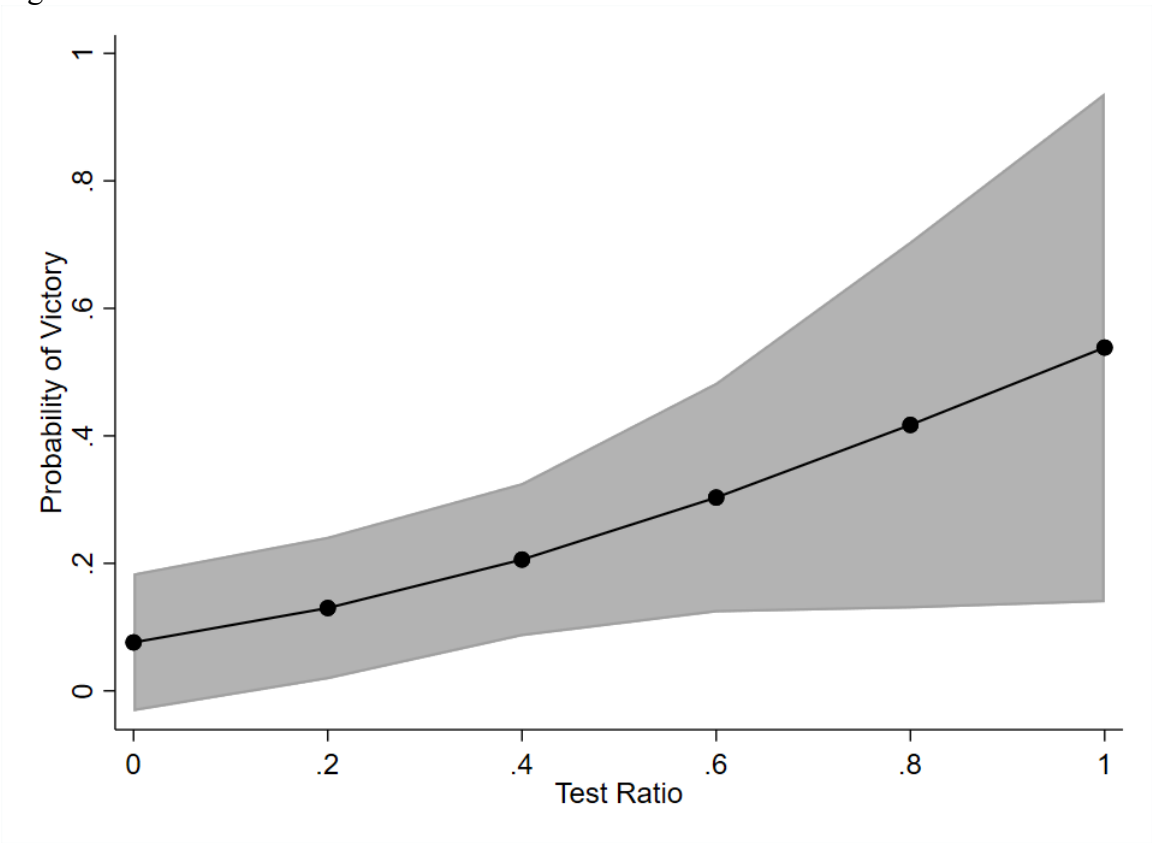


Figure 2.2 graphs the substantive effects for *Test Ratio*. Graphing the substantive effects of *Test Ratio* allows for a more fine-grained look at how nuclear weapons testing impacts crisis outcomes. Similarly to *Test Superiority*, moving from the lowest value of *Test Ratio* to the highest value increases a state's probability of victory in a crisis. States that do not conduct any nuclear weapons tests during a crisis (*Test Ratio*=0) have a predicted probability of victory in that crisis of about 7.6%. The predicted probability of victory increases to about 53.9% for states that conduct all of the nuclear weapons tests (*Test Ratio*=1) in their dyad during a crisis, a difference of 46.3%. Moving from the lowest level of *Test Ratio* to the highest level results in a percent change of 609.2%.

While these results largely match what we saw with *Test Superiority*, a more interesting picture emerges when we take a deeper dive into the differences among the different values for *Test Ratio*. In general, the differences in predicted probability of victory between the different values of *Test Ratio* grow faster as we move to higher levels of *Test Ratio*. For example, the difference in the probability of victory between states with a *Test Ratio* of 0.2 and 0.0 is about 5.4%, while at the opposite end of spectrum, the difference in the probability of victory between states with a *Test Ratio* of 1.0 and 0.8 is about 12.2%. While the differences in the probability of victory grow faster as *Test Ratio* increases, the substantive effects actually decrease, as the percent change in the predicted probability of victory decreases as *Test Ratio* increases. For example, the percent change in the predicted probability of victory as a state moves from conducting 0% of a dyad's nuclear weapons tests during a crisis to conducting 20% of the tests is about 71.1%. The percent change in the predicted probability of victory as a state moves from conducting 80% of a dyad's nuclear weapons tests during a crisis to conducting 100% of the tests is only about 29.3%. This difference in the percent change in the probability of crisis victory from the lower values of *Test Ratio* to the higher values of *Test Ratio* reflects how nuclear weapons tests reveal the balance of resolve. As one of the states in a crisis dyad reaches higher values of *Test Ratio*, it is clear that they are the the more resolved state, and so moving from conducting 80% of the tests to 100% of the tests does not reveal any extra information to their opponent.

Overall, the results presented above show that states that conduct more nuclear weapons tests than their adversaries during a crisis are more likely to win that crisis and also demonstrate that nuclear weapons testing can have a large substantive impact on

nuclear crisis outcomes. The results presented here and the uncertainty of the existing literature beg the question: do states that possess nuclear weapons but are nuclear weapons testing inferior (*Test Superiority*=0) have a greater probability of crisis victory than states that do not possess nuclear weapons at all? In the next section, I conduct additional empirical analysis to answer this question.

2.7 Additional analysis

To answer the question posed above I expand my sample and bring in crisis dyads where at least one state posseses nuclear weapons or neither state possess nuclear weapons, increasing the sample size to 810. I also create 3 new independent variables based on a state's nuclear status during a crisis and whether or not they conduct more nuclear weapons tests than their crisis opponent. The first variable, *Test Superior*, is the exact same as the main independent variable from the main analysis above. The second variable, *Test Inferior*, equals 1 when a state possesses nuclear weapons but conducts less tests than their opponent during a crisis. Finally, the third variable is *No Nukes*, which equals 1 when a state does not possess nuclear weapons. I run 3 additional probit models and change the baseline category in each model, while keeping all control variables the same. The results are presented in Table 3 below.

Model 1 includes *Test Superior* and *No Nukes* with *Test Inferior* as the baseline category. The results from Model 1 lend additional support to the *Nuclear Testing Superiority hypothesis*, as states that are nuclear weapons testing superior have a greater probability of victory in a crisis than states that are nuclear weapons testing inferior. Interestingly, the results for the *No Nukes* variable indicates that states that do not possess

nuclear weapons have a greater probability of victory in a crisis than states that possess nuclear weapons but are testing inferior, although this fails to reach significance.

Table 3: Additional Analysis

	Model 1	Model 2	Model 3
Test Superior	0.44**	0.12	
	(0.21)	(0.29)	
Test Inferior		-0.31	-0.44**
		(0.24)	(0.21)
No nukes	0.31		-0.12
	(0.24)		(0.29)
Warhead Superior	0.39	0.39	0.39
	(0.26)	(0.26)	(0.26)
Violence	0.04	0.04	0.04
	(0.04)	(0.04)	(0.04)
Gravity	-0.05	-0.05	-0.05
	(0.11)	(0.11)	(0.11)
Proximity	-0.06	-0.06	-0.06
	(0.14)	(0.14)	(0.14)
Capabilities	0.66**	0.66**	0.66**
	(0.21)	(0.21)	(0.21)
Population	-0.04	-0.04	-0.04
	(0.04)	(0.04)	(0.04)
Democracy	0.24**	0.24**	0.24**
	(0.11)	(0.11)	(0.11)
Security	-0.07	-0.07	-0.07
	(0.08)	(0.08)	(0.08)
Second Strike	-0.12	-0.12	-0.12
	(0.20)	(0.20)	(0.20)
Constant	-0.92*	-0.61*	-0.59
	(0.48)	(0.36)	(0.50)
Observations	810	810	810

Standard errors in parentheses, * p<.10, ** p<.05

In model 2, both *Test Superior* and *Test Inferior* fail to reach signficance with *No Nukes* as the baseline category. Finally, the results from Model 3, where *Test Superior* is the baseline category, offers more support for the *Nuclear Testing Superiority hypothesis*, as *Test Inferior* is negative and statistically significant. Once again, this demonstrates that

when compared to states that possess nuclear weapons and conduct more nuclear weapons tests than their adversaries during a crisis, states that possess nuclear weapons but are nuclear weapons testing inferior have a lower probability of victory in a crisis.

Taken altogether, the results from this additional analysis demonstrate support for the *Nuclear Testing Superiority hypothesis* and the notion that nuclear-armed states that conduct more nuclear weapons tests than their adversary during a crisis are more likely to win that crisis. The results do not definitively answer the question that motivated this additional analysis, as it appears that there is no statistical difference in the likelihood of crisis victory between testing inferior states and states that do not possess nuclear weapons. This has implications for policymaking still, as it shows that states that possess nuclear weapons must undertake provocative actions with those weapons, in order to demonstrate their resolve during a crisis. Just possessing nuclear weapons is not enough to improve your chances of victory. I now turn to a brief case study of the 1969 Sino-Soviet Border Conflict to illustrate how nuclear testing superiority can translate to crisis victory.

2.8 Sino-Soviet Border Conflict (1969)

The most immediate causes of the 1969 Sino-Soviet Border Conflict began in 1963, when China first raised the issue of 'unequal treaties' that were imposed on the weak Qing Dynasty in the 19th century and looked to remedy some of these historic complaints (Segal 1985). Throughout the 1960s, both the Soviets and the Chinese began the process of reinforcing their military presence along their shared border. Direct tensions began to arise again in 1968 with the Soviet invasion of Czechoslovakia and the issuing of the Brezhnev doctrine, which scared Chinese leaders into believing the Soviets may interfere

with Chinese domestic politics under the guise of "strengthening international Communism." The crisis officially began on 2 March 1969, when Chinese forces ambushed Soviet troops on Zhenbao Island in the Ussuri River, which formed the border between the two states (Gerson 2010). The island had seen numerous clashes between Chinese and Soviet border forces in the past, but the fighting usually devolved into fistfights, with the occasional use of clubs, sticks, and firehoses (Gerson 2010). The aim of the Chinese ambush was to deter the Soviets from interfering in China (Segal 1985).

Soviet leadership, believing the 2 March 1969 ambush was emblematic of an increasingly aggressive and revisionist China, initiated[36] another skirmish on 15 March 1969 (Gerson 2010). This skirmish was much larger and more deadly than the 2 March incident, with the Soviets using tanks, armored personnel carriers, artillery, and other heavy weapons in order to teach the Chinese a 'lesson'. The nuclear aspect of the crisis also began around the same time as the 15 March incident, with Soviet radio broadcasts and newspapers beginning to make nuclear threats toward China (Gerson 2010). Also, the Soviets conducted 2 nuclear weapons tests less than a month after the second skirmish, on 4 April and 13 April 1969, making it clear nuclear weapons would play a role in the crisis. These initial threats were originally not seen as credible by Chinese leadership but by August 1969, Chinese leaders had become increasingly concerned about the prospect of a Soviet first strike (Gerson 2010).

Once convinced that a first strike was imminent, Chinese leaders immediately began to prepare for war, including testing nuclear weapons on 22 September and 29 September (Johnston 2009; Segal 1985). The crisis finally ended on 20 October 1969 when

[36] While both sides accused the other side of initiating the 15 March 1969 skirmish most scholars and analysts, including a CIA assessment, believe the Soviets initiated this battle (Gerson 2010).

the Chinese finally agreed to sit down for negotiations over the border (Gerson 2010). Most scholars and analysts consider this crisis a Soviet victory[37], as the crisis was not over what a final border deal between the Soviets and Chinese would look like, but instead whether they would hold talks over the border at all (Kroenig 2018). This was clearly settled in the Soviet Union's favor as the Chinese leadership unwillingly submitted to negotiations over the border issues.

2.8.1 Why did the Soviet Union prevail in the 1969 Sino-Soviet Border Conflict?

The main reason why the Soviets won was that once the crisis started, they revealed themselves to be more resolved and risk-accepting than the Chinese. Kroenig (2018) argues that the Soviets won the crisis because their larger nuclear arsenal allowed them to be more risk-accepting, but this cannot be the entire story as the Soviet nuclear arsenal was superior to China's before the Chinese initiated the crisis and during the entire length of the crisis. Chinese leaders were not persuaded by their relative nuclear inferiority to not initiate the crisis or to quickly give in once the crisis began, as they held off on Soviet demands for negotiations for over 7 months. It was only after the Soviets undertook provocative, dynamic actions, such as nuclear weapons tests, that the Chinese backed down.

The Soviets undertook a few actions during the crisis that demonstrated their superior resolve and risk-acceptance to the Chinese. First, the Soviets conducted 6 nuclear weapons tests between the beginning of the crisis and August 1969, with 6 more to come while the crisis was still ongoing in September and October 1969 (Johnston 2009). This was compared to 0 tests conducted by the Chinese during that same March to August pe-

[37] Two notable exceptions: the ICB Project (Brecher et al. 2021) and Sechser and Fuhrmann (2017) consider this crisis to be a stalemate or inconclusive.

riod. Second, the Soviets continued to initiate border skirmishes, including a 13 August 1969 attack that killed 38 Chinese soldiers (Gerson 2010). Not only did they continue to initiate these skirmishes, but the Soviets also escalated the skirmishes by being the side that introduced heavy weapons, like tanks and artillery, into the skirmishes. Finally, in the summer of 1969, Soviet officials began approaching foreign government officials and asking how their governments would respond to a Soviet first strike against China. This was publicly revealed to Chinese leadership and the world on 27 August 1969, when CIA Director Richard Helms announced that the Soviets had been approaching governments in Eastern Europe about an attack on China's nuclear program (Gerson 2010).[38]

I argue that the Chinese backed down in October 1969 because by then the balance of resolve between them and the Soviet Union had been clearly revealed to be in the Soviets' favor. Soviet actions clearly demonstrated they were more resolved than the Chinese during the crisis and more willing to run the risk of nuclear war. Consistent with my theory, the Soviet Union conducted 12 nuclear weapons tests during the crisis compared to only 2 for the Chinese. More importantly, 6 Soviet nuclear weapons tests were conducted during the crucial period between March 1969 and August 1969 (compared to 0 for the Chinese) in which Chinese leaders became concerned about a Soviet first strike. While we lack direct evidence of the effects of Soviet nuclear testing superiority on Chinese decision-making, it is not hard to see that these tests, in combination with other actions like continued border skirmishes and the approaching of foreign governments about a first strike, demonstrated superior Soviet resolve. Additionally, these same actions also

[38] Helms left out that the Soviets had also approached American officials for their thoughts about a Soviet first-strike against Chinese nuclear facilities (Gerson 2010).

had the effect of raising the risk of nuclear war. This caused the Chinese to back down and begin negotiations over the border, a clear victory for the Soviet Union.

2.9 Discussion and Conclusion

The findings from the analyses conducted above lead to some policy recommendations. First, the results presented above demonstrate that states are incentivized to test nuclear weapons during crises. These tests obviously come with some dangerous consequences, including increasing the risk of nuclear war breaking out and increasing the amount of harm done to the physical environment and the health of human populations. Due to these dangerous consequences, a policy recommendation would be to convince states that have not signed and ratified the Comprehensive Test Ban Treaty to take steps to sign and ratify the agreement. Getting holdout states to sign and ratify the CTBT presents a significant hurdle with current politics, as competition between Israel-Iran and the United States-China make it unlikely that any of these states would agree to the terms of the CTBT without their main competitor moving first. Additionally, even absent international competition, it is unlikely the United States would be able to sign and ratify this treaty because of likely opposition from Republicans in the United States' Senate (Hansen 2006). Even with these hurdles, pushing for the CTBT to enter into force is a worthwhile goal as evidence from other international security institutions, such as the Nuclear Nonproliferation Treaty (NPT), demonstrates that increased systemic complicance with the institution decreases the likelihood of individual states violating the agreement (Kaplow 2022).

Additionally, the findings from the additional analysis can be used to strengthen and support the international nuclear non-proliferation regime because the results show

that nuclear weapons only help you win crises if you are willing to incur additional costs by undertaking provocative actions with those weapons, like conducting nuclear weapons tests. States considering acquiring or developing nuclear weapons because of the benefits it may bring them in coercive diplomacy and interstate disputes should reconsider if they are not willing to take provocative and risky actions with those weapons.

This project is the first to theoretically and empircally examine nuclear weapons tests in the international relations literature. Furthermore, I argue that the empirical analyses carried out in this project offer a more direct test of the the brinkmanship model by capturing dynamic, provocative actions taken during a crisis that can reveal the balance of resolve. The main finding from this paper is as follows: as predicted in the *Nuclear Testing Superiority hypothesis*, states that conduct more nuclear weapons tests than their adversaries during a crisis are more likely to win that crisis. Future extensions of this project could empirically test other provocative actions that could be linked to brinkmanship, such as missile tests or the deployment of new weapons systems, or examine the relationship between nuclear weapons tests and other portions of the crisis process such as onset or escalation.

CHAPTER 3. DANGEROUS DESIGN: THREAT AND ARMS CONTROL AGREEMENT DESIGN

3.1 Introduction

From ancient Greece to the Cold War and today, humans have long been concerned with what type of weapons are to be made, sold, deployed and used on the battlefield. These concerns have only grown in importance in the modern era as weapons technology has become increasingly deadly, especially with the invention and use of nuclear weapons at the end of World War II. The main way political actors, including modern states, have sought to control the manufacturing, spread, deployment, and use of these weapons is through negotiating international arms control agreements. Modern arms control agreements (post-WWII) have not only sought to limit what type of weapons are built, but also how many of them are built, where they are deployed, who they can be sold to, and how their effectiveness can be tested. Like other international agreements, not all of these agreements are equal; some are much more effective in constraining state behavior than others. One of the keys in determining their effectiveness is to examine how they have been designed, specifically by looking at the level of obligation in these agreements. This leads to the following question: what factors influence arms control agreement design?

There are multiple domestic and international factors that may plausibly influence the design of arms control agreements, one of which I examine here. I focus on the impact that threats from other states can have on design choices, once again, specifically examining how threat influences the level of obligation in these agreements. I examine threats originating from two different sources: threats from states with who you are negotiating and designing the arms control agreement with (called "partner threat") and threats

from states who are not involved in the negotiating process (called "external threat"). I choose to focus on threat instead of other factors because, in general, arms control agreements are formed to control the spread of arms and arms-related technologies, which states can and have used to threaten other states. Many arms control agreements are designed and formed in response to threats from other states. Thus, threat and arms control are intimately linked. While other scholars have examined how the design of arms control agreements impacts the likelihood that these agreements are ratified and entered into force (Kreps 2018), little is known about what influences states to choose that specific agreement design in the first place. By better understanding what drives states to make the design choices that they do, scholars and policymakers can have a better understanding of how these agreements come together, and how to best design them so they are effective in constraining state behavior when it comes to deadly weapons and potentially help save lives.

I choose to focus on the level of obligation in arms control agreements instead of other dimensions of agreement design, because the level of obligation determines how constraining the agreement will be on state behavior, and thus, how easy (or difficult) it is for states to cheat on the agreement. As a result, obligation is the most important dimension of design in determining whether an arms control agreement achieves the goal of controlling the spread of arms and arms-related technologies. Additionally, obligation is the most important design dimension in determining whether arms control agreements will be ratified and enter into force (Kreps 2018). Obligation is defined as the degree to which an agreement is legally binding (Kreps 2018), more specifically, whether the agreement includes escape, withdrawal, or sunset clauses that provide states mechanisms

for withdrawal should circumstances change in the future (Milner and Rosendorff 2001) or allow for reservations, declarations, and interpretive statements that exempt states from particular aspects of the agreement (Helfer 2012). Placing reservations, declarations, and interpretive statements (also known as "understandings") on an agreement communicates a state's intent to exempt itself from particular provisions of a treaty (Hill 2016), thus nullifying some of the constraining effects the agreement may have on state behavior.

Obligation can be thought of as how much flexibility the agreement allows for in state behavior. A low obligation agreement, like the Interim Agreement between the United and the Soviet Union, give states the greatest amount of flexibility and are the least constraining on their behavior. This 1972 agreement is considered low in obligation because Article VIII of the agreement includes a sunset clause, which means the agreement is in place for only a set amount of time, instead of indefinitely (Interim Agreement). In the case of the Interim Agreement, the sunset clause lasted for five years, allowing both sides to resume their arms build-up after it ended. Agreements high in obligation, like the Comprehensive Test Ban Treaty (CTBT), allow for the least amount of flexibility and are the most constraining. The CTBT is considered to be high in obligation because Article XV of the agreement does not allow for reservations to be placed on the agreement, meaning states cannot exempt themselves from any part of the agreement (Comprehensive Test Ban Treaty).

In this paper, I argue that states purposefully design the level of obligation in arms control agreements based on the levels of threat that the designing states are feeling from the two different sources described above. I expect the two different sources of threat to

lead to different design choices, with increases in partner threat leading to an increase in the likelihood of a high obligation agreement, while increases in external threat are expected to lead to a decrease in the likelihood of a high obligation arms control agreement.

The paper will proceed as follows: I begin by discussing examples of low obligation and high obligation arms control agreements before introducing a theory that draws on the different incentives states face when they are experiencing threats from different sources. I test this theory on all arms control agreements, including formal treaties and informal arrangements, negotiated between 1945 and 2010. This includes agreements that have been ratified and entered into force and agreements that have been designed but for one reason or another have either not been ratified or entered into force. I find some support for the notion that different sources of threat lead to different design choices, specifically, increases in external threat result in a decrease in the likelihood of designing a high obligation arms control agreement. I follow this up with a set of case studies to provide additional evidence before concluding with a discussion of the important implications of these findings for policy-making and future research.

3.2 General arms control

Although the frequency with which arms control agreements are negotiated and formed has increased with time, modern arms control agreements are still exceedingly rare (Coe and Vaynman 2020). The general goal of any arms control agreement is to control the spread of arms and arms-related technologies. Modern arms control agreements seek not only to regulate the manufacturing of weapons and related technologies, but also to regulate how those weapons are tested, where they can be deployed to, how many can be deployed, what international agencies monitor compliance, and which states weapons

and weapons-technologies can be sold too. Even with these diverse set of goals, one thing all arms control agreements share in common is a desire to constrain states' behavior to some degree when it comes to arms and arms-related technologies. As discussed above, arms control agreements vary in their ability to constrain states, with the level of constraint being determined by the level of obligation in these agreements. At the same time, states are usually worried about designing agreements that are too constraining and usually still seek to preserve some flexibility in their ability to manufacture, test, deploy, sell, and acquire arms in order to be able to react to an uncertain future.

This leads to a trade-off between the ability to constrain state behavior and preserving flexibility in state behavior. This trade-off is present in the design of all international agreements. Koremenos (2005) examines the role that international uncertainty plays in international agreement design, with states being hesitant to commit themselves to an agreement if there are large degrees of uncertainty because they are afraid changing circumstances will alter their expected benefits. She finds that as uncertainty increases, the probability that states will design agreements with flexible duration provisions increases. While more flexible agreements allow states to have a wider latitude in what behavior (or when the behavior) is considered acceptable by the agreement, they also make it more difficult to constrain state behavior and lessen the effectiveness of the agreement. For arms control agreements, more flexibility will allow states to build, sell, deploy, and test more types of weapons, allow them to easily leave an agreement in order to build-up their arsenal, or place reservations and understandings on the terms of the agreement. Flexible agreements are those that have lower degrees of obligation. Higher obligation arms control agreements are by nature less flexible and more constraining, but they are

superior in preventing cheating and limiting states' abilities to build, sell, deploy, and test weapons.

One aspect of uncertainty in international politics is which states will threaten other states politically, economically, and militarily in the future. As we have observed over the years, relationships between states can become better or worse depending on a number of developments that are often difficult to foresee. I argue that what determines which side of the constraining-flexibility trade-off that an arms control agreement falls on is the level and source of threat that states designing the agreement face. Before a deeper discussion on the role of threat in influencing arms control agreement design, I discuss two examples of arms control agreements, to give the reader a better understanding of what low and high-obligation agreements look like.

3.2.1 Low obligation: the Wassenaar Arrangement

An example of a low obligation agreement is the Wassenaar Arrangement (hereafter known as the Arrangement). The Arrangement entered into force on December 19th, 1995 with a declaration issued at the Peace Palace in The Hague (Wassenaar Arrangement). The Arrangement was intended to succeed the Coordinating Committee for Multilateral Export Controls (COCOM), a Cold War-era arms control arrangement between seventeen non-Communist, Western states led by the United States, that sought to restrict the export of sensitive military technology and dual-use items that could contribute to the military capabilities of Communist states.

As the successor organization to COCOM, the Arrangement shares many of the same goals and a similar structure. As of 2020, there are forty-two states who take part in the Arrangement, including many former Communist states in Eastern Europe. The pur-

pose of the Arrangement is to promote transparency in regards to transfers of conventional arms and dual-use goods and technologies, to ensure through national policies that transfers of these items do not contribute to the development or improvement of military capabilities, to complement and reinforce existing arms control agreements, and to improve cooperation between the member states of the Arrangement when it comes to preventing the acquisition of weapons and dual-use technologies by actors in unstable regions or by actors whose behavior is threatening peace and security.

This arms control agreement is considered a low obligation agreement for a few reasons. First, like COCOM, there is no official treaty text which governs the organization. There is the Wassenaar Arrangement Initial Elements document that was created at the founding of the organization, but this is not a standard international treaty and has not been deposited with the UN. With no treaty text, states maintain maximum flexibility and do not have their behaviors constrained at all, except by their own voluntary choices. With no treaty to fall back on, the Arrangement is a voluntary arrangement with no treaty-specific mechanism to enforce compliance.

While states must meet certain criteria in order to become a member of the Arrangement, once a member, there is no treaty text which would constrain their behavior. Instead, member states of the Arrangement are only constrained by their own choices and national policies, not by the Arrangement itself. This allows states to maintain maximum flexibility as they can decide to change course at the national level and easily withdraw from the Wassenaar Arrangement or place reservations and understandings on the terms of the arrangement that would except them from prohibitions against exporting specific technologies or weapons to rogue states. This flexible structure is common among other

arms control export regimes like the Missile Technology Control Regime or the Proliferation Security Initiative and is in direct contrast to a high obligation agreement, like the Treaty of Pelindaba which I will discuss next.

3.2.2 High obligation: Pelindaba Treaty

Pelindaba Treaty, also known as the African Nuclear-Weapons-Free Zone Treaty (hereafter ANWFZ) entered into force on July 15th, 2009 after the deposit of the twenty-eighth instrument of ratification. The treaty has been signed and ratified by forty-one states in Africa, with additional protocols to the treaty being signed and ratified by every member of the Permanent Five members of the U.N. Security Council except the United States. Discussions for creating a nuclear-free Africa had begun much earlier than this though, with the "Declaration on the Denuclearization of Africa" being issued by the heads of state and government of the Organization of African Unity (predecessor to the African Union) in 1964. Building off this declaration, the ANWFZ Treaty's main goal is to prevent the deployment, manufacturing, and testing of nuclear weapons on the continent of Africa, with secondary goals including contributing towards global non-proliferation and disarmament through regional non-proliferation and disarmament, promote cooperation in the peaceful development and uses of nuclear energy, and to protect African states from possible nuclear attacks or the threat of nuclear attacks. To achieve these goals, a high obligation arms control agreement has been designed.

Two articles in this treaty specifically, Articles XVI and XVII, make this treaty a high obligation arms control treaty. Article XVI ensures that no reservations are placed on this treaty, so states who sign and ratify this treaty are not allowed to have exceptions to the treaty (Pelindaba Treaty). This constrains state behavior by not allowing states the

avenues to argue that the treaty does not apply to them under specific circumstances or to specific weapons, facilities, tests, deployments, and other actions. Article XVII ensures the treaty is in force for an unlimited duration and does not mention any specific sunset or escape clauses. With no sunset clauses states' behavior is constrained by ensuring that they are not able to begin to develop a nuclear weapons program after the treaty ends; instead, the unlimited duration of the treaty prevents states' the opportunity to pursue a nuclear weapons program indefinitely.

Not having escape clauses in the treaty constrains behavior by not giving states' the option of leaving the treaty and immediately pursuing nuclear weapons. Escape clauses differ from formal withdrawal mechanisms, in that escape clauses allow the state who exercises the clause to immediately be free of the terms of the treaty, albeit often with some type of penalty imposed. Formal withdrawal mechanisms allow states to leave agreements early but only after some period of time has passed. For example, Article XX of the ANFWZ Treaty allows for states to withdraw from the treaty but only after they have given the treaty body 12 months' notice of their intent to leave. This is not unusual among arms control agreements.

A high obligation arms control agreement like the ANWFZ Treaty is relatively rare, as low obligation arms control agreements are more common than high obligation agreements. Of the forty-eight arms control agreements examined in this paper, thirty-nine are considered low obligation agreements, while nine are considered high obligation agreements. This is puzzling because high obligation agreements are universally acknowledged as being better at constraining states' behavior and preventing the manufacture, spread, deployment, and testing of weapons. If the goal of arms control is to pre-

vent the spread of weapons, why are states more likely to design these weak, low obligation agreements that are less effective in constraining state behavior? Once again, I argue that regime type and partner and external threat all play a role in influencing the design of arms control agreements. I now briefly review the international agreement design literature before moving into my theory and expectations.

3.3 Threat and arms control agreement design

Threat is defined as a situation in which a state has both the capabilities and intentions to inflict a negative consequence on another state (Davis 2000). Threat is essentially the ability to coerce a state to do something they would not do by themselves, with this often leading to negative consequences for the state being threatened. Threats are probabilistic in that they may or may not be acted on. Threat has two key dimensions that must be realized in order for something to be considered a threat: intentions and capabilities. With intentions, states must have negative intentions towards another state and want to force negative consequences on that state in order for that dimension to be satisfied. It is not enough just to want to be able to inflict a negative consequence on another state, in order to be considered threatening, a state must be able to act on those intentions, thus bringing in the capabilities dimension. Without the capabilities to inflict a negative consequence, negative intentions will not be considered threatening. Consider the hypothetical example of Peru and Botswana. Even if Peru and Botswana greatly disliked each other, had opposing foreign policies, and had negative intentions toward one another, we would not consider them to be threats to each other because neither state has the capability to project sufficient military or economic force to inflict a negative consequence on the other.

In the context of arms control agreements, states may face threats from two sources: partner and external threats. Partner threat refers to threats originating from states with who you are directly negotiating and designing the arms control agreement with. External threats are threats that come from states that are not involved with the original design of the arms control agreement. The source of the threat should lead to differences in arms control treaty design.

With partner threat, states are directly negotiating with the sources of that threat and as such, have the opportunity to constrain their adversaries' behavior in order to reduce the level of threat they are facing. This should lead to states wanting to "lock it in" and create less flexible and thus, higher obligation agreements, but the level of partner threat felt by each state is different. These differences mean that it is actually the least threatened state, among the states negotiating and designing the agreement, that determines the level of obligation in the agreement. This is because the least threatened state does not need to constrain their "partners" as much or as urgently as the other states do. With that being said, the least threatened state will still want to constrain their partners to a certain degree.

States have different preferences for arms control treaty design when facing external threats. There are not the same incentives to "lock it in" and design agreements high in obligation. Instead, because the threats are coming from outside the arms control agreement, states will not want to be constrained by the arms control agreement and will favor flexibility in agreement design. If states are constrained by an agreement, they will be less able to build, acquire, and deploy certain types of weapons that may help them successfully face an external threat. This leads to the following hypotheses:

Partner threat hypothesis: As the minimum partner threat increases, the probability of designing a high obligation agreement should increase.

External threat hypothesis: As the level of external threat that states engaged in an arms control agreement design face increases, the probability of designing a high obligation agreement should decrease.

In the next section, I describe the empirical strategy I use to evaluate my hypotheses. I then follow that up with a discussion of the results from the statistical analysis in this chapter before moving on to a set of case studies.

3.4 Research design

I test my theoretical expectations on data from Kreps (2018) that covers all arms control agreements concluded, negotiated, and seriously considered from 1945 to 2010. This dataset includes both agreements that were ratified and entered into force and those that were negotiated and designed but either were not ratified or did not enter into force. The dataset includes 48 different arms control agreements, which are listed in Table 3.4 in Appendix 2. The unit of analysis is at the arms control agreement level, so all variables are aggregated up to the agreement level. Since the independent variables of interest are all measures of state characteristics that are then aggregated up to the arms control agreement level, I identified all states that participated in negotiations over each agreement. Not all states that sign and ratify treaties participate in the design and negotiation of the treaty. For example, one hundred and ninety states are recognized as having signed the Nuclear Non-proliferation Treaty (NPT Treaty) since it was opened for signature; but only eighteen states actually took part in the design of this treaty. In cases where it was not entirely clear which states participated in the design of an agreement, I code all states who signed the agreement within a week of it being open for signature as participants in

the design of the treaty. I am assuming that if states are willing to sign a treaty almost immediately when it opens for signatures then they are probably aware of its design and gave their input somewhere in the negotiating process. I use logistic regression to test my theory because the dependent variable is binary.

3.4.1 Dependent variable: Level of agreement obligation

The dependent variable of interest is the level of obligation within arms control agreements. Once again, arms control agreement obligation refers to how legally binding the agreement is. Kreps (2018) codes obligation on a 4-point scale that ranges from 0 to 3, with "0" referring to an agreement that is informal, not legally binding and is instead based on a voluntary arrangement. A "1" refers to a formally binding agreement that is legally binding, but allows for reservations, understandings, declarations, and escape or sunset clauses. A "2" refers to a formally binding agreement that has no reservations, declarations, or understandings but does allow for escape and sunset clauses. Finally, agreements coded as "3" are the most binding agreements and are those without escape or sunset clauses or reservations, declarations, or understandings. Agreements with high obligation scores are the most constraining on state behavior.

I then turn obligation into a binary dependent variable by recoding the categories so that all the agreements are coded as either 0 or 1. Low obligation agreements that were originally coded as 0 or 1 are now all coded as equaling 0, while high obligation agreements that were originally coded as 2 and 3 now equal 1. I move from a categorical dependent variable to a binary variable because of the small number of agreements that fell into the highest categories of obligation, especially agreements where obligation equaled 3. This small number of agreements precludes me from producing any reliable estimates

when it comes to the obligation=3 category. By decreasing the number of possible categories to two, I increase the number of agreements that fall into each category and thus, increase the reliability of these estimates.

3.4.2 Main independent variables: Partner threat and external threat

The main two independent variables in this study are the level of threat experienced by states designing an arms control agreement that originates from the other states who are also taking part in the design process of an agreement ("partner threat"), and the level of threat experienced by states designing an arms control agreement that originates from states not involved in designing a specific agreement ("external threat").

To calculate threat, I adopt the procedure used by Leeds and Savun (2007). I will briefly describe this process. First, threat is a function not only of capabilities but also of goals and intentions (Walt 1988). We know that states that share foreign policy orientations are less likely to engage in military conflict than those states with opposing foreign policies. But not all states are able to reach other states militarily because of a lack of capabilities, so I first identify all states that are able to reach a state militarily. States that share land borders and great powers, who have the military capabilities to reach any state in the world, are identified as being able to reach a state militarily. I identify the great powers as the Permanent Five members of the UN Security Council. I then employ Signorino and Ritter's (1999) S score to identify states with similar or opposite foreign policy orientations. Any state whose S score is below the population median and shares a land border or is a great power is coded as threatening (Leeds and Savun 2007). Using S scores captures the goals and intentions aspect of threat. Finally, to capture the capabilities aspect of threat, I simply sum the capabilities of all threatening states using the Cor-

relates of War CINC scores. This gets me the overall level of threat that a state is facing in a given year.

For partner threat, I then subtract out the CINC scores of every threatening state who was not identified as taking part in the arms control agreement negotiation or design, leaving me with the threat each state is feeling from other states involved in the design process. I then use the lowest partner threat score, as the least threatened of the partners has the ability to act as a veto player and have the greatest influence on agreement design. The least threatened state is the least incentivized to design a high obligation agreement because they do not have the same need to constrain their partners as the more threatened states do, and so they can greatly impact the design of the arms control agreement.

For external threat, I sum together the CINC scores of all threatening states who were not involved in the arms control agreement negotiation and design process. I then add all these up for each state involved in an agreements' design process and divide by the number of states in the design process to get an average external threat score for each agreement. I use the average to ensure that the level of external threat is not being driven by how many states are involved in the design of the agreement. Agreements vary in the number of states involved in the design from bilateral agreements like SALT I between the United States and the Soviet Union to the large multilateral agreements such as the NPT which involved a large number of states. I then take the natural log of both the partner and external threat variables to approximate a normal distribution.

3.4.3 Control variables

I control for a number of factors that may influence arms control agreement design. All of these control variables come from Kreps (2018), except for regime type. I use

a dummy variable to indicate whether the agreement was negotiated in the Cold War. I code every observation before 1991 as a "1" and every observation after 1991 as a "0". Additionally, I include a dummy variable to indicate whether an agreement was multilateral or not. Multilateral agreements may be more difficult to design because you have to take into account the conflicting preferences and interests of three or more actors instead of just two.

I also control for the depth of the agreement. This is also a dummy variable in which "1" indicates that the agreement changes arms outputs such as the number of weapons or the locations that they may be deployed. Agreements with more depth are more costly to states since they will most likely have to change their behavior in order to meet the requirements of the agreement. Depth may influence arms control agreement design because states may want to circumvent a "deep" agreement by making the agreement low in obligation, presenting them with an easy exit avenue if it becomes too costly to meet the terms of the agreement.

Finally, I include a control variable that measures the proportion of states involved in the design process of a specific arms control agreement who are democracies in the year before the agreement's design is finalized to ensure that regime type is not driving my results. There is a deep literature that shows that democracies prefer more flexible, low obligation agreements because these types of agreements are easier for states to uphold, and less likely to result in the leaders of democracies being punished by their constituents for not upholding the terms of the agreement (Leeds 1999; McGillivray and Smith 2008; Chiba et al. 2015). To create this measure, I first identify all the states that took part in the original design of the agreement and then use Polity scores to code

whether that state was a democracy in the year before the design of the agreement was finalized. I code any state that scores a 6 or above on Polity2 as a democracy and every state that scores below 6 as a nondemocracy. I then divide the total number of democracies by the number of states that were involved in the original design of the agreement. I use the proportion of democracies because my unit of analysis is arms control agreement-year and so this allows me to capture how democratic the original states in the agreement are.

3.5 Results

To test my two hypotheses, I run multiple models all using logistic regression[39] with robust standard errors. Table 4 presents the results from these tests. Coefficients from logistic regression are difficult to interpret; to make interpretation easier I use percent change in the odds and graph the predicted probability associated with each of the main independent variable while keeping the control variables constant at their means.

[39] Using maximum likelihood estimation with a small sample size biases toward finding results (Long 1997). Long (1997) states that the minimum is at least 100 observations, with 500 or more being "adequate". I recognize these issues, which is why I attempt to provide additional evidence for my hypotheses by probing the results in Table 5 and by bringing in two case studies.

Table 4: Level of obligation in arms control agreements

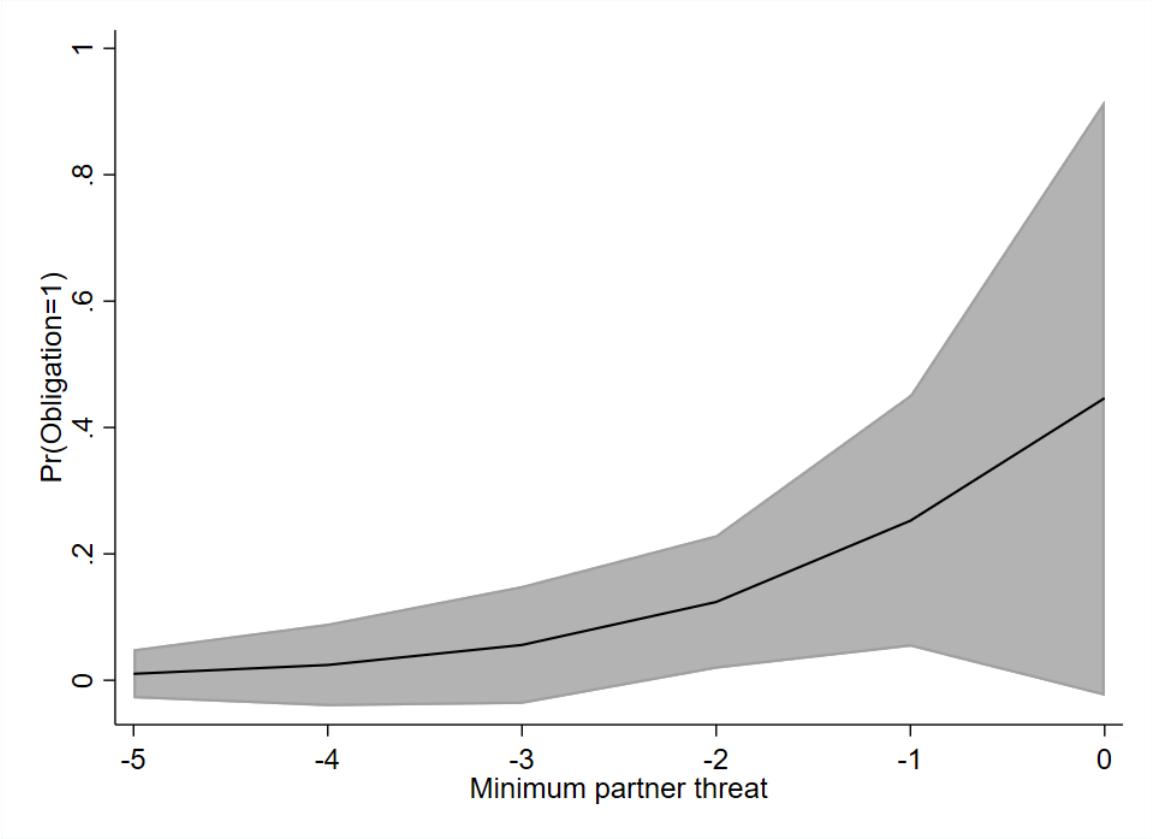
	Model 1	Model 2	Model 3	Model 4	Model 5
Partner Threat	0.70	0.87			2.37
	(0.42)	(0.55)			(1.73)
External Threat			-0.39**	-0.37*	-0.79
			(0.16)	(0.19)	(0.52)
Multilateral		0.57		0.79	-1.81
		(1.18)		(1.13)	(1.27)
Depth		2.19		1.39	1.75*
		(1.13)		(0.93)	(1.02)
Cold War		-0.69		-0.04	0.53
		(0.86)		(0.97)	(1.93)
Regime Type		0.10		-1.16	1.12
		(2.23)		(2.97)	(2.13)
Constant	-0.36	-0.75	-2.68**	-3.08*	-1.11
	(0.70)	(1.43)	(0.73)	(1.72)	(1.60)
Observations	48	48	48	48	48

Standard errors in parentheses, * p<.10, ** p<.05

In the partner threat hypothesis, I predict that as the partner threat of the states involved in negotiating each arms control agreement increases, the probability of designing a high obligation arms control agreement should increase. I do not find support for this hypothesis. While the coefficients for both models 1 and 2 are positive, indicating that as partner threat increases so does the probability of designing a high obligation arms control agreement, neither is statistically significant. In model 2, for each one unit increase in the level of partner threat, the odds of designing a high obligation arms control agreement increases by 138.8%, holding all other variables constant. Although using odds ratios and percentage changes to interpret the coefficients of logistic regression are common methods of interpretation, they do not indicate the magnitude of the change in the probability of the outcome (Long and Freese 2014). To indicate the magnitude in the change in the probability of the outcome, I plot the predicted probabilities.

Figure 3.1 plots the predicted probabilities of being in the highest obligation (obligation=1) category using the results from model 2. I hold all the other variables at their means, with 95% confidence intervals. I interpret the substantive effects below.

Figure 3.1: Substantive effects for *Partner threat*



On average, going from an arms control agreement where the minimum partner threat is at 0 and states are generally on friendly terms, like the Treaty of Tlatelolco, to an arms control agreement where the minimum partner threat is at its greatest, increases the probability of designing a high obligation arms control agreement by 42.5%. This indicates that higher levels of partner threat are more likely to result in high obligation agreements, although once again, these effects are not statistically significant. Overall, the results here do not indicate support for the partner threat hypothesis. It seems that even when receiving threats from a state with whom you are directly negotiating an arms control agreement with, states are hesitant to constrain themselves and surrender their flexibility, even if it would mean constraining the source of the threat as well.

In the external threat hypothesis, I predicted that as the level of external threat states designing an arms control agreement face, the less likely they should be to design a constraining, high obligation agreement, in order to preserve their flexibility and be able to meet those external threats. Model 3 in Table 4 shows the relationship between external threat and obligation, while model 4 shows the relationship with all control variables included. The coefficient for Model 4 is negative and statistically significant, indicating strong support for the external threat hypothesis. Similar to the results for the partner hypothesis, to interpret the substantive results for external threat, I first use the percent change in odds before plotting the predicted probabilities of moving from the lowest levels of external threat to the highest levels of external threat, while keeping all the control variables at their means, and with a 95% confidence interval. In model 4, for each one unit increase in the level of external threat, the odds of designing a high obligation arms control agreement decreases by 30.8%, holding all other variables constant.

Figure 3.2: Substantive effects for *External threat*

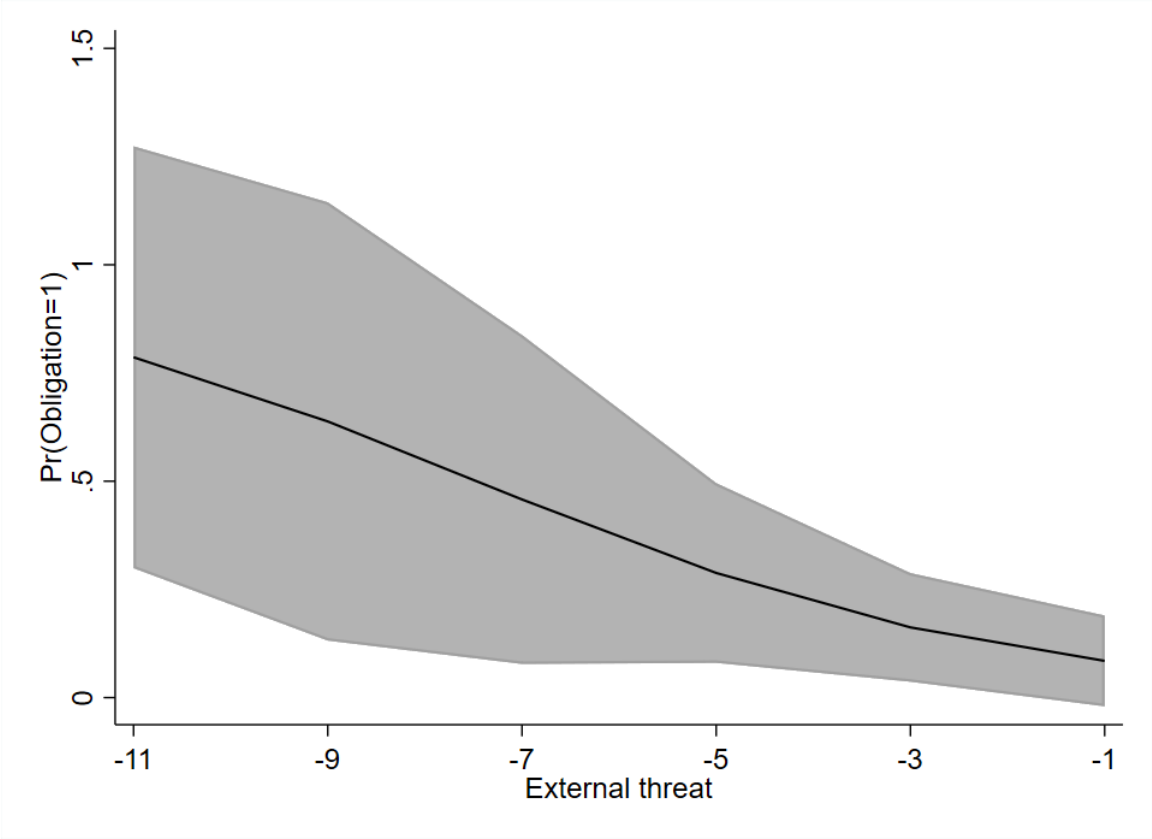


Figure 3.2 plots the predicted probability of designing the highest-level obligation agreement, as the level of external threat increases from its lowest value to its greatest value, using model 4 from Table 4. In Figure 3.2, on average, moving from the lowest level of external threat to the highest level of external threat decreases the probability of designing the highest-level obligation arms control agreement by 66.3%. These results indicate support for the external threat hypothesis. As the levels of external threat experienced by states designing arms control agreements increases, the probability of designing a constraining, high obligation agreement decreases greatly. Similar to the dynamic described in the discussion of the Iran nuclear deal, states that are in the process of shaping an arms control agreement and are experiencing threats from states not involved in the same negotiation process, opt to preserve their flexibility to better be able to meet those

external threats. In the context of arms control agreements, preserving flexibility means states would have greater capabilities to deploy, build, sell, test, and acquire arms and arms-related technologies.

Finally, for the main set of results, Model 5 in Table 4 shows the full model with both of the independent variables and all control variables included. The results mirror those discussed above. Both partner threat and external threat maintain the same relationship with the level of obligation in arms control agreements as they did in the models discussed above. Increases in partner threat are associated with increases in the probability of designing high obligation arms control agreements, while increases in external threat are associated with decreases in the probability of designing high obligation agreements. A one-unit increase in partner threat, increases the odds of designing a high obligation arms control agreement by 971.9%, holding all other variables constant. While for external threat, a one-unit increase is associated with a 54.6% decrease in the odds of designing a high obligation agreement, with all other variables constant. Once again, I graph predicted probabilities to demonstrate the magnitude of the change in the probability of the outcome.

Figure 3.3 shows the predicted probabilities of designing a high obligation agreement based on the levels of partner threat, using the results of model 5 in Table 4. For Figure 3.3, on average, moving from an arms control agreement where the minimum partner threat is at 0 to an arms control agreement where the minimum partner threat is at its greatest, increases the probability of designing a high obligation arms control agreement by 73.0%.

Figure 3.3: Substantive effects for *Partner threat* using model 5

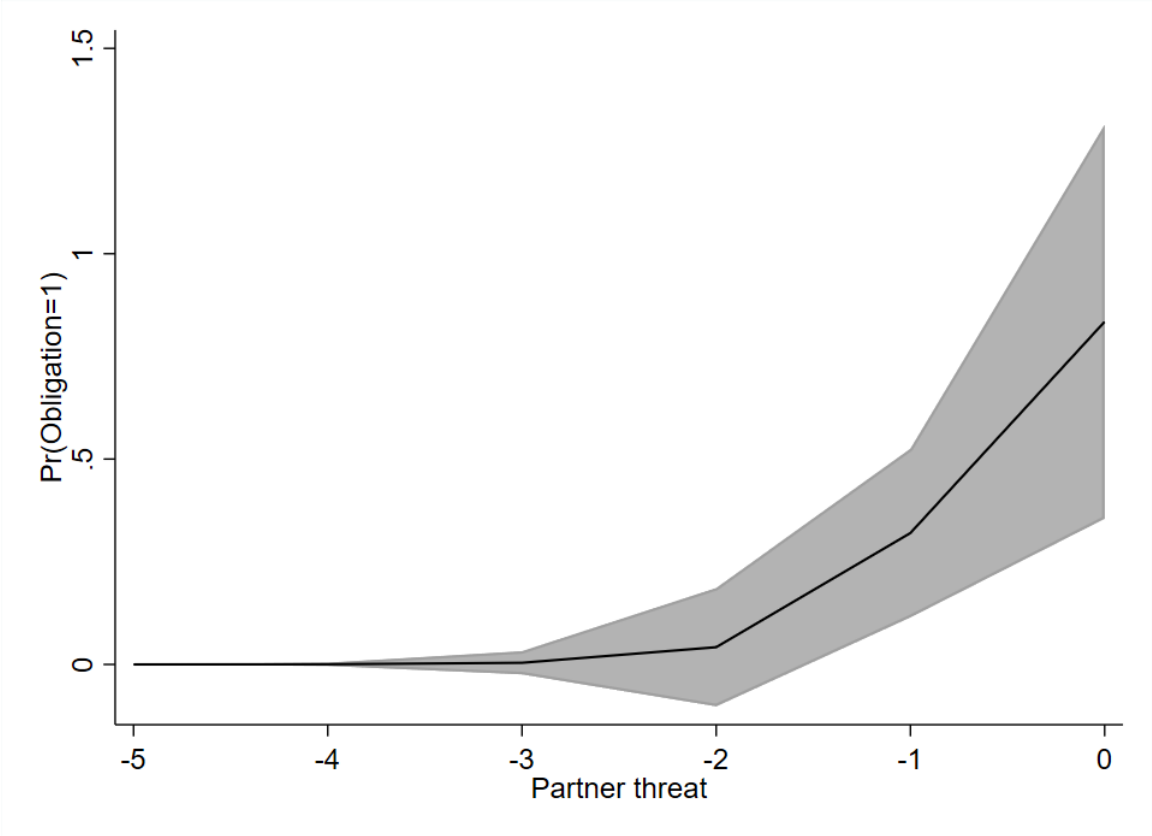
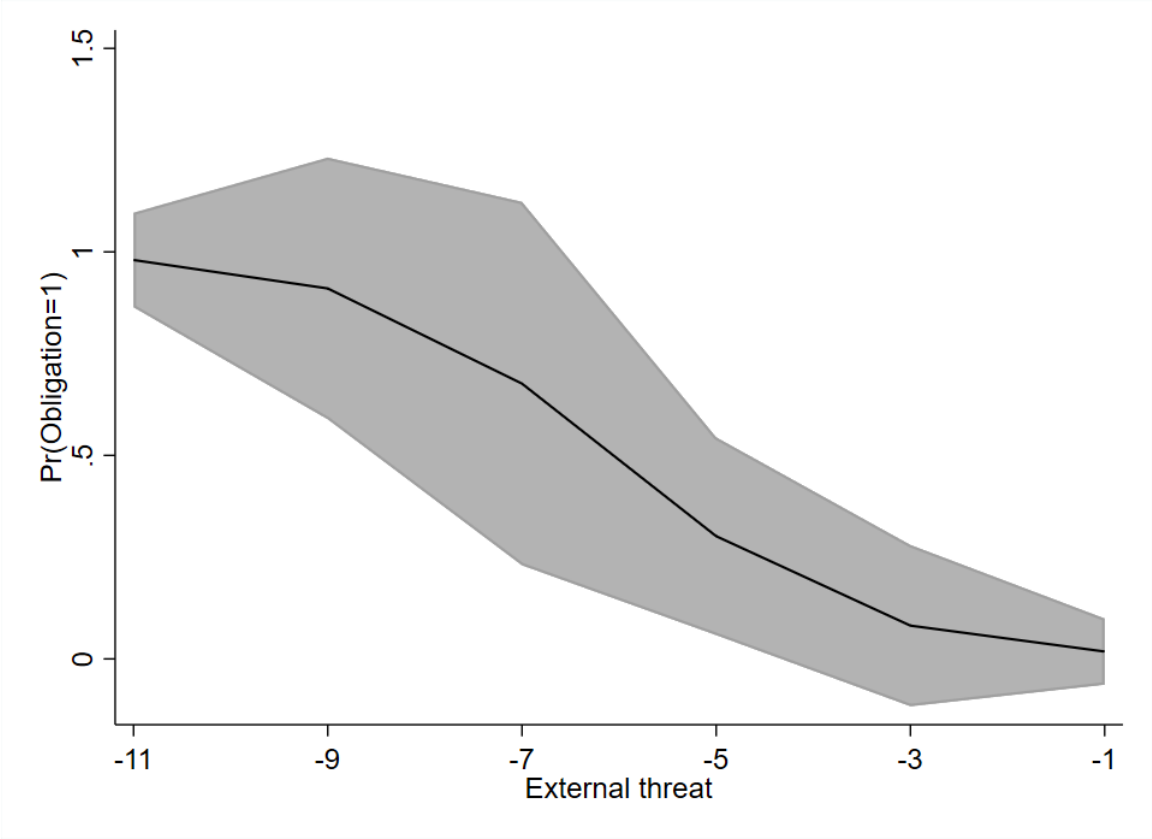


Figure 3.4 shows the predicted probabilities of designing a high obligation agreement based on the levels of external threat, using the results of model 5 in Table 4. In Figure 3.4, the substantive effects are quite large. On average, increasing external threat from the lowest levels to the highest levels decreases the probability of designing a high obligation arms control agreement by 84.7%.

Figure 3.4: Substantive effects for *External threat* using model 5



3.5.1 Robustness check

The small sample size with which I am testing my theory naturally leads to questions about the robustness of the results presented above. To check the robustness of the results, I used Pregibon's delta-beta influence statistic to identify observations that are influential in determining the relationship between partner and external threats and the level of obligation. I then re-ran models 2 (partner threat) and 4 (external threat) from Table 3.1 while excluding the most influential observations. For model 6, which is testing the partner threat hypothesis, I removed the four most influential observations, while in model 7, I removed the three most influential observations. I removed an extra observation in model 6 because two observations had the same exact Pregibon's delta beta value. If the results from above hold, especially concerning the external threat hypothesis, then

we can be more confident that threat plays a role in determining how states design arms control agreements. Table 5 presents the results below.

Table 5: Robustness check

	Model 6	Model 7
Partner Threat	0.02	
	(0.71)	
External Threat		-0.93*
		(0.55)
Multilateral	2.57	0.47
	(1.81)	(1.72)
Depth	2.56	0.90
	(1.61)	(1.05)
Cold War	-0.54	3.76
	(1.28)	(3.16)
Regime Type	-5.43	-0.78
	(4.57)	(6.44)
Constant	-1.55	-8.21
	(2.36)	(6.73)
Observations	44	45

Standard errors in parentheses, * p<.10, ** p<.05

The results shown here mirror those above and indicate support for the external threat hypothesis. In model 6, the coefficient for partner threat is positive, but like the results in model 2, partner threat is still not statistically significant. In model 7, as the level of external threat that states designing an arms control agreement face increases, the probability of those same states designing a high obligation agreement decreases, even after I have accounted for the most influential observations. Overall, the results presented in this section and the section above offer strong support for the external threat hypothesis. States' design preferences for arms control agreements are influenced by threats they receive from states who are not participating in the design of that arms control agreement.

Do these results indicate that threats coming from arms control agreement design partners are inconsequential for the level of obligation observed in those agreements? When, if ever, is partner threat influential in determining the level of obligation in arms control agreements? For many states, in any given year they are likely to experience both partner and external threats and so determining the extent that partner threat matters could be useful in helping scholars determine what threats states are more likely to respond to. In the next section, I determine when partner threat is influential in determining what the level of obligation will be in an arms control agreement.

3.5.2 When does partner threat matter?

To determine when partner threat is influential in determining the level of obligation in arms control agreements, I begin by restricting my sample to arms control agreements where external threat is at its lowest. Partner threat should be significant in cases where states are experiencing low levels of external threat because states do not face the same decision calculus as they did when external threat was high. When external threat is low it should lose its salience and partner threat should dominate state decision-making. If partner threat is not significant when external threat is at its lowest, then it is unlikely that partner threat is influential in determining the level of obligation in arms control agreements.

To empirically test this, I run two logistic regression models that are similar to the models run above, with the results being shown in Table 6. In model 8, I restrict the sample to only the arms control agreements where the level of external threat is in the 75th percentile or below. Model 9 restricts the sample to agreements where the level of external threat is in the 50th percentile or below.

Table 6: When does partner threat matter?

	Model 8: 75th percentile	Model 9: 50th percentile
Partner Threat	1.72*	-2.46
	(0.98)	(2.85)
Multilateral	1.12	2.93
	(1.25)	(3.35)
Depth	1.51	6.40
	(1.10)	(3.91)
Cold War	-1.40	-4.30
	(1.61)	(2.31)
Regime Type	-4.44	-23.86**
	(3.47)	(10.65)
Constant	3.39	4.98
	(3.14)	(7.11)
Observations	36	24

Standard errors in parentheses, * p<.10, ** p<.05

In model 8, partner threat is positive, indicating that as partner threat increases, the probability of designing a high obligation arms control agreement also increases. In model 9, the coefficient for partner threat flips and becomes negative, indicating a lack of support for the partner threat hypothesis. This is not too concerning considering the small sample size. Overall, the results from model 8 indicate some support for the partner threat hypothesis, but only when the level of external threat is low or nonexistent. It seems that partner threat is influential only when external threat is low, and that external threat is more important to state's decision-making when it comes to the level of obligation in arms control agreements than partner threat.

Next, I use a set of case studies to provide additional evidence. For partner threat, I discuss the negotiations around and the design of the SALT II agreement between the United States and the Soviet Union during the Cold War, with each sides' high level of threats toward one another leading a relatively high-level of obligation agreement. For

external threat, Israel's exclusion from the Iran nuclear deal epitomizes how external threats influence arms control agreement design and influences the level of obligation in these agreements. Israel's continued threats towards Iran over the last few decades, as well as their exclusion from the constraints of the Iran nuclear deal, led Iran to refuse to agree to a high obligation agreement and resulted in the low obligation Joint Comprehensive Plan of Action (JCPOA) that allowed Iran to maintain flexibility in order to deal with Israeli threats.

3.6 SALT II: Arms control during the Cold War

The Strategic Arms Limitation Talks (SALT) were a series of arms control discussions and agreements designed and negotiated by the United States and the Soviet Union during the Cold War. Beginning in 1968 and continuing almost continuously until the SALT II agreement was signed in 1979, the SALT talks resulted in three arms control agreements between the United States and the Soviet Union, with discussions around SALT II beginning in November 1972 (U.S. State Department 2017). The general goal of the SALT II talks was to design and commit to a long-term treaty between the United States and the Soviet Union to limit each sides' quantity and quality of strategic offensive (i.e., nuclear weapons) weapons and delivery vehicles and to bolster and replace both the SALT I agreement and the Interim Agreement (Nuclear Threat Initiative 2011).

With these goals in mind, both sides agreed to the following reductions and limitations to their nuclear arsenals: 2,400 vehicle limit on strategic nuclear delivery vehicles (i.e., intercontinental ballistic missile launchers (ICBMs), submarine-launched ballistic missiles (SLBMs), and heavy bombers); 1,320 limit on multiple independently targetable re-entry vehicle (MIRV) systems; prohibition in the construction of new land-based

ICBM launchers; and limits on the deployment of new types of strategic offensive arms. While these reductions were important in removing and eliminating dangerous weapons from deployment, for this paper, the more pressing issue is whether a low or high obligation arms control agreement was designed.

Based on the data collected by Kreps (2018), scholars would consider the SALT II agreement a relatively high obligation arms control agreement. The reason that SALT II is considered a relatively high obligation agreement is that it is a legally binding agreement that does not allow for reservations, understandings, and declarations (SALT II Article XII); it is not considered at the highest levels of obligation because clause one of Article XIX limits the duration of the treaty to December 31st, 1985, thus including a sunset clause in the agreement. The inclusion of a sunset clause precludes the agreement from being considered to be at the highest level of obligation. What led to the United States and the Soviet Union to design a relatively high obligation arms control agreement, instead of opting for a less constraining low obligation agreement? Consistent with the argument laid out in the section above, I argue and present evidence that the "partner" threats felt and issued by the United States and Soviet Union towards one another led them to design a constraining, high obligation agreement that would "lock" in their behavior when it came to strategic offensive weapons and delivery vehicles.

Since SALT II is a bilateral agreement, the United States and the Soviet Union represent partner threats to each other. I predicted that as the level of partner threat in an arms control agreement increases, then the probability of designing a high obligation agreement should also increase. For both states, the SALT II talks represented a unique opportunity to directly constrain their biggest adversaries' ability to deploy and build

strategic offensive weapons. This opportunity to constrain the source of each state's biggest threat occurred in the midst of the Cold War, when the United States, the Soviet Union, and both states' allies and proxies battled for global political, economic, and military dominance.

It is without question that during this time period, the biggest threat to the United States was the Soviet Union and vice versa. Both sides dominated their spheres of influence of politically and economically, and together they possessed a large amount of the world's military capability and nuclear weapons. Additionally, both sides were either just exiting or about to enter long-lasting military conflicts during the SALT II negotiation and design period (November 1972-early 1979). The United States finally withdrew all military personnel from Vietnam in March 1973 (Herring 2001), while the Soviets were still 6 months away from invading Afghanistan at the time of the signing of SALT II. Thus, both the U.S. and Soviet Union were not focused on conflicts on the periphery of the international system and could devote more time and resources to countering threats from each other.

If both sides were the main threats to each other during this time period, then why did the SALT II talks not result in an agreement with the highest level of obligation? While direct evidence of this is difficult to gather, I argue that the reason the SALT II agreement did not reach the highest levels of arms control agreement obligation is because it was negotiated and designed during the period of détente between the United States and the Soviet Union. Détente was a period of easing tensions between the U.S. and the Soviet Union in the midst of the Cold War and saw an increase in attempts to cooperate on a range of issues, from arms control to the economy and lasted throughout

most of the 1970s (Cox 1990). It should be noted that even during this period of relative calm and cooperation between the superpowers there were still significant disagreements with one another and both sides viewed the other as the primary threat in the international system.

With that being said, the relatively positive and cooperative relationship between the United States and Soviet Union during the period of détente may have led the SALT II agreement to being a step-below the highest-level obligation agreements because both sides had more trust for one another than before and did not see it as imperative that they constrain their adversaries' capabilities to the greatest degree possible. Once again, the reason why the SALT II agreement is not considered to be at the highest level of obligation is because it features a sunset clause that limits the duration of the agreement. It is possible that this sunset clause was put into the agreement, during a time of improved relations, because it was believed by either side that relations would continue to improve and there would be less of a need to constrain each other in the uncertain future. I now turn towards the discussion of how external threat influences the level of obligation in arms control agreements by digging deeper into the negotiation and design process of the Iran nuclear deal.

3.7 Israel and the Iran nuclear deal

The Joint Comprehensive Plan of Action (JCPOA), also known as the Iran nuclear deal, is one of the most recent arms control agreements to have been negotiated, designed, and entered into force. The deal was finalized in 2015, with the main accomplishment of the agreement being the slowdown of Iran's nuclear program in exchange for the lifting of sanctions that had been placed on the Iranian economy by the United

States and others. The most common criticism of the deal was that it would only delay and not indefinitely prevent Iran from being able to produce a nuclear weapon (Lewis, Siddiqui, and Jacobs 2015; Kershner and Sanger 2015). If fully implemented, the JCPOA was predicted to prevent Iran from producing fissile material for nuclear weapons for at least 10 to 15 years, after that period, the deal expires and Iran would be free to pursue a nuclear weapon (Samore 2015). The JCPOA was designed and signed by Iran, the United States, the United Kingdom, France, Germany, Russia, and China. The one notable state who was left out of the design process of the agreement was Israel, one of Iran's greatest threats.

While leaving Israel out the design process makes sense in that if they were included in the design process then it would make finalizing a deal even more difficult than it was without them in the agreement, leaving Israel out of the deal influenced how the agreement was designed. Similar to Republicans in the United States, Israel's main concern with the deal was that it limited the duration of the agreement to 15 years, with Iran being free to pursue a nuclear weapon after that time period (Kershner and Sanger 2015). With this sunset clause in the JCPOA limiting its' effectiveness to only 15 years, this deal would be considered a weak, low obligation arms control agreement. Israeli Prime Minister Benjamin Netanyahu, in a rather infamous speech to the United States' Congress railed against the deal, and stated that Iranian acquisition of a nuclear weapon would pose a threat to the very existence of the state of Israel (Kershner and Sanger 2015). With the close relationship between the United States and Israel, why was the United States not able to help design an agreement that would do a better job of allaying the Israeli's fears of Iranian acquisition of a nuclear weapon?

I argue, based on my theory discussed above, that the reason why the JCPOA was designed to be a weak, low obligation agreement, and not a constraining, high obligation agreement was because of Iranian fears of Israeli's threats. Very simply, Iran was never going to agree to a high obligation agreement that seriously constrained their long-term ability to acquire nuclear weapons because they knew the Israelis were not going to be constrained by that same agreement in any way and wanted to preserve some flexibility in order to deal with an uncertain future. While finding direct evidence of this is difficult because of the secretive nature of the Iranian regime, and because some of the people involved in the design process and negotiation of the agreement are still in power in certain states, there is enough evidence to indicate that Israel's failure to be included in the design process led to the JCPOA being a weak, low obligation agreement.

Since Israel was not involved in the design of the JCPOA, they would be considered an external threat to the Iranians. Remember, I found that as external threat increases, we should see an increase in the probability of a weak, low obligation being designed, as states involved in designing the agreement aim to preserve their flexibility to meet outside threats. For Iran, Israel is especially a threat when it comes to the acquisition of nuclear weapons, as the Israelis have shown a unique willingness to use force to ensure that regimes they deem hostile are unable to acquire nuclear weapons. Besides being a nuclear power in their own right, the Israelis have used a combination of traditional military force, cyber capabilities, and assassination to ensure that Syria, Iraq, and Iran did not acquire nuclear weapons. With Iraq and Syria, Israel used more traditional methods to prevent these states from being able produce their own nuclear weapons, bombing an Iraqi nuclear reactor in 1981 and a Syrian reactor in 2007.

With Iran, Israel has taken a more covert, and less direct approach to slowdown the Iranian nuclear program. In 2010, it was discovered that Israel had partnered with the United States to use what has been described as one of the world's first cyberweapons to physically destroy Iranian nuclear centrifuges (Langner 2011). Called Stuxnet, this cyberweapon did not steal, manipulate, or erase information, but instead caused Iran's nuclear centrifuges, components vital for enriching uranium and creating a nuclear weapon, to uncontrollably spin and eventually damage themselves, thus leading to lengthy delays in repairing or replacing the broken centrifuges. This has not been the only action Israel has taken against Iran's nuclear program. Another high-profile incident occurred in 2020, when Iran's top nuclear scientist, Mohsen Fakhrizadeh, was assassinated (Atwood 2020). This follows a series of assassinations of Iranian nuclear scientists in the last 10 years. With both Stuxnet and the series of assassinations, Israel has not openly claimed responsibility, but has been acknowledged as the source of these attacks by the global media and other governments, with US officials noting that Israel was behind the assassination of Fakhrizadeh (Atwood 2020).

Fully knowing Israel's history of meddling in the nuclear programs of other states and the tense relationship between themselves and the Israelis, the Iranians were not going to constrain themselves by agreeing to a high obligation agreement. As an external threat, the Israelis are not constrained by the JCPOA in any way, and are free to hassle the Iranians and their nuclear program. By avoiding the design of a high obligation agreement with a sunset clause, the Iranians preserved their future flexibility to deal with threats from Israel. If the Iranians had agreed to a high obligation agreement without a sunset clause, they would have permanently surrendered their ability to build a nuclear

weapon, severely hampering their ability to deter future threats from Israel, especially nuclear threats.

3.8 Conclusion

This study examined what factors influence the design of arms control agreements, specifically focusing in on how different sources of threat influence the design choices of states when negotiating arms control agreements. Overall, the results suggest that my theory can explain levels of obligation in arms control agreements. The results also show that different sources of threats can lead to different preferences, as increases in external threats were reliably associated with increases in the probability of designing a high obligation agreement, while there less evidence of the role that partner threat plays in arms control agreement design. The findings illustrate the importance of states' security environment in influencing their foreign policy preferences and that states can not design and negotiate agreements in the international system without taking into account the considerations of other actors.

Additionally, one implication from this paper is that the influence of external threat goes beyond agreement design and moves into compliance. Increases in external threat should be associated with an easier ability to comply with an arms control agreement because increases in external threats decrease the probability of designing a high obligation agreement. High obligation agreements are more difficult for states to comply with because of the constraints that are placed on state behavior. Thus, higher levels of external threat should lead to low obligation agreements that are easier for states to comply with. Future work could test this relationship more broadly, examining how external threats influence compliance with different types of international agreements.

One major policy recommendation that arises out this study is which states should be included in the design of arms control agreements, in order to be able to design constraining, high obligation agreements that are more effective in preventing the spread of arms and arms-related technologies. With arms control looking poised to be one of the most salient areas of foreign policy in the upcoming years, especially among great powers like the United States, Russia, and China, it will be important for policymakers to know which states to get onvolved in the design of these agreements in order to craft the most effective arms control agreement possible. Coming back to the Iran nuclear deal, one recommendation based on this line of thinking would be to include Israel in the negotiation and the design of any subsequent agreements. By including the Israelis in the design of any subsequent agreements, Iran can be sure that one of their biggest adversaries and a source of threat towards them will be constrained by that agreement, thus, making it much more likely that the Iranians agree to a constraining, high obligation agreement.

CHAPTER 4. IMPLEMENTING THE DUAL-TRACK APPROACH: CONSTRAINTS ON U.S. COUNTERPROLIFERATION POLICY

4.1 Introduction

The signing of the Joint Comprehensive Plan of Action (JCPOA) on July 14th, 2015 by Iran and the P5+1 states (the U.S., U.K., France, Germany, Russia, and China) represented a win for both the Obama Administration and the international counterproliferation regime. After years of sanctions and threats of military force by the United States and its' allies, Iran agreed to limit their pursuit of nuclear weapons, in exchange for some key rewards including the ability to acquire materials in order to operate their civilian nuclear facilities, the lifting of the oil embargo, sanctions relief, and the resumption of diplomatic ties with the United States and other states. The success of the Obama Administration's (and the P5+1 broadly) negotiating strategy has been credited to the "dual-track" approach that was used, in which negative inducements[40] (the impositions of sanctions and threat of military force) were used in combination with positive inducements (civilian nuclear cooperation and political aid) to get Iran to forego nuclear weapons (Mehta 2020).

However, the fragility of the JCPOA, which President Trump quickly pulled the United States out of once he entered office, demonstrates the importance of executive branch-Congressional relations for counterproliferation policy, as the Obama Administration's reliance on executive branch power to implement the deal, doomed the deal once President Obama left office. Additionally, it has been relatively rare for the United States to use the dual-track approach as part of its counterproliferation policy, with the use of positive inducements only being a much more popular strategy than the dual-track ap-

[40] I use the terms "inducements" and "policies" or "policy instruments" interchangeably.

proach. This raises the following question: why does the United States struggle to implement the dual-track approach in its' counterproliferation policy?

Understanding the conditions under which a dual-track approach is implemented is important because past research demonstrates that the dual-track approach is the most effective in inducing nuclear reversal among potential proliferators (Mehta 2020).[41] This goes against conventional wisdom, in which economic sanctions and the threat of military force were thought to be the most effective tools the United States had in the fight against nuclear proliferation. Using economic sanctions against potential proliferators decreases the probability of nuclear reversal by 27%, while the threat or use of military force decreases the probability of nuclear reversal by 21% (Mehta 2020: 91). Clearly, these negative inducements do not lead to nuclear reversal when used by themselves, demonstrating the need to use both positive and negative inducements. By better understanding the conditions under which this approach is most likely to be implemented, scholars and policymakers can develop a better understanding of determinants of the United States' counterproliferation policy and possibly develop solutions to overcome some of the obstacles that prevent dual-track implementation. I argue that United States' foreign policy towards potential nuclear proliferators, and whether the dual-track approach is implemented, is constrained by both domestic politics and the international strategic environment.

At the domestic level, U.S. presidents are constrained by their relationship with Congress, and specifically whether they have high levels of concurrence with Congress. When facing a hostile Congress with whom they have low levels of concurrence, the

[41] It should be noted that even the dual-track approach does not guarantee success. If a potential proliferator is determined enough and can direct enough resources into their nuclear program, then they can develop nuclear weapons, regardless of what actions the United States and the international community take.

president's power over foreign policy is weakened and it becomes much more difficult for them to implement their preferred, dual-track approach. At the same time, presidents are also constrained by the international strategic environment, specifically, by whether the potential proliferating state has a friendly relationship with the United States or not. U.S. presidents are unlikely to want to upset allies or potential allies through the use of sanctions or the threat of military force, and so they are less likely to use negative inducements against these types of states than they would be when facing adversaries interested in acquiring nuclear weapons.

While scholars have previously examined how U.S. presidents make foreign policy (Gries 2014; Milner and Tingley 2015), there has been a lack of scholarly attention on U.S. foreign policy-making towards potential nuclear proliferators, who pose unique challenges to the United States. Additionally, the counterproliferation literature has failed to theorize about when the United States is likely to use *both* positive and negative inducements in order to convince potential proliferators to end their nuclear programs, instead usually focusing on either type of inducement. This project presents a novel explanation about what determines the United States' counterproliferation strategy, while giving scholars and policymakers a better understanding of what U.S. and global counterproliferation policy will look like under certain conditions.

The paper proceeds as follows: I begin with a brief overview of the counterproliferation literature before introducing my theory on the role of domestic politics and the international strategic environment on influencing the United States' counterproliferation policy. I then test my theory on all states who pursued nuclear weapons between 1953 and 2007, and find that the when the president has a good relationship with Congress and

the U.S. is dealing with a potential proliferator with whom it has had a militarized interstate dispute within the past 3 years, the dual-track approach is more likely to be implemented than when those conditions are not present. I conclude with a discussion of the important implications these results have for policy-making and future academic research.

4.2 The outline of counterproliferation policy

While the literature surrounding the pursuit and acquisition of nuclear weapons is well developed and has received plenty of attention over the years (e.g., Sagan 1996; Jo and Gartzke 2007; Fuhrmann 2009), the empirical literature concerning what influences states to give up nuclear weapons (and the pursuit thereof) has only received intense examination within the last decade. This literature has largely focused on theorizing about positive and negative inducements separately, with only some more recent work examining the effectiveness of the dual-track approach in inducing nuclear reversal. I will first discuss the relevant research that examines the role of positive inducements in counterproliferation policy before turning towards negative inducements, and finally ending with a discussion of nuclear reversal and the dual-track approach.

4.2.1 Positive inducements and counterproliferation policy

The United States, as the leading state in the international counterproliferation regime, often uses positive inducements in order to persuade potential proliferators to forego acquiring nuclear weapons. For example, Mattiacci et al. (2022) find that nuclear latent states that possess overt lab-scale enrichment and reprocessing facilities are more likely to receive cooperative overtures from the United States. The authors argue this is

because the United States wants to incentivize these states away from pursuing nuclear weapons. At the same time, if the potential proliferator decides in the future that they want to pursue nuclear weapons, the United States now has leverage and can threaten to withdraw its' cooperative overtures to coerce the potential proliferator into not pursuing nuclear weapons. Additionally, nuclear latent states who attempt to keep their nuclear facilities covert are less likely to receive cooperative overtures from the United States because the covert nature of the nuclear program signals the dark intentions of the program to the United States, who may be unwilling to spend valuable political capital on potential proliferators who seem more resolved in their pursuit of nuclear weapons.

4.2.2 Negative inducements and counterproliferation policy

Other scholars have focused on the role of negative inducements in counterproliferation policy. Fuhrmann and Kreps (2010) focus on the threat of or actual use of military force as counterproliferation policy and examine when states consider attacking nuclear infrastructure in nonnuclear states. They argue that states attack or consider attacking nuclear programs when they are threatened by the potential proliferator's acquisition of nuclear weapons. The authors find that attacks on or the consideration of an attack on nuclear infrastructure are more likely when: the attacking state and potential proliferator have had violent conflict in the past; the potential proliferator is an authoritarian state, and; the attacking state and potential proliferator have dissimilar foreign policies. Additionally, the authors find that contrary to conventional wisdom, states are actually more likely to attack or consider attacking powerful potential proliferators than weak potential proliferators because the risk of retaliation is not salient in the attacking state's calculus and the more powerful potential proliferators pose a bigger threat if they acquired the

bomb. While understanding the role of positive and negative inducements in counterproliferation policy is important on their own, the research from Mehta (2020) that is described below, combines some of these strands and demonstrates that using both positive and negative inducements simultaneously is the most effective strategy in limiting potential proliferators' nuclear weapons ambition.

4.2.3 Nuclear reversal and the dual-track approach

One of the first empirical examinations of nuclear reversal comes from Mattiacci and Jones (2016), who argue that the process of nuclear proliferation is nonlinear and identify three transition possibilities that can either move states closer to a nuclear weapon (starting a nuclear program or acquiring a nuclear weapon) or further from a nuclear weapon (nuclear reversal). As it relates to nuclear reversal, they argue and find support for the idea that a strong nonproliferation regime is associated with a decreased likelihood of nuclear reversal, as a strong nonproliferation regime increases the costs associated with restarting (or starting) a nuclear program and so potential proliferators are less likely to give up their current pursuit of nuclear weapons to avoid the future costs associated with restarting it. Additionally, they find that states with high levels of nuclear latency are more likely to give up their pursuit of nuclear weapons when compared to states with low levels of nuclear latency. This is because states with high levels of nuclear latency face lower costs of restarting their pursuit of nuclear weapons in the future, and thus, are incentivized to give up their pursuit of nuclear weapons in the present time period in order to avoid the costs and pressures associated with an active pursuit of nuclear weapons, while still knowing that they will be able to restart their pursuit in the future.

Similarly, Mehta (2020) argues and finds support for the idea that nuclear reversal is significantly influenced by external factors, specifically, whether or not external actors offer both negative and positive inducements to the potential proliferator. Mehta (2020) argues that states interested in preventing the spread of nuclear weapons have a number of different policy instruments they can use to persuade potential proliferators to either give up their pursuit of nuclear weapons or their actual nuclear weapons. These policies can be broadly divided into "positive" inducements and "negative" inducements[42] (Mehta 2020). Positive inducements include policies such as military aid, economic (foreign) aid, political aid, and civilian nuclear cooperation. Negative inducements include economic sanctions and the threat of or use of military force.

Mehta (2020: 26) finds that nuclear reversal is most likely when states are threatened with sanctions (a negative inducement) but also offered rewards (positive inducements) to compensate them for a lost nuclear weapons deterrent, with the threat of U.S. military force lurking beneath the negotiation process. Additionally, it is unlikely that the use of sanctions alone is enough to delay or stop nuclear development. Sanctions must be paired with positive inducements in order to motivate the leaders of potential proliferators to end their nuclear programs. Also, leaders play an important role in nuclear reversal, as new leaders may be better able to end nuclear programs than incumbent leaders. New leaders can use the positive inducements they receive in exchange for ending their nuclear program as political cover to satisfy their domestic constituencies while also pointing out the detrimental effects negative inducements may have on their state (Mehta 2020: 28). Overall, Mehta (2020) argues and finds support for the idea that the dual-track ap-

[42] It may be useful to think of positive inducements as "carrots" while negative inducements are "sticks", and when used together form a carrot-and-stick approach to counterproliferation.

proach is the most effective strategy for inducing potential proliferators to give up their nuclear programs.

Mattiacci and Jones (2016) focus on the role of potential proliferators' domestic capabilities (nuclear latency) and the broader international community's efforts (nonproliferation regime), while Mehta (2020) focuses on the role of individual states in influencing potential proliferators' decision to end their pursuit of nuclear weapons. Both contribute to our understanding of nuclear reversal, but neither provides an explanation of why the United States is unable to always use the effective dual-track approach with potential proliferators. As previously stated, understanding under what conditions a dual-track approach is likely to be used is an important next step for scholars and policymakers to better understand the obstacles to successful counterproliferation policy. In the next section, I introduce my theory on the domestic and international-level constraints that make it difficult for the United States to implement a dual-track approach.

4.3 Determinants of the United States' counterproliferation policy

The United States' counterproliferation policy is constrained at both the domestic and international levels. As I will discuss in the following sections, both the domestic and international levels influence U.S. presidents when it comes to determining counterproliferation policy for different states. At the domestic level, the president's implementation of the dual-track approach is constrained by whether the president has concurrence with Congress because of the obstacles Congress can put in the way of counterproliferation policymaking. Internationally, U.S. presidents are constrained by the potential proliferator's relationship with the United States. I will begin by discussing the constraints U.S. presidents face at the domestic level.

4.3.1 Congress and the United States' counterproliferation policy

While much of the U.S. foreign policy literature focuses on the unique role and power of the president when it comes to making foreign policy, it is important to acknowledge that the president is not the only actor who influences U.S. foreign policy. Both houses of Congress also exert influence over the shape of U.S. foreign policy through their ability to constrain the president's actions by refusing to pass legislation, refusing to ratify international agreements, or by criticizing the policy in the media so that it becomes politically problematic to implement, among other actions. While I assume that U.S. presidents will prefer to use a dual-track approach for counterproliferation policy whenever possible, Congress has often been one of the main obstacles to successfully implementing and maintaining this dual-track approach.

For example, after the 1994 midterm elections in which Republicans gained unified control over Congress, Republicans began to heavily criticize the Agreed Framework that was signed between the United States and North Korea right before the midterm elections in October 1994 (PBS 2003). In the agreement, North Korea agreed to freeze its' nuclear program in exchange for civilian nuclear cooperation and improved economic and political relations with the United States. This agreement was negotiated during a time in which the U.S. had placed sanctions on North Korea and was considering the use of an air strike to cripple North Korea's nuclear infrastructure. After gaining power in Congress, many Republicans accused the Clinton Administration of appeasing the "rogue" North Korean regime, with Senator John McCain even going so far as to call the lead negotiator for the United States, Robert Gallucci, a "traitor" (PBS 2003). The intense criticism and political backlash over the deal led the Clinton Administration to drag its' feet implementing some of the provisions of the agreement, upsetting the North Koreans

and eventually being one of the factors that contributed to the demise of the Agreed Framework. This example demonstrates the role Congress can play in influencing the United States' counterproliferation policy. The Agreed Framework was negotiated and designed during a period in which the president and Congress had high levels of concurrence, the importance of which will be described below, while its' implementation phase took place during a period of more hostile relations between the Clinton Administration and Congress. I will now discuss some of the barriers to implementing the dual-track approach.

4.3.2 Congressional barriers to dual-track implementation

I identify three barriers to implementing the dual-track approach. The first barrier is distributive politics, and specifically how positive and negative inducements can generate large costs and benefits for domestic actors. Different inducements can either be high on distributive politics, when they create large costs and benefits (i.e., creating "winners" and "losers") for domestic actors or low on distributive politics, when these concentrated costs and benefits are not created. For example, economic (foreign) aid, a commonly used positive inducement, creates high levels of distributional politics in the United States because it generally favors capital-abundant individuals and highly-skilled exporting interests who benefit from donor countries becoming wealthier and better able to buy their goods and services, while it harms unskilled workers and interests who now may face increased competition from foreign workers and interests (Milner and Tingley 2015: 43-45). While the positive and negative inducements used for counterproliferation policy by the United States vary in their levels of distributional politics, the way they are paired together in a dual-track approach means that it is likely that at least one of the in-

ducements is going to have high levels of distributional politics. This is likely to lead to opposition from some members of Congress, whose constituents may lose out from the policy.

Another barrier to implementing the dual-track approach is ideological divisions. Similar to distributional politics, the inducements that make up the dual-track approach vary in the levels of ideological divisions they create. I adopt Milner and Tingley's (2015) conceptualization of ideological divisions as taking place along a left-right orientation, with those on the right being less supportive of government intervention in the economy and governmental efforts to redistribute income (both at home and abroad), while those on the left are more supportive of government intervention in the economy and governmental efforts to redistribute income. For example, military aid, which is a positive inducement, is given with the goal being to enhance the military capacity of the recipient nation and scores low on ideological divisions. Military aid scores low on ideological divisions because the left-right divide does not consistently map onto it. American liberals, for example, may support military aid to a state trying to defend itself against an aggressor, like Ukraine in 2022, but oppose military aid to oppressive regimes like Saudi Arabia. Some American conservatives may support military aid because it can enhance American security, but others who are more concerned about government spending may oppose it as wasteful spending. Similar to distributional politics, the positive and negative inducements used for counterproliferation policy by the United States vary in their levels of ideological divisions. Once again, the way that positive and negative inducements are paired together in a dual-track approach means that it's likely that at least one of the in-

ducements is going to create high levels of ideological divisions, making some members of Congress hesitant to support its' implementation.

Finally, the last barrier to the dual-track approach is electoral politics, and specifically which party controls Congress. While this barrier is closely related to the ideological divisions discussed above, both Democrats and Republicans have a range of ideological views when it comes to foreign policy within their parties. When the president's party does not have unified control over Congress, Congress may oppose agreeing to a dual-track approach because they do not want to give the president's party a victory that can help in elections, regardless of whether there are distributional politics or ideological divisions present. Divided government is not an insurmountable barrier for presidents, as they can reach across the aisle to negotiate bipartisan deals with the opposition party. A key to overcoming this barrier, and the others described above, is having high levels of concurrence between the president and Congress.

4.3.3 Concurrence and counterproliferation policy

None of the barriers described above are insurmountable. Presidents can overcome these barriers in a variety of different ways. An obvious way to overcome the barrier of divided government would be for the president's party to win elections and gain unified control over Congress. Presidents can overcome both distributional politics and ideological divisions by being skilled negotiators who are able to strike deals with the opposition party or members of their own party who oppose the dual-track approach. One indicator of the president's ability to overcome the barrier to the dual-track approach is the level of concurrence between the president and Congress. Concurrence is the degree to which the president and a majority of Congress agree on policy outcomes. Concur-

rence is important because it demonstrates the degree to which the president is able to overcome the barriers Congress can enact when it comes to implementing counterproliferation policy. For example, President's Clinton high levels of concurrence in 1993 contributed to his administration's ability to negotiate and begin to implement the 1994 Agreed Framework with North Korea, while the low levels of concurrence that came after the Republican victory in the 1994 midterm elections prevented this dual-track approach from being used from 1995 onwards.

High levels of concurrence can indicate two things that increase the probability of a dual-track approach being implemented. First, it could indicate that the president and their administration is skilled in navigating the issues arising from distributional politics, ideological divisions, and divided government and able to come to some level of agreeance with a majority of Congress. Presidents can achieve this in numerous ways, from backroom negotiations to using the influence of their office to publicly pressure members of Congress to agree with them, or some combination of the two. High levels of concurrence could also indicate that the barriers identified above are not present in a specific case because the president's party control both houses of Congress or there is bipartisan consensus on counterproliferation policy. Either way, high levels of concurrence are going to increase the probability that the president's preferred dual-track approach to counterproliferation policy is implemented because it shows that Congress is likely to agree with the president on a range of policy issues, including counterproliferation policy. While the domestic constraints, and specifically the level of concurrence between the president and Congress, is an important determinant of the United States' counterprolif-

eration policy, I now turn towards discussing the other important determinant of counter-proliferation: the international strategic environment.

4.3.4 International politics and United States' counterproliferation policy

The international strategic environment is an important part of determining the United States' counterproliferation policy. Foreign policy, such as counterproliferation policy, is not made in a vacuum, and states often consider how their policies are going to interact with the existing international environment, and also how allies and adversaries are going to react to these policies. While the United States would almost always prefer to use the dual-track approach because of its effectiveness, it is not always feasible, as some potential proliferators are insulated from being the targets of the negative inducements that make up one half of the dual-track approach. I argue that the most important determinant at the international level for U.S. counterproliferation policy, and what insulates some potential proliferators from being the target of the dual-track approach, is the potential proliferators' relationship with the United States.

Potential proliferators who have had friendly and productive relationships with the United States are insulated from being the targets of a dual-track approach because U.S. presidents are more likely to use positive inducements and less likely to use negative inducements in order to convince them to forego nuclear weapons. This is because presidents are unlikely to want to upset potential proliferators with whom the United States has friendly relationships with by using negative inducements, such as sanctions or military force. Thus, the dual-track approach, which always includes negative inducements, is less likely to be implemented when the potential proliferator and the United States share a friendly relationship. These same incentives do not hold when the potential proliferators

and the United States have an adversarial relationship. In cases where there is a contentious relationship or a history of disputes between the potential proliferator and the United States, U.S. presidents have much more freedom to use both positive and negative inducements in a dual-track approach, because they do not have to worry about upsetting an ally or friendly state with the use of negative inducements.

Fuhrmann and Kreps (2010) demonstrates that states are more likely to consider using or actually using military force, a negative inducement, against another state's nuclear infrastructure if they have had violent conflict in the past, such as a militarized interstate dispute. Building off this finding, I use a MID (or lack thereof) between the potential proliferator and the United States as an indicator of their relationship. This means the United States is more likely to use the dual-track approach with potential proliferators with whom it has had a MID, because there is less of a fear of upsetting an ally or a potential ally by using negative inducements. In combination with the domestic politics discussion above, this leads to the following hypothesis:

H1:When the president has high levels of concurrence with Congress and when facing a potential proliferator with whom the U.S. has had a MID in the past few years, the United States should be more likely to use the dual-track approach of both positive and negative inducements.

4.4 Research design

To test my theory of U.S. counterproliferation policy, I draw on a variety of data sources. My sample includes 31 states[43] who have pursued nuclear weapons and is taken from Mehta (2020: 20-21), which is built off existing lists of states who have pursued nuclear weapons including Bleek (2017) and Mueller and Schmidt (2008). States enter the sample the year they begin their pursuit of nuclear weapons and they exit the sample

[43] See Appendix for the full list of states.

when they end that pursuit. The sample includes states who successfully build nuclear weapons, with those nuclear powers staying in the sample even after they acquire nuclear capabilities because it is still possible that those states might decide to give up their nuclear weapons due to changes in their domestic or international environment. Using a state-year unit of analysis, the temporal domain ranges from 1953 to 2007, and the analysis includes 743 observations.

4.4.1 Dependent variable: dual-track approach

The dependent variable in this analysis is a binary variable called *Dual-track*. *Dual-track* measures whether the United States implements a dual-track approach against a potential proliferator in a given year, and equals 1 when the U.S. uses both positive and negative inducements in the same year.[44] The U.S. implements a dual-track approach in less than 6% (43 out of 743) of the observations in the sample. To determine when the United States uses a dual-track approach, I draw on a variety of data sources, each of which measures when one of the policy instruments is used. For negative inducements, I use Fuhrmann and Kreps (2010)[45] to capture when the United States has threatened or actually used military force against a potential proliferators' nuclear facilities and I use the TIES dataset (Morgan et al. 2014) to measure when the United States implements sanctions against a potential proliferator. For the sanctions, I only include cases where the United States actually imposes sanctions, is the primary sender, and one of the top 3 reasons listed for why the United States implemented sanctions was to "deny strategic mate-

[44] The United States does not have to use all of the positive inducements or all of the negative inducements in a given year for it to count as a dual-track approach. As long as the U.S. uses at least one positive inducement and at least one negative inducement in the same year, I count it as a dual-track approach.

[45] Fuhrmann and Kreps (2010) data ends in 2000. I also code for the United States' (in cooperation with Israel) use of the Stuxnet cyberweapon against Iranian nuclear facilities in 2006.

rials." This arguably underestimates the United States use of sanctions as part of their counterproliferation policy and biases against finding positive results.

Positive inducements include military aid, economic (foreign) aid, political aid, and civilian nuclear cooperation. Following Mehta (2020), I used USAID data to measure when the United States gave military or economic aid to a potential proliferator in a given year (USAID 2019). The U.S. is counted as providing military and economic aid to a potential proliferator if they gave any amount of aid in a given year. To capture when the United States uses political aid, I adopt Mehta's (2020) definition and operationalize political aid as the existence of a defense pact between the United States and a potential proliferator in which the U.S. promises to come to that state's aid if military conflict breaks out. To capture the existence of a defense pact between the U.S. and a potential proliferator I use ATOP's directed dyad-year version 5.1 dataset (Leeds et al. 2002). Finally, to measure when the United States provides civilian nuclear technology to a potential proliferator, I use Fuhrmann's (2009) dataset of civilian nuclear cooperation agreements. Since the dependent variable, *Dual-track,* is binary I use logistic regression models and employ robust standard errors that are adjusted for clustering by state to account for the interdependence of observations.

4.4.2 Main independent variables: Concurrence X MIDs

My theory focuses on constraints at both the domestic and international levels that presidents face when determining the United States' counterproliferation policy. To capture the domestic-level constraints, I use *Concurrence* to measure the president's degree of concurrence with Congress. *Concurrence* is based off a measure from The American Presidency Project which measures the frequency with which a majority of members of

both chambers of Congress vote with the president's position on roll call votes in a given year.[46] High levels of legislative concurrence between the Presidency and Congress indicate that the president and a majority of Congress have a positive relationship since they are taking similar positions on roll call votes. Low levels of legislative concurrence indicate a disconnect between the president and Congress, and make implementing the president's preferred counterproliferation policy difficult. *Concurrence* equals 1 when the level of roll call concurrence between the president and Congress is greater than or equal to 72.4% (the median value of the sample), and equals 0 otherwise.

To capture the international-level constraints, I use militarized interstate disputes (MIDs) from the Correlates of War (Palmer et al. 2020). Using COW's dyadic MID version 4.02 dataset, I focus on MIDs between the United States and the potential proliferators. *MID* equals 1 when the United States has a militarized interstate dispute with a potential proliferator within the past 3 years. 3 years is a reasonable amount of time since the negative feelings caused by a MID are likely to last a few years beyond the end of the dispute.[47] Since my theory argues that the domestic and international-level constraints interact to determine U.S. counterproliferation policy, I interact *Concurrence* with *MID* to generate my main independent variable *Concurrence* X *MID*, which equals 1 when the level of roll call concurrence between the president and Congress is greater than or equal to 72.4% and the potential proliferator has had a militarized interstate dispute with the U.S. in the past 3 years. *Concurrence* X *MID* equals 1 in 95 observations.[48] In the analy-

[46] For 1974, I used the average of Nixon and Ford's level of concurrence with Congress.
[47] I also use MIDs within the past 5 years and within the past 10 years. Both have similar results to using MIDs within the past 3 years. See Appendix for robustness checks.
[48] Not all 95 of these cases make it into the analysis, as I lose a few cases because of the 1 year lag I use for *Concurrence* X *MID*.

sis, I lag *Concurrence* X *MID* by 1 year to ensure that it precedes my dependent variable, *Dual-track*.

4.4.3 Control variables

I adopt a number of control variables that could possibly influence U.S. counter-proliferation policy. First, I control for the regime type of the potential proliferator, as Fuhrmann and Kreps (2010) found that attacks on or the consideration of an attack on nuclear infrastructure is more likely when the potential proliferator is an authoritarian state, thus the dual-track approach is less likely to be used against democratic states. *Democracy* equals 1 when the potential proliferator is a democracy in a given year, and equals 0 otherwise. To determine whether a potential proliferator is a democracy in a given year, I use Polity V data (Marshall and Gurr 2020) and adopt the common approach of coding states as democracies if they receive a polity2 score of 6 or higher.

Additionally, I control for the traditional military capabilities in the U.S.-potential proliferator dyad, using the Correlates of War's (COW) National Material Capabilities data set (Singer et al. 1972). I control for traditional military capabilities because Fuhrmann and Kreps (2010) find that states are more likely to attack or consider attacking the nuclear infrastructure of militarily powerful states because of the greater risk they pose if they do acquire a nuclear weapon, when compared to relatively weaker states. I create a variable, *Capabilities*, that equals the capabilities of the potential proliferator divided by the total combined capabilities of both the potential proliferator and the United States in a given year. Each state's military capabilities are assessed using CINC scores.

I also control for the trade relations between the United States and the potential proliferators. The United States may be unlikely to use negative inducements (the threat

of or use of military force or economic sanctions) against potential proliferators with whom it has significant trade relations, even if the relationship (outside of trade) between that state and the United States is contentious. To determine the level of trade relations between the United States and potential proliferators, I use COW's international trade data (Barbieri and Keshk 2016). Using the variable, smoothtotrade, which measures the total amount of trade between states in a given year, I calculate the median value for the total amount of trade for all dyads involving the United States. I then create *Major trade partner* which equals 1 if the total amount of trade between a potential proliferator and the United States is greater than or equal to that median value.

Finally, I control for which political party the president is a member of. Anecdotally, since the end of World War II, Republicans are generally considered to favor a more aggressive, muscular foreign policy than their Democratic counterparts. This means they may be less likely to implement positive inducements against a potential proliferator with whom the United States has an adversarial relationship with, such as Iran in the 21st century. *President party* equals 1 when a Republican is in the White House.

4.5 Results

To evaluate my hypothesis, I use logistic regression models as well as robust standard errors that are clustered by state (the potential proliferator). I predicted that the dual-track approach is more likely to be implemented when the president has a positive, working relationship with Congress as measured by legislative concurrence, and the potential proliferator has had a MID with the United States in the past 3 years. Table 7 presents the results for the main analysis, with model 1 being a trimmed model that examines the relationship between *Dual-track* and $Concurrence\ X\ MID_{t-1}$. Model 2 is the full model and

includes all control variables. The results across both models indicate some support for hypothesis 1.

Table 7: Logit models with MIDs within past 3 years

	Model 1	Model 2
Concurrence X MID $_{t-1}$	2.40**	2.51**
	(1.12)	(1.22)
Concurrence$_{t-1}$	-0.91**	-0.96**
	(0.40)	(0.38)
MID$_{t-1}$	-0.75	-1.09
	(1.17)	(1.07)
Democracy		-0.24
		(0.99)
Capabilities		0.74
		(3.50)
Major trade partner		-0.60
		(0.61)
President's party		-0.35
		(0.46)
Constant	-2.60**	-1.75**
	(0.51)	(0.81)
Observations	768	743

Standard errors in parentheses, * p<.10, ** p<.05

In model 1, $Concurrence\ X\ MID_{t-1}$ has a positive coefficient and is statistically significant.[49] In model 2, $Concurrence\ X\ MID_{t-1}$ once again has a positive coefficient and is statistically significant.[50] These results largely hold, albeit at higher p-values, when MIDs within the past 5 years and MIDs within the past 10 years are used instead. These results are presented in the appendix. I will discuss the substantive effects of $Concurrence\ X\ MID_{t-1}$ in model 2 below.

[49] The p-value for $Concurrence\ X\ MID_{t-1}$ in model 1 equals 0.032.
[50] The p-value for $Concurrence\ X\ MID_{t-1}$ in model 2 equals 0.039.

4.5.1 Substantive effects for *Concurrence* X *MID*

I assess the substantive effects using *Clarify*, and examine the probability of implementing the dual-track approach for each of the four possible conditions from the interaction between *Concurrence* and *MID*.[51] The results are displayed in Table 8 below. As expected, the probability of implementing the dual-track approach is greatest, at 37.6%, when there are high levels of concurrence between the president and Congress and the potential proliferator has had a MID with the United States. The high levels of concurrence either indicate a president skilled at overcoming the domestic barriers Congress can put in the way of the president's preferred counterproliferation policy or that the barriers are not present, while the presence of a MID makes it acceptable to use negative inducements against the potential proliferator.

Unexpectedly, the second highest probability of implementing the dual-track approach is when there are low levels of concurrence and no MID between the United States and the potential proliferator. After taking a look at the data, the potential proliferator who drives this result is South Africa from 1986 to 1992. South Africa may be an outlier case because of its' domestic policy of apartheid, which led the United States to enact sanctions on it in 1986, even though South Africa was considered a key security partner in containing communism in Africa during the Cold War. This may just be an outlier case or it could highlight a weakness in the theory, specifically the need to consider the domestic policy of the potential proliferator and how that impacts its' relationship with the United States.

[51] See King, Tomz, and Wittenberg (2000) and Tomz, Wittenberg, and King (2003) for more information on *Clarify.* All substantive interpretations are based off Table 3.1 model 2, with all control variables held at their means.

Table 8: Probability of dual-track implementation

Variable	Pr(*Dual-track*=1)	95% CI
Low Concurrence with no MID	0.212	[0.044, 0.516]
Low Concurrence with a MID	0.011	[0.0004, 0.062]
High Concurrence with no MID	0.017	[0.002, 0.067]
High Concurrence with a MID	0.376	[0.065, 0.825]

The expected probability of implementing the dual-track approach is lowest when there are low levels of concurrence between the president and Congress and the potential proliferator has had a MID with the United States. This makes sense, as the low levels of concurrence indicate that Congress is most likely putting up barrier to implementing the dual-track approach. The expected probability is second lowest when there are high levels of concurrence and the potential proliferator has not had a MID with the United States. The expected probability of implementing the dual-track approach is low in this case, because the United States is unlikely to want to use negative inducements (sanctions and the threat or use of military force) against states with whom they have a friendly relationship, as evidenced by the lack of a MID.

4.6 Discussion and conclusion

There are not many clear policy recommendations based off the theory and results presented above. For U.S. presidents interested in countering nuclear proliferation, a rec-

ommendation would include fostering positive relationships with Congress. It would most likely not even be necessary to develop a positive working relationship across all issue areas, but only on foreign policy and counterproliferation specifically. Even this could prove difficult though, as the Iran example from the introduction illustrates. While it was in the best interest of both the Obama Administration and Republicans in Congress to come together and implement a dual-track approach to counter Iran's nuclear ambition, years of hostile interactions between Obama and Republicans made this extremely difficult and contributed to the fragility of the JCPOA. While the Obama Administration was eventually able to marshal the necessary political capital to get this done, it had to be done through an executive order and not through a more durable agreement passed by the Senate.

This project is the first to empirically and theoretically examine when the United States, the most important state in the international fight against the spread of nuclear technology and weapons, is likely to implement the most effective counterproliferation policy, the dual-track approach. The dual-track approach uses both positive and negative inducements in order to convince potential proliferators to forego nuclear proliferation and begin nuclear reversal (Mehta 2020). My theory builds off existing research and assumes that U.S. presidents always favor implementing the dual-track approach, but are constrained at both the domestic and international levels. At the domestic level, presidents are constrained by their relationship with Congress. Presidents with high levels of concurrence with Congress are in a better position to implement their preferred dual-track approach. At the international level, presidents are constrained by the United States' relationship with the potential proliferator. Thus, the dual-track approach is most likely to be

used when the interaction between domestic and international politics results in a situation where the president has a friendly relationship with Congress *and* the U.S. has had a MID with the potential proliferator in the past 3 years.

Future extensions of this project could do a number of things. First, they could develop a more direct measurement of presidential power over foreign policy and not have to extend this logic to argue that presidential power is maximized when the president has a positive working relationship with Congress. Additionally, future research could examine different framings that could be used to convince Congress members and their voters of the need to implement the dual-track approach. This would be important because as this chapter demonstrates, opposition from Congress is one of the key obstacles in preventing the dual-track from being implemented. Understanding what convinces Congress and their voters that this strategy is important, effective, and needs to be implemented would be an important contribution to the literature. A survey experiment presenting different vignettes could test which of these framings is most effective in getting hesitant Congress members and voters to support using the dual-track approach.

CHAPTER 5. CONCLUSION

This research project began by outlining the historical process through which a few states tested nuclear weapons on large-scale, angering the international community and leading to attempts to regulate testing through arms control agreements. The difficulty the international community had in designing effective arms control agreements that would not only regulate nuclear weapons testing, but also nuclear weapons in general, as well as the barriers to implementing the effective dual-track approach to counterproliferation policy by the United States demonstrates how difficult it is to control the spread of nuclear technology. One of the consequences of these failures to control the spread of nuclear technology is the ability of states to use their nuclear arsenals as part of a brinkmanship strategy, in which states take provocative actions to manipulate the shared risk of nuclear war in order to get the less resolved state to back down (Schelling 1966; Powell 2003; Kroenig 2003). In chapter 2, I contributed to this brinkmanship literature by tying nuclear weapons tests, and the hesitancy to ban them through an arms control agreement, to states' use of nuclear weapons tests as part of their brinkmanship strategies. During a crisis, nuclear weapons tests help reveal the balance of resolve between crisis participants, with the state who tests more being more resolved, and thus, more likely to win the crisis. Using data on nuclear weapons tests that occur during crises captures the dynamic, provocative actions that make up brinkmanship strategies.

The next step in this process was figuring out why the international community struggles to control the spread of nuclear technology, especially if they are opposed to some of the consequences that accompany it, such as large-scale nuclear weapons testing. In chapter 3, I argue that even though states spend time and resources on negotiating and

designing arms control agreements, different sources of threat make it difficult for them to design effective, constraining arms control agreements. External threats especially, motivate states to design flexible, less effective arms control agreements that give them greater flexibility in building and acquiring weapons to meet those external threats. Disaggregating data on threats based on the source of the threat lends some evidence to this argument. This research contributes to scholars' understanding about what motivates the design of arms control agreements, which is important because of how the agreement's design influences its' effectiveness, but also because of the findings of previous research, which indicate that the design of arms control agreements influences the likelihood of the agreement being ratified and entering into force (Kreps 2018).

The last step in exploring the historical process that motivated this project is examining what prevents the United States, the leading state in the fight against nuclear technology proliferation, from using the most effective strategy at its' disposal: the dual-track approach. Past research has demonstrated that the dual-track approach of using both positive and negative inducements is the most effective way to convince potential proliferators to engage in nuclear reversal (Mehta 2020). My research in chapter 4 addresses the conditions under which the United States is likely to implement the dual-track approach and argues that U.S. presidents are constrained at both the domestic and international levels. At the domestic level, presidents must overcome Congressional barriers, which include distributive politics, ideological divisions, and electoral politics, to implementing the dual-track approach. Internationally, presidents are constrained by the relationship between the United States and the potential proliferator in that presidents are unlikely to want to use the negative inducements that are part of the dual-track approach

with allies or other friendly states. I find that the dual-track approach is most likely to be used when presidents are able to overcome these Congressional barriers, as measured by the level of legislative concurrence between the president and Congress, and when the potential proliferator has had a militarized interstate dispute with the United States in the past few years. This contributes to scholars' understanding of the conditions under which the United States is most likely to use the most effective counterproliferation strategy. I now revisit some of the implications of this research project, for both policy makers and future research.

5.1 Policy implications

For policy makers, the findings of each of the three empirical chapters leads to a variety of policy recommendations. The findings of chapter 2 demonstrate that nuclear powers are incentivized to test nuclear weapons during crises, which comes with the dangerous consequence of increasing the risk of nuclear war. For policy makers interested in making the world safer, there should be a diplomatic push to get reluctant nuclear powers to sign the Comprehensive Test Ban Treaty in order to prevent these states from using nuclear weapons tests during a crisis, and pushing the crisis adversaries closer to nuclear war. While this is a difficult goal, it could be paired with other concessions to convince nuclear powers to forego nuclear weapons tests.

The findings of chapter 3 also lead to a policy recommendation. In order to mitigate the effects that external threat has on arms control agreement design, states should try and include the sources of external threat in the arms control agreement negotiation and design process so that all parties would be subject to the restrictions in the agreement. This would incentivize states to design more constraining, less flexible agreements that

are going to be more effective in constraining state behavior when it comes to building and acquiring nuclear weapons and other weapons technologies. This would most likely make the negotiation and design process more contentious and difficult, but any agreement that came out of this process would at least have a greater chance of being effective. Finally, there are not clear policy recommendations that arise out of the findings from chapter 4, as it would be difficult for presidents to manipulate the conditions that make implementing the dual-track approach more likely. The best recommendation would be for presidents to use their influence to cultivate positive relationships with Congress, especially on counterproliferation policy. This would allow presidents to overcome the Congressional barriers to implementing the dual-track approach.

5.2 Implications for future research

Future research could also build off this project. Based off the theoretical model and findings of chapter 2, researchers could examine whether other weapon tests, such as missile launches, have a similar effect on the probability of crisis victory as nuclear weapons tests. Other actions, such as military exercises, could also be tied into a brinkmanship strategy, especially when said exercises are simulating the use of nuclear weapons. The findings from chapter 2 could also be tied into the larger bargaining literature, and examine how nuclear weapons tests, as demonstrations of resolve, impact crisis bargaining. There may be other observable implications, besides crisis victory, that come with being nuclear testing superior, such as pursuing a specific bargaining strategy. Chapter 3 could be extended by focusing on the two other aspects of arms control agreement design: precision and delegation. While less important than the level of obligation in determining how effective an arms control agreement is, they still matter when it

comes to constraining state behavior. Future research could also examine how threat influences which states enter into arms control negotiations in the first place. Finally, chapter 4 could be extended in a number of ways. Most importantly, future research could examine different framings that could be used to convince Congress members and their voters of the need to implement the dual-track approach in cases of nuclear proliferation. A survey experiment presenting different vignettes could test which of these framings is most effective in getting hesitant Congress members and voters to support using the dual-track approach.

APPENDICES

APPENDIX 1. Supplemental information for chapter 2

This appendix includes supplemental information from chapter 2, including the list of nuclear crises that were analyzed, cross-tabulations of the main independent variable and dependent variable, and a set of robustness checks that probe the statistical results further.

Table 9 (continued): Nuclear crises, 1945-2016

Crisis Name	**Year**	**Nuclear Participants**
Korean War	1950	Soviet Union*, United States
Suez Crisis	1956	Soviet Union*, United Kingdom, United States*
Berlin Deadline	1958	Soviet Union, United Kingdom, United States
Berlin Wall	1961	France, Soviet Union*, United Kingdom, United States
Cuban Missile Crisis	1962	Soviet Union, United States*
Congo Crisis	1964	Soviet Union, United States*
Six Day War	1967	Soviet Union, United States*
Sino-Soviet Border War	1969	China, Soviet Union
War of Attrition	1970	Israel, Soviet Union
Yom Kippur War	1973	Israel, Soviet Union, United States*
War in Angola	1975	Soviet Union*, United States
Afghanistan Invasion	1979	Soviet Union*, United States
Kashmir	1990	India, Pakistan
India/Pakistan Nuclear Tests	1998	India, Pakistan

Kargil Crisis	1999	India*, Pakistan
India Parliament Attack	2001	India*, Pakistan
Kaluchak	2002	India, Pakistan
N. Korea Nuclear Crisis III	2006	North Korea, United States
N. Korea Nuclear Crisis IV	2009	North Korea, United States
N. Korea Nuclear Crisis V	2013	North Korea, United States
India-Pakistan Border Firing	2014	India, Pakistan
Turkey-Russia Jet Incident	2015	France, Soviet Union, United Kingdom, United States
N. Korea Nuclear Crisis VI	2016	North Korea, United States
Uri Base Attack	2016	India, Pakistan

Source: Hewitt (2003)
Note: The victor in each crisis is denoted by an asterisk. Not all crises have victors, and some have multiple victors.

Table 9 lists all crises featuring dyads where both states are nuclear powers. This is the main sample that is analyzed in Table 2. Table 10 (below) shows that states have only won a clear victory in 27% of nuclear crises between 1945 and 2016. The table also shows that being nuclear testing superior during a crisis greatly enhances a state's probability of crisis victory. Nuclear testing superior states have won 50% of the nuclear crises in which they were involved, compared to only 21% for states that were nuclear testing inferior, and 27% for all crisis participants. The results of a chi-square test allow me to reject the null hypothesis that there is no relationship between nuclear testing superiority and nuclear crisis outcomes. The cross-tabulations presented here show that states who test more nuclear weapons during a crisis than their crisis adversaries are more likely to win nuclear crises.

Table 10: Cross-tabulations of nuclear crises, 1945-2016

		Outcome		
		Win	*Lose*	*Total*
Testing Superiority	Yes	7 (50%)	7 (50%)	14 (100%)
	No	10 (21%)	38 (79%)	48 (100%)
	Total $\boldsymbol{\chi^2}$= 4.6331 (p=0.031)	17 (27%)	45 (73%)	62 (100%)

Table 11 re-runs model 1 from table 2.2 on all crisis dyads. The key variable here is *State A challenger* which controls for whether State A in the dyad initiated the crisis. Since deterrence is easier to achieve than compellence (Schelling 1966), it may be the case that the results from table 2 occur because we are observing only cases of deterrence. The results from this analysis show this is not the case, and even when we account for whether a state initiates a crisis or not, *Test Superiority* is still significant and have a positive coefficient.

Table 11: Controlling for which state initiates the crisis

	Model 1
Test Superiority	0.37*
	(0.21)
State A challenger	-0.65**
	(0.12)
Warhead Superior	0.19
	(0.19)
Violence	0.02
	(0.04)
Gravity	-0.04
	(0.11)
Proximity	0.02
	(0.14)
Capabilities	0.75**
	(0.21)
Population	-0.07*
	(0.04)
Democracy	0.16
	(0.12)
Security	-0.10
	(0.12)
Second Strike	-0.19
	(0.22)
Constant	-0.12
	(0.36)
Observations	810

Standard errors in parentheses, * p<.10, ** p<.05

Table 12 removes the United States from all observations. Since the United States conducted the most nuclear weapons tests of any state, it may be the case that we were only observing U.S.-influenced effects in table 2. Model 1 in table 12 includes crisis dyads where both states possess nuclear weapons, while model 2 includes all crisis dyads. *Test Superiority* is positive and significant in both models.

Table 12: Removing the U.S. from the analysis

	Model 1	Model 2
Test Superiority	8.64**	0.56*
	(2.19)	(0.29)
Warhead Superior	8.59	0.06
	(7.40)	(0.21)
Violence	-1.16	0.01
	(0.84)	(0.04)
Gravity	3.87	0.003
	(2.39)	(0.12)
Proximity	-1.93	-0.10
	(1.47)	(0.14)
Capabilities	30.59*	0.275**
	(15.77)	(0.22)
Population	-4.54**	-0.08**
	(2.26)	(0.04)
Democracy	12.91	0.26*
	(9.28)	(0.12)
Security	5.88	-0.03
	(3.64)	(0.09)
Second Strike	-13.07**	-0.13
	(5.33)	(0.28)
Constant	23.58**	-0.17
	(10.36)	(0.37)
Observations	32	674

Standard errors in parentheses, * p<.10, ** p<.05

APPENDIX 2. Supplemental information for chapter 3

This appendix includes supplemental information from chapter 3, including the list of arms control agreements that were analyzed. The process for coding the agreements is as follows: Identify if negotiation of the treaty was referred to a subset of states (this is usually the case for the large multilateral treaties) or carried out by all states who were involved. For those treaties in which I cannot find which states were apart of the negotiations, I code the states who signed the treaty within a week of when it was opened for signature. I am assuming that if states are willing to sign a treaty immediately when it opens for signatures then they are probably aware of its design and even took part in part of the design. Table 13 includes the year the negotiation over the treaty was initiated, the source(s) I used to find the states negotiating the agreement, as well as any additional notes about my coding process. Some of these agreements are very obscure and never entered into force and so finding information on them was difficult.

Table 13 (continued): Arms control agreements, 1945-2010

Agreement	**Year Initiated**	**Source(s) and additional notes**
1. Baruch Plan	1946	• https://history.state.gov/milestones/1945-1952/baruch-plans
2.COCOM	1949	• Yasuhara, Yoko. 1991. "The Myth of Free Trade: The Origins of COCOM 1945-1950." *The Japanese Journal of American Studies* 4. • Coded the 17 members of the organization.
3.IAEA	1953	• Fischer, David. 1997. "History of the International Atomic Energy Agency: The First Forty Years." *International Atomic Energy Agency,* Vienna, Austria.
4.Limited Test Ban Treaty	1954	• https://www.archives.gov/milestone-documents/test-ban-treaty • Trilateral between the U.S., UK, and USSR.
5.Threshold Test Ban Treaty	1955	• https://2009-2017.state.gov/t/isn/5204.htm.
6.Antarctic Trea-	1956	• https://2009-

ty		2017.state.gov/t/avc/trty/193967.htm.
7.NPT	1959	• https://quod.lib.umich.edu/e/endc/4918260.0001.001/1?page=root;rgn=full+text;size=100;view=image. • Negotiation was referred to Eighteen Nation Committee on Disarmament
8.African NWFZ	1961	• http://disarmament.un.org/treaties/t/pelindaba/deposit/desc • Included all countries who signed on 11 April 1996 plus Russia who signed 5 November 1996.
9.Treaty of Tlatelolco	1961	• https://treaties.unoda.org/t/tlatelolco. • Included all countries who signed on 14 February 1967 plus Nicaragua who signed 15 February 1967.
10.Hotline Treaty	1963	• https://www.armscontrol.org/factsheets/Hotlines.
11.Space Treaty	1964	• https://treaties.unoda.org/t/outer_space. • Included all countries who signed on 27 January 1967 plus Brazil who signed on 30 January 1967.
12.ABM Treaty	1967	• https://www.armscontrol.org/factsheets/abmtreaty.
13.Seabed Treaty	1969	• https://treaties.unoda.org/t/sea_bed. • Negotiation referred to Conference of the Committee on Disarmament.
14.Interim Agreement, US-USSR	1969	• https://www.britannica.com/event/Strategic-Arms-Limitation-Talks#ref117729. • Interim Agreement is part of SALT I talks.
15.Accident Measures Agreement	1969	• https://nuke.fas.org/control/accident/intro.htm. • Difficult to code this one since the treaty was opened for signatures on 5 December 1979 but no states actually signed on that date. I instead code all states who signed before it entered into force in 1984 (11 total).
16.SALT II	1969	• https://1997-2001.state.gov/www/global/arms/treaties/salt2-1.html#1
17.Treaty of Bangkok (Southeast Asia NWFZ)	1971	• https://www.nti.org/education-center/treaties-and-regimes/southeast-asian-nuclear-weapon-free-zone-seanwfz-treaty-bangkok-treaty/.
18.Zangger Committee	1971	• http://zanggercommittee.org/history.html • https://www.iaea.org/sites/default/files/infcirc209.pdf

		• https://www.iaea.org/sites/default/files/publications/documents/infcircs/1974/infcirc209a1.pdf • https://www.iaea.org/sites/default/files/publications/documents/infcircs/1974/infcirc209a3.pdf • https://www.iaea.org/sites/default/files/publications/documents/infcircs/1974/infcirc209a4.pdf • https://www.iaea.org/sites/default/files/publications/documents/infcircs/1974/infcirc209a5.pdf • All of the 5 above documents are letters sent from the 15 original members of the Zangger Committee to the IAEA indicating that they will be acting in accordance with the trigger list developed by the Zangger Committee and the obligations laid out in the NPT regarding the export of nuclear materials. ○ The 15 members are as follows: Australia, Denmark, Canada, Finland, Norway, USSR, U.S., UK, Netherlands, West Germany, East Germany, Poland, Hungary, Czechoslovakia, and Ireland.
19.Moon Agreement	1972	• https://treaties.un.org/Pages/ViewDetails.aspx?src=IND&mtdsg_no=XXIV-2&chapter=24&clang=_en
20.Prevention of Nuclear War Agreement	1973	• https://fas.org/nuke/control/prevent/text/prevent1.htm • Bilateral between US-USSR
21.Conventional Forces in Europe Treaty	1973	• https://www.armscontrol.org/factsheet/cfe • Seems that it was negotiated by all members of NATO and the Warsaw Pact. Could not find a source that explicitly states which states were involved in negotiations. Code the 16 NATO members and 6 Warsaw Pact members who signed.
22.Nuclear Suppliers Group	1973	• https://nuclearsuppliersgroup.org/en/participants1 • Coded first 7 states that were involved.
23.Peaceful Nuclear Explosion Treaty	1974	• https://www.ctbto.org/nuclear-testing/history-of-nuclear-testing/peaceful-nuclear-explosions/ • Bilateral between US and USSR
24.Treaty of Ra-	1975	• https://treaties.un.org/doc/Publication/UNTS

rotonga		/Volume%201445/volume-1445-I-24592-English.pdf • Coded 6 states who signed it. Left out Cook Islands and Niue because they do not reach the population threshold to have a country code in COW.
25.Trilateral Comprehensive Test Ban Negotiations	1978	• https://history.state.gov/historicaldocuments/frus1977-80v26/d162 • US, UK, and USSR
26.Antisatellite Agreement	1978	• https://www.nap.edu/read/11/chapter/7 • Not really sure who was involved here but the best I can find are the US and USSR. The only countries who currently have anti-satellite capabilities are the US, Russia (USSR), China, and India (as of 2019).
27.Convention on the Physical Protection of Nuclear Material	1979	• https://www.iaea.org/sites/default/files/infcirc274r1.pdf • Treaty lists which states participated in the meeting to draft the convention. 58 in total.
28.Intermediate Range Nuclear Material	1979	• https://www.armscontrol.org/factsheets/INFtreaty • Bilateral between the U.S. and USSR
29.Proposed Internationally Legally Binding Security Assurances	1980	• https://www.armscontrol.org/factsheets/negsec • Negative security assurances are when nuclear weapons states promise not to use or threaten to use nuclear weapons against non-nuclear weapons states. Difficult to figure out which states hold them but I have found a 1995 UNSC Resolution (UNSC Resolution 984) in which the P5 members issue some form of negative security assurances. This is what I code.
30.START I	1982	• https://www.armscontrol.org/factsheets/start1 • Agreement was originally signed between the US and USSR in 1991 but when the USSR collapsed the agreement was expanded to include the successor states of the USSR who possessed nuclear weapons: Russia, Belarus, Ukraine, and Kazakhstan. I only code the agreement between the US and USSR because these were the states who negotiated it.
31.Missile Technology Control Regime	1987	• https://www.armscontrol.org/factsheets/mtcr • Coding the members who signed the agreement in 1987 since that is when it was opened for signature.
32.India-Pakistan	1988	• https://www.nti.org/learn/treaties-and-

Non-Attack Agreement		regimes/india-pakistan-non-attack-agreement/ • Bilateral agreement between India-Pakistan
33.START II	1991	• https://fas.org/nuke/control/start2/index.html • Bilateral between U.S. and Russia
34.Lisbon Protocol	1991	• https://www.armscontrol.org/node/3289
35.Mongolia Nuclear Weapon Free Zone	1992	• https://www.nti.org/learn/treaties-and-regimes/nuclear-weapon-free-status-mongolia/ • Unilaterally declared by Mongolia in 1992 so that is the year I use. It entered into force in 2000 when it was recognized as legally-binding by the UN.
36.Fissile Material Cut-off Treaty	1993	• https://www.unog.ch/80256EE600585943/(httpPages)/6286395D9F8DABA380256EF70073A846?OpenDocument • Proposed by the UN Conference on Disarmament which currently has 65 members. It was originally proposed in 1993 but no action has been taken on it. I code members during the year 2010 since this is the last year in the dataset.
37.Comprehensive Test Ban Treaty	1994	• https://www.un.org/disarmament/wmd/nuclear/ctbt/ • https://www.nti.org/education-center/treaties-and-regimes/conference-on-disarmament/ • Treaty was negotiated by the 65 members of the Conference on Disarmament (CD). CD originally consisted of 40 members but additional members were added in 1995 before the conclusion of negotiations over the CTBT so I code all members of the CD.
38.Trilateral Statement	1994	• https://www.armscontrol.org/factsheets/Ukraine-Nuclear-Weapons • Included U.S., Russia, and Ukraine
39.Proposed Nuclear Weapons Convention	1994	• http://inesap.org/sites/default/files/inesap_old/mNWC_2007_Unversion_English_N0821377.pdf • First proposed in 1997 by Costa Rica and subsequently updated by Costa Rica, Malaysia, and the International Campaign to Abolish Nuclear Weapons. I code Costa Rica and Malaysia in the year 2007.
40.Wassenaar Agreement	1995	• https://www.wassenaar.org/about-us/ • 33 original members including the 17 that were part of COCOM

41.Nuclear Terrorism Convention	1996	• https://undocs.org/en/A/RES/51/210 • Diaz Paniagua, Carlos Fernando, *Negotiating Terrorism: The Negotiation Dynamics of Four UN Counter-Terrorism Treaties, 1997-2005* (2008). City University of New York, 2008. Available at SSRN: https://ssrn.com/abstract=1968150 or http://dx.doi.org/10.2139/ssrn.1968150 • The UN established an *Ad Hoc Committee* in resolution A/RES/51/210 to negotiate a Nuclear Terrorism Convention. I cannot find a source that lists which states were apart of the *Ad Hoc Committee* but Diaz Paniagua (2008: 32) states that membership on the *Ad Hoc Committee* was open to all states, even non-UN members, since they would be eligible to be apart of the subsequent treaty. With this being said, I have followed the rule I laid out earlier: I code all states who have signed the treaty within a week of its opening as taking part in the negotiations.
42.START III	1997	• https://www.armscontrol.org/factsheets/start3 • Bilateral between U.S. and Russia
43.Central Asia Nuclear Weapon Free Zone (Treaty of Semei)	1997	• https://www.nti.org/learn/treaties-and-regimes/central-asia-nuclear-weapon-free-zone-canwz/ • Includes 5 states of central Asia
44.Lahore Declaration	1998	• https://www.nti.org/learn/treaties-and-regimes/lahore-declaration/ • Bilateral agreement between Pakistan and India.
45.Strategic Offensive Reductions Treaty (SORT)	2001	• https://www.armscontrol.org/factsheets/sort-glance • Bilateral agreement between U.S. and Russia.
46.Proliferation Security Initiative	2003	• https://www.armscontrol.org/factsheets/PSI • Originally proposed and shaped by 10 states.
47.UNSC Resolution 1540	2004	• https://www.un.org/disarmament/wmd/sc1540/ • Designed and adopted by the 15 members of the UNSC at the time.
48.New Start	2009	• https://www.armscontrol.org/factsheets/NewSTART • Bilateral agreement between U.S. and Russia.

APPENDIX 3. Supplemental information for chapter 4

This appendix includes supplemental information from chapter 4, including a robustness check of the main statistical results. Table 14 model 1 uses MIDs between the United States and potential proliferators within the past 5 years. Similar to the results in the main analysis, $\text{Concurrence X MID}_{t-1}$ has a positive coefficient and is statistically significant. Model 2 uses MIDs within the past 10 years. Once again, $\text{Concurrence X MID}_{t-1}$ has a positive coefficient and is statistically significant. Overall, the results presented here demonstrate that the results presented in the main analysis, Table 4.1, are not contingent on only analyzing MIDs within the past 3 years. The results hold when analyzing MIDs within the past 5 years and within the past 10 years.

Table 14: Logit models with MIDs in the past 5 and 10 years

	Model 1	Model 2
$\text{Concurrence X MID}_{t-1}$	2.49**	2.56*
	(1.23)	(1.34)
Concurrence_{t-1}	-0.96**	-1.03**
	(0.38)	(0.41)
MID_{t-1}	-1.46	-2.11**
	(1.07)	(1.06)
Democracy	-0.33	-0.47
	(0.98)	(0.95)
Capabilities	1.09	1.19
	(3.41)	(3.59)
Major trade partner	-0.81	-1.27**
	(0.62)	(0.62)
President's party	-0.32	-0.41
	(0.48)	(0.47)
Constant	-1.51*	-0.83
	(0.81)	(0.85)
Observations	743	743

Standard errors in parentheses, * p<.10, ** p<.05

REFERENCES

Atwood, Kylie. 2020. "US official says Israel was behind assassination of Iranian scientist." *CNN*, Dec. 2. https://www.cnn.com/2020/12/02/politics/iran-scientist-israel-assassination-us/index.html.

Barbieri, Katherine and Omar M. G. Keshk. 2016. *Correlates of War Project Trade Data Set Codebook*, Version 4.0. Online: https://correlatesofwar.org. Accessed January 2023.

Beardsley, Kyle and Victor Asal. 2009. "Winning with the Bomb." *Journal of Conflict Resolution* 53(2): 278-301.

Betts, Richard K. 1987. *Nuclear Blackmail and Nuclear Balance.* Washington D.C.: Brookings Institution Press.

Bleek, Philipp C. 2017. "When Did (and Didn't) States Proliferate? Chronicling the Spread of Nuclear Weapons." Discussion Paper. Cambridge, MA: Project on Managing the Atom, Belfer Center for Science and International Affairs, Harvard Kennedy School and the James Martin Center for Nonproliferation Studies, Middlebury Institute of International Studies.

Brecher, Michael and Johnathan Wilkenfield. 2000. *A Study of Crisis.* Ann Arbor, MI: University of Michigan Press.

Brecher, Michael, Johnathan Wilkenfield, Kyle Beardsley, Patrick James and David Quinn. 2021. *International Crisis Behavior Codebook, Version 14.* https://sites.duke.edu/icbdata/data-collections/. Accessed 1 September 2021.

Carr, Alan B. 2020. "Of Clouds and Craters: The History of U.S. Nuclear Weapons Testing." *Los Alamos National Laboratory* and *U.S. Department of Energy*. https://www.osti.gov/biblio/1615651-clouds-craters-history-nuclear-weapons-testing. Accessed 2 March 2022.

Cha, Victor. 2013. "North Korea's Third Nuclear Test." *Center for Strategic and International Studies.* https://www.csis.org/analysis/north-korea%E2%80%99s-third-nuclear-test. Accessed on 9 September 2021.

Chiba, Daina, Jesse C. Johnson, and Brett Ashley Leeds. 2015. "Careful Commitments: Democratic States and Alliance Design." *The Journal of Politics* 77(4): 968-982.

CNN. 2001. "U.S. lifts final sanctions on Pakistan." *Cable News Network*. http://articles.cnn.com/2001-10-29/us/gen.us.pak.sanctions_1_pakistan-and-india-sanctions-nuclear-weapons?_s=PM:US. Accessed 18 November 2021.

Coe, Andrew J. and Jane Vaynman. 2020. "Why Arms Control Is So Rare?" American Political Science Review 114(2): 342-355.

Comprehensive Nuclear-Test-Ban-Treaty Organization Preparatory Commission. 2012. "Overview of the Verification Regime." *Comprehensive Nuclear-Test-Ban-Treaty Organization*. https://www.ctbto.org/our-work/verification-regime#:%7E:text=The%20verification%20regime%20of%20the,underwater%20or%20in%20the%20atmosphere. Accessed 7 March 2021.

Cox, Michael. 1990. "From the Truman Doctrine to the Second Superpower Détente: The Rise and Fall of the Cold War." *Journal of Peace Research* 27(1): 25-41.

Davis, James W. 2000. *Threats and promises: The pursuit of international influence.* Baltimore, MD: John Hopkins University Press.

de Villiers, J.W., Roger Jardine, and Mitchell Reiss. 1993. "Why South Africa Gave Up the Bomb." *Foreign Affairs* 72(5): 98-109.
Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41(1): 68-90.
Fuhrmann, Matthew. 2009. "Spreading temptation: Proliferation and peaceful nuclear cooperation agreements." *International Security* 34(1): 7-41.
Fuhrmann, Matthew and Sarah E. Kreps. 2010. "Targeting nuclear programs in war and peace: A quantitative empirical analysis, 1941-2000." *Journal of Conflict Resolution* 54(6): 831-859.
Gartzke, Erik A. and Dong-Joon Jo. 2009. "Bargaining, Nuclear Proliferation, and Interstate Disputes." *Journal of Conflict Resolution* 53(2): 209-233.
Gartzke, Erik A., Shannon Carcelli, J. Andres Gannon, and Jiakun Jack Zhang. 2017. "Signaling in Foreign Policy." In *Oxford Research Encyclopedia of Politics*.
Gelpi, Christopher F. and Michael Griesdorf. 2001. "Winners or Losers? Democracies in International Crises, 1918-1994." *American Political Science Review* 95(3): 633-647.
Gerson, Michael S. 2010. "The Sino-Soviet Border Conflict: Deterrence, Escalation, and the Threat of Nuclear War in 1969." *Center for Naval Analyses*. https://www.cna.org/archive/CNA_Files/pdf/d0022974.a2.pdf. Accessed 22 June 2022.
Gries, Peter Hays. 2014. *The Politics of American Foreign Policy: How Ideology Divides Liberals and Conservative over Foreign Affairs.* Palo Alto, CA: Stanford University Press.
Hansen, Keith A. 2006. *The Comprehensive Nuclear Test Ban Treaty: An Insider's Perspective*. Stanford, CA: Stanford University Press.
Helfer, Laurence. 2012. "Flexibility in International Agreements." In *Interdisciplinary Perspectives on International Law and International Relations*, ed. By Juffrey Dunoff and Mark Pollack, 175-196.
Herring, George C. 2001. *America's Longest War,* 4th ed. Blacklick, Ohio: McGraw-Hill.
Hewitt, J. Joseph. 2003. "Dyadic Processes and International Crises." *Journal of Conflict Resolution* 47(5): 669-692.
Hill, Daniel W. 2016. "Avoiding Obligation: Reservations to Human Rights Treaties." *Journal of Conflict Resolution* 60(6): 1129-1158.
Hoffman, Stanley. 1965. *The State of War: Essays on the Theory and Practice of International Politics.* New York: Praeger.
Horschig, Doreen. 2022. "Israeli Public Opinion on the Use of Nuclear Weapons: Lessons from Terror Management Theory." *Journal of Global Security Studies* 7(2).
Interim Agreement Between the United States of America and the Union of Soviet Socialist Republics on Certain Measures with Respect to the Limitation of Strategic Offensive Arms, art. 8, 1972.
International Physicians for the Prevention of Nuclear War. 1991. *Radioactive Heaven and Earth: The health and environmental effects of nuclear weapons testing in, on, and above the earth.* New York: The Apex Press.

Jervis, Robert. 1984. *The illogic of American nuclear strategy.* Ithaca, NY: Cornell University Press.
Jervis, Robert. 1989. *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon*. Ithaca, NY: Cornell University Press.
Jo, Dong-Joon and Erik Gartzke. 2007. "Determinants of nuclear weapons proliferation." *Journal of Conflict Resolution* 51(1): 167-194.
Johnston, Robert. 2009. "Nuclear Tests—Databases and Other Material." *Johnston's Archive.* https://www.johnstonsarchive.net/nuclear/tests/. Accessed on 2 March 2021.
Kaplow, Jeffrey M. 2022. "State Compliance and the Track Record of International Security Institutions: Evidence from the Nuclear Nonproliferation Regime." *Journal of Global Security Studies* 7(1).
Kershner, Isabel and David E. Sanger. 2015. "Skeptical of Iran Nuclear Deal, Israel Calls for Changes." *The New York Times*, April 6. https://www.nytimes.com/2015/04/07/world/middleeast/israel-iran-nuclear-deal.html?searchResultPosition=1.
Kimball, Daryl. 2020a. "The Nuclear Testing Tally." *Arms Control Association*. https://www.armscontrol.org/factsheets/nucleartesttally. Accessed 2 March 2022.
King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2): 347-361.
Kormenos, Barbara. 2005. "Contracting around International Uncertainty." *The American Political Science Review* 99(4): 549-565.
Kreps, Sarah. 2018. "The Institutional Design of Arms Control Agreements." *Foreign Policy Analysis* 14: 127-147.
Kroenig, Matthew. 2013. "Nuclear Superiority and the Balance of Resolve: Explaining Nuclear Crisis Outcomes." *International Organization* 67: 141-171.
Kroenig, Matthew. 2018. *The Logic of American Nuclear Strategy: Why Strategic Superiority Matters.* New York, NY: Oxford University Press.
Krzyzaniak, John. 2019. "Double Flash: Forty years ago, the Carter administration covered up a presumed Israeli nuclear test." *Bulletin of Atomic Scientists*. https://thebulletin.org/2019/09/double-flash-vela-incident/. Accessed 13 November 2021.
Langner, Ralph. 2011. "Stuxnet: Dissecting a Cyberwarfare Weapon." *IEEE Security & Privacy* 9(3): 49-51.
Leeds, Brett Ashley. 1999. "Domestic Political Institutions, Credible Commitments, and International Cooperation." *American Journal of Political Science* 43(4): 979-1002.
Leeds, Brett Ashley and Burcu Savun. 2007. "Terminating alliances: Why do states abrogate agreements?" *Journal of Politics* 69(4): 1118-1132.
Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long. 2002. "Alliance Treaty Obligations and Provisions, 1815-1944. *International Interactions* 28(3): 237-260.
Lewis, Paul, Sabrina Siddiqui, and Ben Jacobs. 2015. "Republicans fume over Iran nuclear deal but hope for undermining accord is slim." *The Guardian,* July 14.

https://www.theguardian.com/world/2015/jul/14/republicans-iran-nuclear-deal-reaction.
Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: SAGE Publications Inc.
Long, J. Scott and Jeremy Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*, 3rd ed. College Station, TX: Stata Press.
Marshall, Monty G. and Ted Robert Gurr. 2020. *Polity V Project: Dataset User's Manual.* College Park: University of Maryland. https://www.systemicpeace.org/inscr/p5manualv2018.pdf. Accessed January 2023.
Mattiacci, Eleonora and Benjamin T. Jones. 2016. "(Nuclear) Change of Plans: What Explains Nuclear Reversals?" *International Interactions* 42(3): 530-558.
Mattiacci, Eleonora, Rupal N. Mehta, and Rachel Elizabeth Whitlark. 2022. "Atomic Ambiguity: Event Data Evidence on Nuclear Latency and International Cooperation." *Journal of Conflict Resolution* 66(2): 272-296.
Mearsheimer, John J. 1984. "Nuclear weapons and deterrence in Europe." *International Security* 9(3): 19-46.
McGillivray, Fiona and Alastair Smith. 2008. *Punishing the Prince: A Theory of Interstate Relations, Political Institutions, and Leader Change.* Princeton, NJ: Princeton University Press.
Mehta, Rupal N. 2020. *Delaying Doomsday: The Politics of Nuclear Reversal.* New York, NY: Oxford University Press.
Milner, Helen V. and B. Peter Rosendorff. 2001. "The Optimal Design on International Trade Institutions: Uncertainty and Escape." *International Organization* 55(4): 829-857.
Milner, Helen V. and Dustin Tingley. 2015. *Sailing the Water's Edge: The Domestic Politics of American Foreign Policy.* Princeton, NJ: Princeton University Press.
Morgan, T. Clifton, Navin Bapat, and Yoshi Kobayashi. 2014. "The Threat and Imposition of Sanctions: Updating the TIES dataset." *Conflict Management and Peace Science* 31(5): 541-558.
Muller, Harald and Andreas Schmidt. 2008. "The Little Known Story of Deproliferation: Why States Give Up Nuclear Weapon Activities." *ISA Convention Paper.*
Nikitin, Mary Beth D. 2021. "North Korea's Nuclear Weapons and Missile Programs." *Congressional Research Service*. https://crsreports.congress.gov/product/pdf/IF/IF10472. Accessed 13 November 2021.
Nuclear Threat Initiative. 2011. "Strategic Arms Limitation Talks (SALT I)." *Nuclear Threat Initiative*. https://www.nti.org/learn/treaties-and-regimes/strategic-arms-limitation-talks-salt-i-salt-ii/.
Nuclear Threat Initiative. 2014. "Israel Nuclear Overview." *Nuclear Threat Initiative.* https://www.nti.org/analysis/articles/israel-nuclear/. Accessed on 9 March 2022.
Palmer, Glenn, Roseanne W. McManus, Vito D'Orazio, Michael R. Kenwick, Mikaela Karstens, Chase Bloch, Nick Dietrich, Kayla Kahn, Kellan Ritter, Michael J. Soules. 2020. "The MID5 Dataset, 2011-2014: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 39(4): 470-482.
Pelindaba Treaty (African Nuclear-Weapon-Free Zone Treaty), art. 16, July 15, 2009.

Pelindaba Treaty (African Nuclear-Weapon-Free Zone Treaty), art. 17, July 15, 2009.
Pelindaba Treaty (African Nuclear-Weapon-Free Zone Treaty), art. 20, July 15, 2009.
Powell, Robert. 1988. "Nuclear Brinkmanship with Two-Sided Incomplete Information." *American Political Science Review* 82(1): 155-178.
Powell, Robert. 1990. *Nuclear Deterrence Theory: The Search for Credibility*. Cambridge, UK: Cambridge University Press.
Powell, Robert. 2003. "Nuclear Deterrence Theory, Nuclear Proliferation, and National Missile Defense." *International Security* 27(4): 86-118.
Roser, Max and Mohamed Nagdy. 2013. "Nuclear Weapons." *Our World in Data.* https://ourworldindata.org/nuclear-weapons. Accessed on 2 March 2021.
Rhodes, Richard. 1995. *Dark Sun: The Making of the Hydrogen Bomb.* New York, NY: Simon & Schuster.
Sagan, Scott D. 1989. *Moving Targets: Nuclear strategy and national security.* Princeton, NJ: Princeton University Press.
Sagan, Scott D. 1996. "Why do states build nuclear weapons? Three models in search of a bomb." *International Security* 21(3): 54-86.
Sagan, Scott D. and Benjamin A. Valentino. 2017. "Revisiting Hiroshima in Iran: What Americans Really Think about Using Nuclear Weapons and Killing Noncombatants." *International Security* 42(1): 41-79.
Samore, Gary. 2015. "The Iran Nuclear Deal: A Definitive Guide." *Belfer Center for Science and International Affairs,* Harvard University.
Sample, Susan. 2016. "Anticipating War? War Preparations and the Steps-to-War Thesis." *British Journal of Political Science* 48: 489-511.
Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, Mass.: Harvard University Press.
Schelling, Thomas C. 1966. *Arms and Influence.* New Haven, Conn.: Yale University Press.
Schwartz, Stephen I. 1998. *Atomic Audit: The Costs and Consequences of U.S. Nuclear Weapons Since 1940.* Washington, D.C.: Brookings Institution Press.
Scott, Sam. 2017. "What Bikini Atoll Looks Like Today." *Stanford Magazine*. https://medium.com/stanford-magazine/stanford-research-on-effects-of-radioactivity-from-bikini-atoll-nuclear-tests-on-coral-and-crab-dna-48459144020c. Accessed 12 March 2021.
Sechser, Todd S. and Matthew Fuhrmann. 2013. "Crisis bargaining and nuclear blackmail." *International Organization* 67(1): 173-195.
Sechser, Todd S. and Matthew Fuhrmann. 2017. *Nuclear Weapons and Coercive Diplomacy*. Cambridge, UK: Cambridge University Press. Kindle Edition.
Segal, Gerald. 1985. *Defending China.* New York, NY: Oxford University Press.
Senese, Paul D. and John A. Vasquez. 2005. "Assessing the Steps to War." *British Journal of Political Science* 35(4): 607-633.
Signorino, Curtis S., and Jeffrey M. Ritter. 1999. "Tau-b or Not Tau-b: Measuring the Similariy of Foreign Policy Positions." *International Studies Quarterly* 43(1): 115-144.
Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." in Bruce Russett (ed.) *Peace, War, and Numbers*. Beverly Hills: Sage 19-48.

Tannenwald, Nina. 1999. "The Nuclear Taboo: The United States and the Normative Basis of Nuclear Non-Use." *International Organization* 53(3): 433-468.

The American Presidency Project. "House and Senate Concurrence with Presidents." *The American Presidency Project*, ed. John T. Wooley and Gerhard Peters. Santa Barbara, CA: University of California. 1999-2016.

The Wassenaar Arrangement. 2020. "About Us." *The Wassenaar Arrangement. https://www.wassenaar.org/about-us/.*

Tomz, Michael, Jason Wittenberg, and Gary King. 2003. "Clarify: Software for Interpreting and Presenting Statistical Results." *Journal of Statistical Software* 8(1):1-29.

Treaty Between the United States of America and the Union of Soviet Socialist Republics on the Limitation of Strategic Offensive Arms, Together with Agreed Statements and Common Understandings Regarding the Treaty (SALT II), art. 12, June 18, 1979.

Treaty Between the United States of America and the Union of Soviet Socialist Republics on the Limitation of Underground Nuclear Weapon Tests (Threshold Test Ban Treaty), U.S.-U.S.S.R, July 3, 1974. https://2009-2017.state.gov/t/isn/5204.htm. Accessed 3 November 2021.

United Press International. 1961. "Senator Jackson Says U.S. Must Resume Nuclear Tests in Atmosphere: RUSSIANS' BLAST HELD 'BLACKMAIL'." *The New York Times*. https://www.nytimes.com/1961/10/31/archives/senator-jackson-says-us-must-resume-nuclear-tests-in-atmosphere.html?searchResultPosition=9. Accessed 10 March 2021.

United States Department of Energy. 2015. "United States Nuclear Tests: July 1945 through September 1992." *U.S. Department of Energy*. https://www.nnss.gov/docs/docs_LibraryPublications/DOE_NV-209_Rev16.pdf. Accessed 9 September 2021.

USAID. 2019. "U.S. Overseas Loans and Grants Greenbook." Accessed via https://foreignassistance.gov/reports.

Vasquez, John A. 1993. *The war puzzle.* Cambridge, UK: Cambridge University Press.

Walt, Stephen. 1988. "Testing Theories of Alliance Formation: The Case of Southwest Asia." *International Organization* 42(2): 275-316.

Waltz, Kenneth N. 1990. "Nuclear myths and political realities." *American Political Science Review* 84(3): 731-745.

Wellerstein, Alex. 2021. "An Unearthly Spectacle: The untold story of the world's biggest nuclear bomb." *Bulletin of Atomic Scientists*. https://thebulletin.org/2021/11/the-untold-story-of-the-worlds-biggest-nuclear-bomb/. Accessed 30 October 2021.

VITA

1. B.A. in Political Science, B.A. in History, Southern Illinois University-Carbondale, 2017
2. Teaching Assistant, Department of Political Science, University of Kentucky
3. 2022 College of Arts and Sciences Outstanding Teacher Award at University of Kentucky, 2022 Peace Studies Program Outstanding Teacher Award at University of Kentucky, 2022 Malcolm E. Jewell Outstanding Graduate Student Award Honorable Mention at University of Kentucky
4. Bryce Allen Beschorner