RICE UNIVERSITY

Publicly-Conducted Missile Tests in International Politics

by

Cem Birol

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

T. Clifton Morgan, Chair
Albert Thomas Professor of Political Science

Ashley Leeds
Professor of Political Science, Department Chair

Songying Fang
Associate Professor of Political Science

Hülya Eraslan
Ralph O’Connor Professor of Economics
ABSTRACT

Publicly-Conducted Missile Tests in International Politics

by

Cem BIROL

This dissertation investigates the role of publicly-conducted missile tests in international politics. The first paper introduces a new dataset including 11 countries’ missile tests between 1949 and 2015. It also overviews these countries’ indigenous missile development programs. The second paper empirically investigates what factors may affect changes in how frequently countries test-fire their missiles publicly. The results suggest that a country’s public missile tests are intrinsic to missile development programs and transpire more frequently with greater economic wealth. On the other hand, being under a military threat, major or severe economic sanctions have either weak or no consistent impact on countries’ public missile tests. The third paper investigates whether conducting public missile tests is associated with countries’ subsequent militarized interstate dispute, or war involvement. Statistical analyses show that there is little or no evidence that suggests that public missile tests have a discernible impact on countries’ militarized dispute or war involvement. Ultimately, this dissertation stands as the first large-N statistical work on publicly-conducted missile tests. In addition, it is the first one to empirically downplay the potential relationship between militarized conflicts and public missile tests.
Acknowledgments

I would like to acknowledge several people who provided me with the intellectual and mental support while I developed my dissertation. First, I would like to thank T. C. Morgan, for accepting to be the chair of my dissertation committee, his patience with me, and his intellectual guidance since my first day at Rice University. Second, I would like to thank Brett Ashley Leeds, and Songying Fang for generously taking their time to carefully read each draft of my work and giving me extremely valuable feedback. Third, I would like to thank Prof. Hulya Erarslan for accepting to be in my committee. Fourth, I would like to thank Johnathan Kurt Simmons, a dear friend and colleague, for giving me great advice in developing my arguments, and similarly Andrew R. Wood for helping me with the editing my work. Fifth, I would like to thank my mother Tulin Birol, and my brother Erim Birol for giving all the mental support I needed during my studies at Rice University. Sixth, I would like to thank my father, Huseyin Hulki Birol, for being my main inspiration behind my passion for academic research, and my grandfather, Nevzat Ayaz for being my main inspiration to study military affairs and international politics. Lastly, I dedicate my dissertation to my late aunts Birgul Bute and Gunseli Sutekin both of whom closely watched my progress while I pursued my PhD, had confidence in me more than anyone I know (because they were the only two people who called me a Dr. before I even dared to dream of my defense day), and who hopefully are still watching me from the heavens…
# Contents

Acknowledgments ...................................................................................................................... ii
Contents ......................................................................................................................................... iii
List of Figures ............................................................................................................................... vi
List of Tables ............................................................................................................................... vii
Introduction ................................................................................................................................. 8
  Plan of the Dissertation ............................................................................................................. 14
  Summary of the Results ........................................................................................................... 17
  What will we be learning? ........................................................................................................ 18
  References for Introduction .................................................................................................... 20

**Paper I: Public Missile Tests – A New Dataset** ..................................................................... 23
  Introduction ............................................................................................................................... 23
  Defining Missile Tests ............................................................................................................. 26
  Available Data on Missile Tests ............................................................................................ 29
  Dataset ..................................................................................................................................... 31
    (ii) Coding: .......................................................................................................................... 35
    (iii) Variables: ...................................................................................................................... 37
  Country Profiles ....................................................................................................................... 46
    (i) United States: ............................................................................................................... 46
    (ii) USSR & Russia: .......................................................................................................... 51
    (iii) China: .......................................................................................................................... 55
    (iv) India: ............................................................................................................................ 57
    (v) Pakistan: ....................................................................................................................... 60
    (vi) North Korea: .............................................................................................................. 63
    (vii) Iran: ............................................................................................................................. 65
    (viii) Iraq: ............................................................................................................................ 67
    (ix) Israel: ............................................................................................................................ 69
    (x) The United Kingdom: ................................................................................................. 71
    (xi) France: .......................................................................................................................... 74
  Discussion: ............................................................................................................................... 77
References for Paper I: ...........................................................................................................81
APPENDIX for Paper I ..............................................................................................................88
  (1) Missile Aspects by Countries (1949-2015) .................................................................88
  (2) Missile Names by Country ............................................................................................89
  (3) Cross-Correlation Plots .................................................................................................94

Paper II: Why do countries test missiles and missile-related technologies? ......................96
What do we know of missile tests so far? ...............................................................................98
Argument 1 – Missile Tests as a Necessary Part of Developing New Missiles. ..................100
Argument 2 – Missile Tests to Respond to Military Threats .............................................102
Argument 3 – Missile Tests as Means to Cope with Sanctions ........................................108
Research Design (Sample and Variables) .............................................................................110
Statistical Analyses ..............................................................................................................125
Robustness Checks ..............................................................................................................137
  Crises instead of MIDs: ...................................................................................................137
  Country Sensitivity: .......................................................................................................139
  Zero-Inflated Negative Binomial Regression: ...................................................................141
  Alternative outcome variables: .........................................................................................142
Conclusion and Discussion .................................................................................................146
References for Paper II: .....................................................................................................150
Appendix for Paper II ...........................................................................................................156

Paper III - Missile Tests, Militarized Interstate Disputes and Wars ...................................158
Theoretical Argument .........................................................................................................161
Missiles and Missile Tests ..................................................................................................163
First Strike Advantages .....................................................................................................165
RESEARCH DESIGN .......................................................................................................168
  (1) MID Onset: ..............................................................................................................168
  (ii) MID Escalation to War: ........................................................................................181
STATISTICAL ANALYSES .............................................................................................185
  (i) MID Onset Results: .................................................................................................185
  (ii) MID Escalation: ....................................................................................................194
Robustness Checks for Paper III: ......................................................................................200
  (i) Rare events Logit: ...................................................................................................200
  (ii) Heatmap Analysis: .................................................................................................200
List of Figures

Figure 1: Public Missile Tests by Country ................................................................. 45
Figure 2: American Missile Tests (1949-2015) ....................................................... 50
Figure 3: USSR and Russian Missile Tests .............................................................. 51
Figure 4: Chinese Yearly Missile Tests ..................................................................... 55
Figure 5: Yearly number of missile tests by India .................................................... 58
Figure 6: Yearly Missile Tests of Pakistan ................................................................. 61
Figure 7: Yearly Number of North Korean Missile Tests .......................................... 64
Figure 8: Yearly Missile Tests of Iran .......................................................................... 66
Figure 9: Yearly missile tests of Iraq .......................................................................... 68
Figure 10: Yearly Missile Tests of Israel ....................................................................... 70
Figure 11: Yearly Missile Tests of England .................................................................. 72
Figure 12: Yearly Missile Tests of France .................................................................... 74
Figure 13: A Comparison of the American and Russian Missile Tests (1949-2015) .... 78
Figure 14: Yearly Missile Test Trends of India vs its Adversaries .............................. 80
Figure 15: US vs USSR/Russia .................................................................................. 94
Figure 16: India vs Pakistan ....................................................................................... 95
Figure 17: India vs China ............................................................................................ 95
Figure 18: Yearly Missile Tests (1) ............................................................................. 113
Figure 19: Yearly Missile Tests (2) ............................................................................ 113
Figure 20: Yearly Missile Tests (3) ............................................................................ 114
Figure 21 Substantive Effects of Various Forms of MID involvement on Missile Tests .... 131
Figure 22: Country Sensitivity Analyses for Models 1 and 3 of Table 2 .................... 140
Figure 23: Country Sensitivity Analyses for Model 1 of Table 2 (New Missiles) ........ 141
Figure 24: Missile Test Variables (Frequency Histograms) ......................................... 156
Figure 25: Substantive Effect Analyses for Model 1 of TABLE 10 ............................ 191
Figure 26: Heat Map Plot of Model 1 of TABLE 9 ..................................................... 222
Figure 27: Heat Map Plot of Model 1 of TABLE 10 .................................................... 222
Figure 28: Heat Map Plot of Model 1 of TABLE 11 ................................................... 223
Figure 29: Country Sensitivity Analysis for Model 1 of TABLE 9 .............................. 224
Figure 30: Country Sensitivity Analysis for Model 1 of TABLE 10 ............................ 224
Figure 31: Country Sensitivity Analysis for Model 1 of TABLE 11 ............................ 225
List of Tables

Table 1: Aspects of the Missiles Tested by Country ................................................................. 88
Table 2: Negative Binomial Estimates of Missile Tests (Hypothesis 2) ........................................ 129
Table 3: Negative Binomial Estimates of Missile Tests (Hypotheses 3 - 4) ................................. 135
Table 4: Negative Binomial Estimates of Already Commissioned Missile Tests (Hypothesis 5) ................................... 136
Table 5: Negative Binomial Estimates of Missile Tests ................................................................. 139
Table 6: Different Missile Types Estimated .................................................................................. 145
Table 7: Summary Statistics ........................................................................................................ 147
Table 8: Summary Stats of Missile Test Variables by Country (N=36,777 by country) ..... 175
Table 9: Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 1) ................................. 187
Table 10: Logit Estimates, being targeted in a MID vs Missile Tests (Hypothesis 2) ............... 190
Table 11: Logit Estimates, MID Escalation to War (Hypothesis 3) ............................................. 196
Table 12: Heckman Probit Estimates of MID Escalation to War .............................................. 198
Table 13: Descriptive Statistics (MID Involvement Models) ....................................................... 214
Table 14: Descriptive Statistics (MID Escalation Models) .......................................................... 215
Table 15: Bivariate Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 1) ............... 216
Table 16: Bivariate Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 2) ............... 216
Table 17: Bivariate Logit Estimates, MID Escalation to War (Hypothesis 3) ........................... 217
Table 18: Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 1) ............................... 218
Table 19: Logit Estimates, Being targeted in a MID vs Missile Tests (Hypothesis 2) ............... 219
Table 20: Rare Events Logit Estimates (Hypotheses 1, 2, and 3) ................................................. 220
Table 21: International Crisis Onset and Escalation Analyses, Probit Estimates ....................... 227
Introduction

In 2005, Pakistan publicly test-fired its first cruise missile named Babur. Babur was reportedly reverse engineered from a Tomahawk missile that Pakistanis salvaged from Afghanistan in early 2000s (BBC 2005, Indo-Asian News 2009, UPI 2011). The public test of Babur took place several months after Pakistan’s traditional rival India test-fired its BrahMos cruise missile which was jointly manufactured by India and Russia (AFP 2004). In 2007 Russian Premier Vladimir Putin ordered the testing of Topol’-M naval-based ballistic missiles. These were upgraded versions of the Cold War-era Topol’ missiles (Isachenkov 2007). In 2017, North Korea’s missile tests were alarming to the rest of the world. Almost each month, the so-called “hermit” nation test-fired one or more missiles and publicized it each time – a pace greater than previous years (Missile Defense Project – North Korea). Towards the end of 2017, North Korea was able to test fire its first nuclear-capable Intercontinental Ballistic Missile (Missile Defense Project – Hwasong 15, Zachary et al 2017).

In all these incidents, a country is observed to have publicly test-fired missiles of its own. These tests are not isolated cases. Public missile tests frequently transpire in various formats. For example, countries often test-launch a missile without a specific target to assess the stability of its booster system and/or the maximum range and apogee that it can reach. These are typically the missile tests that North Korea conducts. Sometimes, countries employ specific mock targets during tests and focus on evaluating the accuracy of a missile. These targets can be replications of enemy technologies. For example, Iran once built a mock ship that resembled USS Nimitz (an aircraft carrier) against which it
launched a short-range missile interceptor – Nasr (Geller 2015). China in 2007 launched an intermediate range missile to destroy an old satellite that was in higher orbit than any American surveillance satellite (Facts on File 2007). Similarly, in 2019, India test-launched a missile to destroy its satellite (BBC 2019). These tests can vary in the level of success: a missile can explode in the launching pad due to failure in the booster system, can hit a friendly target due to problems in the navigation system, or can actually hit the desired target, travel the expected distance, et cetera. Ultimately, missile tests are the most direct way that countries understand how their missiles are performing, and whether it can be used reliably in a combat situation.

In general, a country’s public missile tests are not welcome by other countries. After all, missiles can be highly destructive in warfare, and when a country tests such destructive weapons, its neighbors, and potential adversaries can rightfully be concerned. Pakistan’s Babur test was heavily criticized by India. Similarly, India’s public missile tests have almost always been protested by Pakistan. Russia’s Topol M’ tests were not welcomed by the United States and its allies. South Korea, Japan, the US, many NATO countries and even Russia and China often criticized North Korea’s missile test launches. Given how frequently countries test-fire their missiles, a rightful question is what the role(s) of these unwelcome tests is/are in international politics.

The increase in North Korea’s missile tests made this topic popular in global politics. However, since the beginning of the Cold War, countries have been publicly test-firing their missiles and missile-related systems. Despite how frequent public missile tests have been, media and policy world have paid greater attention to these incidents than the academia. News pieces typically narrated Soviet missile tests as provocative or defiant
against the US in a way similar to how more recent North Korean or Iranian missile tests were covered. Often times, the question was whether these tests would bring the world or the relevant countries on the brink of war, or at least hostile military encounters. Yet, these questions have often been short of an answer that was based on systematic/scientific research.

On the other hand, the scholarship in international politics has not shown much enthusiasm for public missile tests. The focus was rather placed on the possession of lethal weapons, nuclear technology and how this affected relations between countries. For example, many scholars sought to understand how the presence of nuclear weapons shaped countries’ approaches to solving their conflicts (Gartzke and Jo 2009, Geller 1990, Kahn 1960, Powell 1987, Schelling 1960, 1966, Seschser and Fuhrmann 2013). Most of this research concluded that nuclear weapons would remain in their silos and countries would rather use their nuclear possessions to secure better crisis bargaining deals from their opponents. In the end of the day, nuclear warheads were not placed on giant ICBMs and launched against adversaries. However, countries still kept showing to the rest of the world how their missiles performed. So, these destructive weapons flew all over the skies for the world to witness. It was simply the case that they did not land against an enemy, but rather against mock targets. Interestingly, no scholar has investigated whether test-firing missiles would have a different impact than just possessing it.

Overall, there seems to be two different factions with two different attitudes on missile tests among nations. The first is the policy world that is highly alarmed when a rogue nation such as Iran or North Korea test-fires a destructive missile. The second is the academia that remains highly uninterested in missile tests. The major problem is that we
do not know which side is right with its general attitude. Should we be concerned with missile tests because they potentially bring countries closer to militarized conflicts and war? Should we pay little attention since they are a triviality in global politics. The answer is not easy to come up with. This is mainly because we do not have any large-N data on missile tests to derive generalizable patterns, and to see which faction (policy field or academia) is right.

Why do we need data on missile tests? The first reason is that a general understanding on this topic is long overdue. Public missile tests are not confined to Soviets, Iran, or North Korea. Today, more countries keep trying to develop their indigenous missile systems and also test-fire their own missiles. A recent example is Turkey, which has placed much greater focus than usual on developing its own missiles. In the last two or three years, Turks test-fired several short to medium range missiles of their own.\(^1\) At the same time, Turkey is at odds with many of its neighbors as well as the United States. It is not very likely that a North Korea expert will be able to present sound arguments on Turkey, which is a completely different nation from the former that is situated on a different geography. More recently, India destroyed a satellite with a ballistic missile. This event took place almost a month after the country was in a lethal border skirmish with Pakistan (BBC 2019).

How do we interpret Turkish or Indian missile tests? One possibility is to try to become an India or a Turkey expert. However, this will not be helpful if Greece -which recently test-fired its Russian-based S-300 missile interceptors or England which plans to leave European Union becomes chronic missile testers like the Cold War Soviets or the current North Korea.

---

Accordingly, we need to establish generalizable patterns regarding missile tests and their role in international politics. This approach will be at the expense of our ability to capture specific details of a particular country. However, the benefit will be our ability to make informed speculations whenever a new case of missile test occurs. To be able to develop a general understanding of the missile tests in global politics, we need to pursue a scientific method. In doing so, first, we need to figure out what arguments about the causes and consequences of missile tests are there (or can be made). Second, we need to test these arguments against data on a global scale. Therefore, an appropriate data collection effort is a particularly crucial step in developing a general understanding on the role of missile tests in global politics.

The second reason we need data on missile tests is the policy implications that can be obtained through statistical analyses. On one hand, missile tests are overlooked in the academia. Then a possibility is that scholars consider missile tests as trivial, non-consequential events in international politics. On the other hand, pundits and policy experts often call attention to how missile tests can substantially increase hostilities among nations. If the statistical analyses support pundits’ expectations, then policy-makers may be better off shifting resources to prevent countries from test-firing missiles. To develop preventive policies, we also need to know what factors consistently decrease or increase the missiles that countries publicly test-fire. That way, we can better know what to avoid, and what to focus on if the goal is to make a country test-fire fewer or no missiles. If, however, statistical analyses reveal that missile tests are almost non-consequential in international politics, then policy-makers may rather be better off decreasing the resources they allocate to prevent others’ missile tests.
Consider the case of North Korea as an illustration for how statistical analyses can generate important policy implications regarding missile tests. Recently, North Korea received UN-based severe sanctions. The reason behind some of these sanctions were to prevent North Korea from test-firing missiles. The UN’s anticipation was that if North Korea continued testing more missiles, this would contribute to its ballistic capabilities and make it a more dangerous opponent in East Asia. Yet, what if North Korea’s missile tests are not making this country more dangerous in global affairs? Alternatively, what if missile tests are actually particularly dangerous events, but sanctions are found to have no significant effects on missile tests? It is only through a rigorous data collection effort, and statistical analyses that can we answer these questions – and create better policies.

Accordingly, in this dissertation, I investigate two main questions with the goal of contributing to our understanding of the roles of missile tests in world politics: first, why countries sometimes publicly test-fire more missiles than other times? Second, do the missile tests of a country increase the risk of its involvement in militarized disputes and ultimately wars?

To understand the roles of missile tests in international politics, the first step ought to consist of an attempt to clarify what motivations countries have in test-firing missiles publicly. Is it because they worry about their security and that they seek to intimidate their potential adversaries? Alternatively, is it because missiles often cannot be commissioned for service without adequate testing? Yet, why then are these tests still conducted in public?

After identifying potential reasons behind missile tests, the next step becomes the investigation of their consequences. Since missiles tend to be very effective tools of warfare, a reasonable question is whether more public missile tests also increase the risk
of wars among nations? Alternatively, do missile tests have a deterrent effect which actually reduces the risk of militarized conflicts and wars? Lastly, is it possible that missile tests are related to some level of militarization in interstate conflicts but fail to account for wars the latter which may depend on a complete different set of antecedents?

Upon presenting a set of answers regarding the reasons behind, and the consequences of missile tests, I conduct a series of large-N statistical analyses. These analyses compare and contrast different competing or complementary arguments behind missile tests and their role in international politics. More importantly, in this dissertation, I present the first large-N panel data on the missile tests of several countries. Thus, to my knowledge, this dissertation presents the first ever attempt to generalize what we know on missile tests. There will be much more work to be done to improve our understanding on this topic. However, the availability of a workable dataset will hopefully enable more subsequent work.

In what follows, I detail the plan of this dissertation, and then conclude by highlighting how it contributes to our understanding of missile tests, and global security

Plan of the Dissertation

This dissertation has a 3-paper format. In other words, each chapter represents a separately-developed paper. The additional conclusion chapter presents a discussion of the all findings, and what they imply all together. Accordingly, the first paper (Paper 1,

---

2 There is only one exception. The first paper fully develops and details the dataset. Hence, I treat it as a data paper. The second and third papers are treated as papers that build on the data paper. In other words, instead of rewriting all the coding procedures and decisions of the missile tests dataset in Papers 2, and 3, I refer the readers back to Paper 1.
hereby) introduces my data collection project on missile tests. The dataset basically consists of public missile tests conducted by 11 countries - the US, USSR/Russia, England, France, Israel, Iran, Iraq, China, North Korea, India and Pakistan between years 1949 and 2015. All these countries have their indigenous missile development programs and appear to be the most publicly missile-testing countries.³ For each of these countries, I specifically record the exact location, day, month, and year of the tests of the missiles and missile-related technologies.⁴ I also present information regarding the specific features of the missiles such as range, whether it was commissioned at the time of the test or not, or whether it was nuclear-capable or not.

In the first part of the Paper 1, I detail the coding procedure of the dataset. This part includes detailed description of the sources I use as well as the coding decisions I make. I conclude this first part with simple statistics of the whole dataset. In the second part, I present an overview of the 11 countries. This overview involves a plot for each country regarding its yearly number of public missile tests based on the data I collected. I also present a brief summary of the missile development and testing histories of each country.

Paper 1 concludes with some generalizable observations from the yearly missile testing trends of each country in the sample. Overall, a crude look into missile test trends shows that countries’ peak number of missile tests often coincide with the periods of their major international conflict involvement. This observation suggests that we ought to look deeper into the relationship between a country’s conflict involvement and missile tests. Naturally, absent the data I collected, we would not be able to trace how increases in missile

³ A more detailed reasoning for the countries of choice is provided in Paper 1.
⁴ More specifically, I include tests of nuclear warheads, ballistic and cruise missiles of all ranges, missile interceptors and space launches where a military ballistic missile was employed.
tests coincide with countries’ major conflicts. Yet, more rigorous statistical analyses are needed to uncover the story behind missile tests and international conflict.

The second and third papers present multivariate statistical analyses to delve into the suspected association between countries’ missile tests and their conflict involvement. As I highlighted earlier, the first step in understanding the role of missile tests in the world politics is to investigate what factors tend to drive countries to test-fire missiles publicly. In that regard, in Paper II, I lay out three broad set of empirical expectations. The first expectation is that increases in missile tests should follow introduction of new missile development programs. The second expectation is that increases in a country’s missile tests should follow the military threats it receives in various forms. Finally, I expect that the imposition of economic sanctions on a country should not lead to a decrease in the total number of its missile tests, but actually an increase in the test-launches of its already commissioned missiles. In order to test these expectations, I conduct a series of statistical analyses using Negative Binomial models.

Paper III of this dissertation investigates whether the public missile-tests of a country are conducive to its militarized interstate dispute and/or war involvement. Pundits and policy experts often perceive missile tests of countries such as North Korea, Iran, China or Russia as destabilizing acts that increase the risk of militarized conflict or war onsets. Accordingly, relying on Jervis’ (1976) “security dilemma” notion, I explain why public missile tests of a country can potentially make other countries and itself (mis)perceive as insecure. Then, I argue that because of the “first strike incentives” created by missiles, the perception of threat created by missile tests is likely to be conducive to interstate disputes
and wars. In evaluating these expectations, I conduct a series of multivariate Logit and Heckman Probit models.

**Summary of the Results**

In Paper 2, only the introduction of new missile development programs seems to consistently increase subsequent public missile tests. I find that missile tests are only weakly driven by military threats that a country receives. I consider three factors that can present a military threat to a country: its previous militarized dispute involvements, its adversaries’ joint military exercises and missile tests. Of all these three factors, only previous involvement in militarized interstate disputes have a weakly positive impact on countries’ subsequent missile tests. Last but not least, economic sanctions do not have a consistent impact on missile tests. The core findings in Paper 2 are such that countries typically test-fire more missiles if they have a strong economy, and shortly after they introduce new missile programs. Yet, trying to intimidate a nation by imposing of economic sanctions will probably not change how many missiles it tests fires in a year. Also, conducting joint military exercises to intimidate an opponent will typically not backfire in the form of this opponent responding with an increase in the number of missiles it test-fires.

In Paper 3, the statistical analyses do not support the anticipation that increases in countries’ missile tests are associated with their risk of being subsequently involved in militarized conflicts and/or wars. While there is a positive effect of missile tests on increasing a country’s subsequent likelihood of involving in militarized interstate disputes, this effect is very weak, and sensitive to sample specifications. On the other hand, I find no support for the expectation that countries that test-fire more missiles are also more likely
to escalate their subsequent militarized disputes to wars. These findings remain more or less the same when I use a Heckman Probit model to account for empirical selection problems inherent in the data.

**What will we be learning?**

This dissertation aims to contribute to our understanding in two major ways. The first is its introduction of a dataset on missile tests. As argued earlier, absent a general understanding of the role of missile tests in international relations, it is difficult to interpret a new case in which a country test-fires a missile. We can always be an expert of this particular country. Yet, in that case, we will not have much rigorous grounds to speculate on other countries’ missile tests. Thus, the goal of this dissertation is not simply to shed light on why North Korea increased its missile tests in 2017, but also to be able to make informed guesses if Russia, India, Pakistan, China, the US, Israel or many such countries test-fire missiles of their own. The introduction of a dataset on missile tests in a global scale will be instrumental to fulfill this goal.

Additionally, the introduction of a new dataset on missile tests can potentially pave the way for subsequent theoretically rigorous studies on this topic. It is often the case that the introduction of a dataset on a certain topic in international relations increases the number of subsequent publications in this topic. For example, the Correlates of War Project (Small and Singer 1982) facilitated more studies on war and enabled many theories on wars to be tested empirically. The same happened when datasets on economic sanctions (Hufbauer et al 1990, Morgan et al 2009), or civil conflict (Gleditsch et al 2002) were introduced. In fact, the absence of a dataset on missile tests might have been a reason why scholars have thus far overlooked this topic. Accordingly, the introduction of my dataset
can pave the way for future scientific studies that can shed light on countless questions on countries’ public missile tests.

Secondly, this dissertation aims to answer policy relevant questions on missile tests which have not yet been addressed in a scientific approach. With data to conduct statistical analyses, the task becomes identifying questions and puzzles pertaining to missile tests. We keep reading news of North Korea or Iran test-launching missiles. These tests typically worry other nations. Most North Korean tests are met with demarches by not only Japan, South Korea or the US, but also China, Russia, the EU, Turkey and even Iran. Same goes for the tests of Iran, China or many other countries. It goes without saying that public missile tests are not welcome. Hence, one of the important questions that ought to be answered is why countries take these actions. Why do countries not keep their missiles in silos, but publicize the test-launches of these missiles? In doing so, are they trying to intimidate others? Are they preparing for wars? Alternatively, are these tests routine events under missile development programs? If, the last explanation is the case, then we may bet better off shifting our attention to the question of why countries feel the need to acquire or develop new missiles. The analyses in this dissertation basically shed light on these questions.

Lastly, there is always the question of whether missile tests are conducive to war. As I laid out before, academia is silent on this topic whereas pundits and policy-makers typically sound the alarms when a country test-fires a missile. Do the latter’s concerns have merit? Will for example Iran become more likely to attack civilian and military ships in the Persian Gulf after test-firing more cruise missiles in the Strait of Hormuz? Or are there reasons other than missile tests that make Iran more aggressive? Similarly, should we
worry when Russia test-fires missiles from supersonic gliders in space? Will these tests that the American President Trump and Vice President Pence worry about, increase the risk that Russia will attack others and cause danger to the rest of the world? One of the goals of this dissertation is to basically present multivariate statistical analyses and show whether missile tests are conducive to militarized conflict when alternative factors that can increase countries’ risk of conflict involvement are accounted for.

In the end of the day, it is always possible that countries’ missile tests do not make them more aggressive, and they do not test missiles to subsequently attack others with these missiles anyway. There are many scholars who would not be surprised with this possibility. Yet, we do not know if this is the case. Policy-makers’ attempts to curb countries’ missile tests with heavy sanctions or military operations are not inspired by a long line of academic work that already established enough knowledge on this topic. On the contrary, we do not know enough on missile tests to make reliable predictions on the consequences of these acts. This dissertation accordingly can be considered as the first step in establishing a more rigorous and general understanding on missile tests.

References for Introduction


Chin, Jeremy (2018) "Turkey Unveils Long-range Air Defense Missile System," Missile Threat, Center for Strategic and International Studies. Published: October 31, 2018,


Paper I: Public Missile Tests – A New Dataset

Introduction

World War II concluded with the US dropping two atomic bombs over Hiroshima and Nagasaki in 1945. It was the only time a country used nuclear weapons against another in a war. Shortly after the end of World War II, the two main victors of the war - the US and the USSR - hired German scientists and started developing their own destructive missiles that quickly dwarfed those that were used in World War II. The dominant scholarly approach during this post-World War II era was to investigate how simply possessing these nuclear-capable missiles influenced the risk of war (Blainey 1973, Kahn 1960, Schelling 1960, 1966). In the meantime, an important detail was overlooked: these missiles did not simply sit idle in their silos. They were often publicly test-launched. More importantly, neither the US nor the Soviets welcomed each other’s missile tests.

While the US and the USSR enjoyed supremacy in ballistic and nuclear technologies, several other countries started to develop and test their own missiles. From the 1950s on, France and (to a lesser degree) England test-launched some of their own missiles (see Encyclopedia Astraunomica - MSBS). In the early 1960s, China joined the club by launching its DF-class ballistic missiles (Encyclopedia Astraunomica - DF 1, China). Starting in the 1970s, news broke of countries like India and Israel launching their own missiles (Encyclopedia Astraunomica – India, - Israel). From the late 1980s and early 1990s and on, countries like North Korea, Pakistan, Iran, and Iraq joined the group of publicly missile testing countries (Missile Threat – North Korea, Pakistan, Iran, Encyclopedia Astraunomica – Iraq). By the end of the Cold War, there were about a dozen
countries which frequently test-fired their own missiles, and they were most certainly not the most peaceful countries in the globe.⁵

In the aftermath of the Cold War, while the American and Russian missile tests were less and less in the news, several countries started to attract global attention and concern. During and shortly after the Taiwan Strait Crisis, China test-fired many missiles that were opposed by Taiwan as well as the US. Shortly after the 1999 Kargil War, India and Pakistan test-launched many missiles at a tit-for-tat pace. Beginning in 2016, North Korea test-launch some its first ICBMs that could reach the western shores of Alaska and Canada (Missile Defense Project – North Korea, Missile Defense Project – Hwasong 15).

Each time these countries test-fired a missile, one question was often asked, yet almost never clearly answered: was global security at risk with these publicized missile launches? Why would countries take such a risk? Surprisingly, the field of international relations has overlooked the reasons behind missile tests and the effects of these tests on the risk of militarized conflict. More specifically, no large-scale empirical analysis has been conducted to understand why countries test-fire missiles publicly and whether these tests actually threaten global security by engendering militarized conflicts and/or wars.

In this paper, I present a dataset on public missile tests conducted by 11 countries between 1949 and 2015. These countries are the United States, USSR/Russia, United Kingdom, France, China, India, Pakistan, North Korea, Iran, Iraq and Israel. More specifically, first, I detail what type of missile tests I focus on. Second, I discuss what relevant data is available in the field. Third, I describe how I used the existing data to

---
⁵ The US, Russia, England, France, Israel, Iran, Iraq, India, Pakistan, North Korea, China are all countries that fought several wars and frequently involved in militarized disputes (see Brecher et al 2017, and Palmer et al 2015 for an empirical record)
compile my missile tests datasets. Fourth, I lay out the coding procedure for each variable I created to describe these countries’ missile tests. Fifth, I present each country’s missile testing history focusing on what sort of missiles they developed and tested within the years of interest. Lastly, I discuss what sorts of potentially-generalizable patterns we observe about these 11 country’s missile tests. This discussion also sheds light on what research questions the dataset can help address.

Overall, it appears that there is little to no regularity in the missile testing trends of the selected countries. Each country had episodes during which they ended up publicly testing particularly more missiles than other times. A deeper examination reveals that most of the peak missile testing periods coincide with significant international crises that countries endure. Thus, missile tests may not always be driven exclusively by the need to complete a missile-program. What remains to be answered is whether international crises trigger missile tests or whether missile tests cause crises which risk escalating to wars. Luckily, the dataset that I present in this paper can help answer these crucial questions of international security.

What makes missile tests an important topic in international relations? First, public missile tests are frequent incidents that rarely go unnoticed by countries. Historically, countries not only sought to observe but also to limit each other’s missile tests. For example, during 1970s, the US and the USSR included clauses in the SALT I and SALT II talks to curb each other’s missile tests (State.Gov). In the aftermath of the Cold War, the US experienced similar issues with Iran and, more recently, with North Korea. From 2006 and on, North Korea received a total of 4 major sanctions some of which sought to curb its
missile tests. Hence, in the world of politics, countries are at least unhappy when their adversaries test-fire missiles. The extent of this unhappiness and whether it leads to militarized conflicts seem to be overlooked.

Second, from a policy perspective, it seems that missile tests are far from losing salience in global politics. Nowadays, countries like Russia and China are test-firing more and more sophisticated missiles. American military officials and executives recently stressed the need to develop and test-fire new weapons just like their adversaries like Russia, China and even North Korea. Thus, it seems like a great more deal of missile tests are going to be conducted by some of the militarily most powerful countries in the world. However, how these tests will affect global security and why they are being conducted publicly in the first place is still not well-understood.

Defining Missile Tests

In this paper, I focus on a particular case of missile tests: countries’ publicly-conducted test-launches of missiles and missile-related systems. Missiles and missile-related systems include short, intermediate, medium and intercontinental range ballistic or cruise missiles, land, as well as air or ship-based missile interceptors. Tests of these missiles have three aspects. First, the test should be publicly conducted by a country.

---


7 The two countries developed and now are testing supersonic gliders that are capable of launching nuclear missiles from space. China tested its glider Wu-14 the first time in 2015 (ValueWalk 2015). Russia’s first glider was first tested in 2016 (Daily Star 2016)

8 Recently, American President Donald Trump stressed the need to develop new weapons as well as Vice President Mike Pence in 2019 (The Oakland Press 2019, Whitehouse 2019).
Second, a missile and/or a missile interceptor should be actually launched during the test. Finally, the launch should have no deliberate goal of causing physical harm against an adversary. In what follows, I briefly discuss these points.

The first aspect of the missile tests I investigate is their publicness. Countries sometimes display their missile tests to the rest of the world, and sometimes they do not. One reason why a test-launch remains secret could be because there is an ongoing test ban that the missile-testing country is under. Another reason for secret tests could be that a country seeks to ensure that the first publicized test of this weapon goes without any accidents in order to avoid embarrassment before the rest of the world. On the other hand, if a missile test is publicized (or is not hidden), we have an opportunity to discuss its impacts on global politics. If North Korea test-fires 5 ballistic missiles and some of these missiles splash into sea 200 miles off the Japanese coast, this focuses our attention on how Japan reacts. We also wonder how South Korea, the US, Russia and even China will react. More importantly, another question is why North Korea launched 5 missiles in such a dangerous fashion. Naturally, secret missile tests will be of little help in answering these questions as they do not raise the same level of attention.

The second aspect of the tests is that they pertain to actual test-fires. This means that the tests I record are not feasibility tests conducted in factories or labs. Such feasibility tests involve the assessment of the navigational systems, engine stability and tests of specific components of a missile that do not involve an actual whole live test-launch. Each weapon typically undergoes such tests before an actual test-fire. Therefore, if, for example,

---

9 Soviets are suspected to have conducted some ICBM tests in secrecy when they signed SALT (Strategic Arms Limitations Talks) and then SALT-II with the US (State.Gov).
the Boeing Company runs operational tests on a Minuteman missile in its factories, or Lockheed Martin runs similar tests on a Trident missile, neither is considered as an actual test launch. Similarly, computer simulated test-fires are not included in the data collection project either, since I only focus on actual test-fires.¹⁰

Third, the missile tests are not conducted to cause harm against an adversary. Making this distinction may seem straightforward, but it is not. A test going wrong may result in harm to an adversary or a friendly country. Conversely, an actual missile strike that misses its target by a large extent can seem like a test. Examples fitting these categories are numerous. For example, during a NATO naval exercise in the Aegean Sea, a US missile hit a friendly Turkish boat (NY Times 1992). In 1986, a malfunctioning Soviet weapon hit a Chinese village (NY Times 1986). In both cases, the missile testing countries issued an apology, and made it clear that their intention was not hostile. However, consider the following example: During the Cold War in 1983, the USSR once hit a South Korean civilian plane, during what it claimed was a missile test. Hundreds of lives were lost. Soviets claimed that they were testing an Su-15 interceptor missile. However, they failed to convince the rest of the world (sfchronicle.com 2019). Thus, intentions are rarely perfectly clear when countries publicly fire missiles. For the present study, I investigate cases in which the launching country specifically claims that the launch was a test and this was confirmed by other countries. In that regard, the Soviet missile against the American jet is not considered as a test – not because it clearly was not, but because it was not historically confirmed by other nations as a test.

¹⁰ Electronic simulated test-fires are not live test-launches of a weapon. It is a computerized simulation of a launch. The US has been conducting such simulations for decades, independent of actual test launches.
In what follows, I discuss what available data we have to build a sample of missile tests which fit the three criteria established in this section.

**Available Data on Missile Tests**

To this day, no large-scale data has been collected on public missile tests to investigate why these tests occur and their effects on international security. This shortage, however, does not necessarily mean that there are no data available to be collected. For example, it is rather easy to come across non-academic newspaper articles that present a chronology of North Korea’s missile tests over a certain period of time. Similarly, America’s MDA (Missile Defense Agency 2018) regularly presents reports of its missile interceptor tests. There are also other country-based reports. However, these individual reports have not yet been combined in a single comprehensive dataset. Such combination is required to explore the reasons behind countries’ public missile tests and their potential association with militarized conflicts. In what follows, I briefly highlight several organizations/media sources which can be particularly helpful in generating the necessary data set.

The first source that can help with data collection is from a Washington-based think tank, CSIS (Center for Strategic and International Studies). One of the CSIS’ projects is called *Missile Threat Initiative* (MT) which presents detailed analyses on various countries’ missile arsenals (missilethreat.org). The MT also includes chronologies of Iran and North Korea’s missile tests from early 1980s until 2017. ¹¹ Most of the information for these countries are obtained from Jane’s Military Balance reports and various media outlets. No

---

¹¹ For North Korea, see Missile Defense Project–North Korea. For Iran, see Missile Defense Project – Iran.
A chronological history of missile tests is offered for other countries. On the other hand, it is possible to directly find detailed information from MT on the specifics of almost every known and currently commissioned missile from France, England, US, Russia, North Korea, Iran, Israel, China, India, and Pakistan. The main issue with CSIS’ MT Initiative is that it does not have enough information to create a missile test event dataset large enough to conduct robust statistical analyses. In addition, most of the Cold War tests of countries are not recorded in this project. Hence, we only get a sense of the current missiles and tests.

Another data source is the Encyclopedia Astronomica (EA) which offers a far-wider case coverage than CSIS’ Missile Threat Project. Designed by Mark Wade, this project is presented in a web site. The goal of the Encyclopedia Astronomica website is to present encyclopedic information on countries’ rocket-launching and space exploration histories. In that regard, one can easily find chronologies of known missile tests and space rocket/shuttle launches of almost all countries in the world. It is even possible to create a missile tests dataset by simply relying on EA. Yet, due to two major issues, exclusive reliance on EA cannot completely help with the current data collection project. First, it is difficult to distinguish between public and secret tests in EA. The website only occasionally provides information regarding whether the test of interest was conducted in secrecy. The second problem is that EA does a limited job in covering missile test launches from cruising ships or unidentified test locations. For example, many North Korean, Pakistani, and

---

12 My data collection borrows many observations from this website. In the next section I detail the features of the site.
13 Encyclopedia Astronomica relies on government archives, academic and military reports some of which potentially might have been classified at the time when the test of interest was conducted.
Iranian missiles launched from naval vessels, aircrafts, or from undisclosed locations are not recorded in EA.\textsuperscript{14}

There are also several other websites that offer information about countries’ missile programs and some of the key tests of these missiles.\textsuperscript{15} Yet, EA and MT appear to have the largest coverage. The issue is that these websites by themselves seem to be inadequate for investigating questions about the causes and consequences of public missile tests. Therefore, I decided to collect data on missile tests using these two websites in conjunction with other sources. In what follows, I detail the data collection procedure.

**Dataset**

The dataset I created involves countries’\textsuperscript{16} missile tests that are (i) public, (ii) actual test-launches, and (iii) that seemingly bear no intention of causing direct physical harm on any adversary. As countries of interest, I included North Korea, Iran, India, Pakistan, China, Soviet Union, Russia, the US, the UK, France, Israel and Iraq between 1949 and 2015.\textsuperscript{17} There are three reasons why I focus on these countries. First, these countries had

\begin{itemize}
\item \textsuperscript{14} A missile can be test-fired from an undisclosed location, however the rest of the test process including the landing / splash area can be reported.
\item \textsuperscript{15} See Global Security (https://www.globalsecurity.org/org/index.html)
\item \textsuperscript{16} It seems that non-state actors can potentially test fire missiles just like a state, yet this rarely happens. For example, Hamas is one such non-state actor which occasionally test-fires short range missiles (Missile Threat - Hamas). For the present study, I only focus on sovereign countries.
\item \textsuperscript{17} 1949 is the year when the Soviets conducted their first nuclear test. From this date and on, the US as well as its allies started to conduct nuclear as well as ballistic weapon tests more frequently than before. For example, the US had already conducted tests of weapons such as V-1 rockets (or its modified versions) before 1949. However, after the Soviets achieved nuclear capabilities, countries started to conduct missile tests publicly. Therefore, the year 1949 is a reasonable start date to collect data on public tests. The end date 2015 serves to capture increases Russian, Chinese, Indian, Pakistani and North Korean missile tests from 2010 and on.
\end{itemize}
their own indigenous missile development programs and did not exclusively rely on imported missiles (or on missiles that were placed within their borders but which they couldn’t use without consulting allies). Second (and probably related to the first), these countries are the most active missile-testing countries in the world between the years of interest. Although there are many other countries that frequently launched rockets to space, these were often sounding rockets exclusively used for scientific purposes. Third, not only did these countries test-fire their own missiles frequently but these tests often attracted global attention and criticism. Thus, their missile tests rarely went unnoticed.

---

18 In fact, none of the countries in the sample have purely ingenious missile technologies of their own. In the late 1940s, American, Soviet and French ballistic missile technologies were directly based on German V-2s developed by Von Braun. For example, one of the earliest French civilian rockets called Veronique was developed by German scientists who escaped Nazi Germany (Encyclopedia Astronautica – Veronique). One of the earliest Soviet missiles called the R-1 was directly based on V-2 schematics (Encyclopedia Astronautica – R1). Chinese presumably got assistance from Soviet scientists to develop their first Dong Feng missiles (Encyclopedia Astronautica – Dong Feng 1). North Koreans reverse engineered Soviet Scud missiles to build their Hwasong missiles (Encyclopedia Astronautica – Hwasong 5). The North Korean Nodong-1 missile schematics were used by Iranians and to a lesser extend Pakistanis. The Iranian Shahab missiles schematics are based on North Korean Nodong-1 missiles (Encyclopedia Astronautica – Iran). Pakistani Hafif V Ghauri missiles are also based on Nodong-1 schematics (Encyclopedia Astronautica – Pakistan). All of the missiles that Iraq test-fired in the sample (with the exception of Babylon Gun) are based on Russian R-17/Scud schematics (Encyclopedia Astronautica – Iraq).

19 There were candidate countries which did not make it to the present sample: Germany was also test-firing a very large quantity of missiles. Yet, these tests took place during the Nazi era, and most specifically during World War II. In addition, the publicness of these tests was hard to confirm. In the Cold War, and Post-Cold War era, Germans predominantly launched rockets for purely scientific purposes, and the launch vehicles were not military-based missiles (Encyclopedia Astronautica – Germany). Australia also had many rocket launches. Some of these were military-based but conducted by England in the Woomera launch site. The remaining launches according to EA were for scientific purposes (Encyclopedia Astronautica – Australia). South Africa was another candidate given its involvement in many conflicts and the global suspicion that it was in possession of nuclear arms. Yet, according to EA, South Africa test-fired military missiles very rarely (Encyclopedia Astronautica – South Africa). Syria was one of the countries that came close to being added to the sample. Yet, this country had only a handful of missile tests in the pre-2010 years (Encyclopedia Astronautica – Syria). This constitutes a problem since some of the widely-used militarized interstate conflict datasets do not cover any event after 2010 (see Maoz et al 2019, Palmer et al 2015). Lastly, Egypt and Libya were candidate countries for the sample. Egypt test-fired a few Soviet R-17s in the 1970s. Yet, from 1973 and Egypt only had civilian space rocket launches conducted elsewhere (Encyclopedia Astronautica – Egypt). Libya only had a handful of missile launches in the 1970s (Encyclopedia Astronautica – Libya). A possible reason for the surprisingly low missile test records for Libya and Egypt could be that both managed to conduct many secret tests. Alternatively, it could be the case that these countries could not afford to test-fire as many missiles as they wanted. More importantly, it could be that neither developed missile technologies of their own.
Overall, one can label the 11 countries of interest as the notorious missile-testers of international politics. This presents the opportunity to assess the reasons behind these tests as well as their consequences in international politics with adequate number of observations to run statistical tests. In what follows, I detail the coding procedure.

(i) Sources:

To extract information on countries’ publicly conducted missile tests, I relied on four main sources of information. The first is CSIS’ Missile Threat (MT) Initiative Project ([https://missilethreat.csis.org/](https://missilethreat.csis.org/)).\(^{20}\) As highlighted in the earlier sections, MT presents chronological orders for Iran and North Korea’s missile tests. In addition, MT presents descriptions of the missiles that each country in the sample except Iraq currently use.\(^ {21}\) Accordingly, as the first step, I used this source to record which missiles are possessed by the countries of interest, and what the key testing dates are for these missiles. However, MT does not record all the public missile tests of the 11 countries of interest. Hence, I turned to 3 additional sources.

The second source is Lexis-Nexis\(^ {22}\) which is an internet database for worldwide news widely used in the field of international relations.\(^ {23}\) For each country, I conducted a keyword-based search to gather news sources using: “country name” + “missile” + “test”

---

\(^ {20}\) Center for Strategic and International Studies is an American Think Tank founded in 1962, and initially affiliated with Georgetown University.

\(^ {21}\) An alternative resource is Jane’s Military Balance which is a well-known project to collect data on countries’ military possessions. The Missile Threat Project of CSIS relies not only on Jane’s’, but also additional resources to present its data. Thus, at least for the countries that I focus on this current project, CSIS presents more detailed and useful information than Jane’s.

\(^ {22}\) During the data collection phase, Lexis-Nexis changed its worldwide news database name to Nexis Uni.

\(^ {23}\) Examples for these data collection projects that use Lexis-Nexis are the Militarized Interstate Dispute (MID) Dataset (Palmer et al 2015), Alliance Treaty Obligations Project- ATOP (Leeds 2018), or Threat and Imposition of Sanctions – TIES dataset (Morgan et al 2013)
 specification. Alternative wording specifications to designate weapon tests revealed fewer news sources. I used a slightly different keyword specification for the United States (“missile name” + “missile” + “test”). The names of the missiles were determined based on MT’s list of missiles for the United States. Using these keywords, I extracted news sources to code the variables of interest for the dataset.

The third source I relied on is Encyclopedia Astraunomica, which is a website developed by Mark Wade. Encyclopedia Astraunomica (EA) records the test-launches of missiles, satellites, and orbital launch vehicles of almost all the countries in the world. This website uses academic publications, government archives, news reports, and company reports to record detailed information on the schematics as well as service histories of missiles, civilian rockets, and missile interceptors. EA presents one of the most detailed

24 Thus, for North Korea, the search terms were “North Korea” and “missile” and “test”. For several countries, the subject (i.e. country name) search term had two versions. For example, for France, not only did I use “France”, “missile” and “test”, but also were “French” and “missile” and “test” combinations. A similar procedure was applied for China (repeating the same search with “Chinese” instead of “China”). In the case of Britain, I used the following alternative words next to “missile” and “test” to look up for the information I needed on weapon tests: “British”, “RAF”, “Royal Navy”, and the “UK”. News sources use one of the five different words to designate a test conducted by the Great Britain. Moreover, the relevant sources can repeat the news of tests such as “Royal Navy tested a cruise missile”, “Britain test-fires a ballistic missile”, “RAF tested an air-to-air missile”. In the last case, marking RAF or Royal Air Force revealed the exact same news sources. Note that if the royal navy in question was the Dutch one, the news source would specify Dutch Royal Navy. If not, then it is British Royal Navy.

25 Alternative wording specifications I tried were: “tested”, “launched”, “test-launched”, “fired”, and “test-fired”. These specifications revealed fewer news resources that “missile” and “test” specifications, and even did not reveal some of the news of weapon tests that could potentially be coded as observations for the dataset.

26 The reason is that most resources in Lexis-Nexis are American-based news outlets. Thus, they rarely use “United States” or “America” as the subject. To get a representative sample of American weapon tests, I had to rely on alternative key word specifications and additional sources. In terms of alternative word specifications, I realized that the news pieces in Lexis-Nexis often used the name of the missile in question. For example, news of a test is often worded as “Minuteman missile tested”, “Navy tests Trident missile”, or “Jupiter missile tested”.

27 See https://missilethreat.csis.org/country/united-states/ for this list, and the weapon details. Specifically, the weapons for which I conducted search were: Harpoon, Snark, Titan (I and II), Minuteman (I, II, and III), Lance, Pershing (1, 2), Atlas, MX (Peacekeeper), Jupiter, ATACMS, ALCM, JASSM, Tomahawk and Trident.

28 http://www.astronautix.com/

29 Encyclopedia Astraunomica is endorsed by partnered with more than 10 different websites including American Astronautical Society. It is also considered as a reliable source by NASA. For more details, see http://www.astronautix.com/
record of missile and rocket launches before 1980s. Despite this advantage, EA has poorer coverage of missile tests conducted at sea compared to Lexis Nexis-extracted data.\textsuperscript{30} For the post-1980 tests, the reports from Lexis-Nexis and Encyclopedia Astraunomica do not contradict each other.\textsuperscript{31}

The final source I use is the World Nuclear Explosions (WNE) Dataset by Yang et al. This source presents US, USSR, England, France, China, India, and Pakistan’s nuclear tests between 1945 and 2000.\textsuperscript{32} For each nuclear test, WNE offers information on test date (month, day, year, time), geographic location of the test (decimal), the yield, and the name of the nuclear device used. A major challenge is that WNE only covers successful detonations; failed attempts are not reported. Unfortunately, countries seldom reveal failed nuclear test attempts. Thus, I do not collect information on failed attempts elsewhere.\textsuperscript{33}

(ii) Coding:

In line with the types of missile tests I focus on in this paper, I ensured that the information source used to code a missile test described a public incident.\textsuperscript{34} Also, no factory tests, 

\textsuperscript{30} Thus, many naval tests conducted by Russia, France, India, Iran and England are not in Encyclopedia Astraunomica.
\textsuperscript{31} Despite its wide coverage of missile and rocket launches, caution is needed when using Encyclopedia Astraunomica. The website presents archival data. Hence, there is always the possibility that some of the tests conducted by one country were not noticed by others at the exact time of the test. For example, if Soviets test-fired R-7 missiles from Kapustin Yar Test range in late 1960s, and did not release this information till 1980s, these tests may not be fully suitable with the dataset. On the other hand, we know that the US was tracking telemetry data of the Soviet missile tests. American experts quite often detected Soviet tests thanks to listening posts in Iran, Turkey as well as in Alaska. Hence, it could be that Americans were able to trace most of the Soviet weapon tests. However, if I did not use EA, and solely relied on Missile Threat or Lexis-Nexis, countries like France, England, and the Soviet Union would appear to have almost no missiles test-fired before 1980 – which is certainly not the case.
\textsuperscript{32} In the case of North Korea, I relied on Lexis-Nexis since the earliest nuclear test for this country took place in 2006.
\textsuperscript{33} Data on attempted nuclear explosions using a missile can actually be found from EA. However, I did not come across any EA observation that was not already recorded by Yang et al.
\textsuperscript{34} To ensure that the recorded tests were public, I checked the source of the news related to the incident of interest – using Lexis Nexis. If the incident of test was only reported by a media source of the missile testing country and not confirmed by any other country or a media source, then I did not record the incident. For the exclusively EA-based observations, it was much harder to confirm the publicness of a missile test. For these
engine simulations, or purely civilian sounding rocket launches were recorded.\textsuperscript{35} Overall, I took into account the following missile types and missile-based weapons that were tested publicly: short range ballistic missiles (SRBM), intermediate range ballistic missiles (IRBM), medium range ballistic missiles (MRBM), intercontinental range ballistic missiles (ICBM), short range cruise missiles (SRCM), intermediate range cruise missiles (IRCM), medium range ballistic missiles (MRCM), intermediate range cruise missiles (ICCM), missile interceptors (MI), submarine- or ship-launched ballistic/cruise missiles, nuclear warheads, and space rockets/vehicles launched with ballistic missiles (which typically have intermediate or greater ranges).

The data matrix I created is such that each row/observation represented a day when a specific missile of interest was test-launched by one of the 11 countries.\textsuperscript{36} Accordingly, if two different missiles were tested the same day, this resulted in two observations.\textsuperscript{37} However, if the same missile was tested multiple times in a day, this resulted in one observation. Lastly, I came across several cases when a country was witnessed to have test-fired multiple missiles in a day, but the names of these missiles were unidentifiable. In such observations, I only recorded the ones for which the day, month, and year as well as the location of the test are fully known.

\textsuperscript{35} A purely civilian rocket launched for aeronomy mission such as Jaguar II of England may not be relevant to the present inquiry. However, a scientific mission conducted by a military missile such as American Atlas D, or North Korean Taepedong III may be particularly relevant. In fact, almost all of North Korea’s, Iran’s and India’s space missions were met with strong criticism by their adversaries – since these missions are often conducted by military launch vehicles (EA – Iran, EA – North Korea, EA – India).

\textsuperscript{36} Recording the missile tests on a daily basis allows for the counting the total number of missile tests conducted in a month, quarter year, year or many other time intervals. That said, I found that for several observations, the test days are not clearly identified. Therefore, the most reliable smallest time interval to aggregate missile test numbers of a country is a year-month.

\textsuperscript{37} For example, if the Soviets tested a nuclear warhead and a RS-18 missile the same day, the two are coded as separate tests. The exception is when the RS-18, one of the most destructive missiles of the Soviet era, was launched with a live nuclear warhead. However, missile tests with a live nuclear warhead are rare throughout the history. Countries have often tested nuclear warheads separately from nuclear capable missiles.
situations, the tests fit with the “publicness” criteria, since they were witnessed by others. However, I created only one observation to capture these particular missile test incidents.\(^{38}\)

(iii) Variables:

**Missile name:** The first variable I created is the name of the missile according to the resource. The name of a missile is different that the family and the aspect of the missile. For example, consider the notorious French M-51 missile. The name of the missile is M-51, the aspect is an ICBM, and the family is MSBS (EA – MSBS M51, Missile Defense Project – M-51). Similarly, in the case of Israel’s Jericho 3, the name is Jericho 3, the family is Jericho and the aspect is IRBM (EA – Jericho III). The name of the missile is necessary to separate one observation from the other. So, if missiles with two different names were tested the same day, this led to two different observations even if these missiles belonged to the same family and/or were of the same type.\(^{39}\)

**Missile aspect:** Next, I created a variable to designate the aspect of the missile. The aspect of a missile includes information about its range, whether it was a cruise or a ballistic missile, whether it was used in a space mission, and whether it was a missile interceptor. Intuitively, the aspect of a missile should matter because tests of missiles with more destructive aspects may also be more threatening to observers. For example, North Korea’s Hwasong-15 (aspect: ICBM) tests may be more concerning to South Korea or to the US

---

38 There were fewer cases of Iranian naval exercises where the source claimed that more than one types of missiles are test-fired. For these cases, if the source also clarified how many different types were test-fired, I created that many observations. If not, I created two observations to designate that more than one type of missile was test-fired in the given day.

39 I did not differentiate between missiles of the same name but that used different engines – unless the country of interest changed the name of the missile due to the new engine type. Naturally, two missiles with the same name, but also with different engines are technically different. However, information about what engine type was used was rarely provided in the sources I used. Thus, I chose not to identify the engine type of the missile of interest.
than its Nodong-I (aspect: IRBM) tests. Similarly, a Pakistani Shaheen-I (SRBM) test may be less threatening to India than a Hatf-VII Babur (SRCM) test. Thus, if we desire to understand to what degree missile tests provoke adversaries, we need information on the aspects of the tested missiles.

In the case of some observations, the range of the tested missile was unidentified. I followed two separate strategies to address this issue. First, I completely left the missile aspect variable as “unidentified”. Second, I used the news source(s) of the relevant incident to extract limited information about the missile aspect. It was often the case that news sources depicted a test incident such as “country X tested a ballistic missile”, “country Y tested a cruise missile”, or “country Q tested a short-range missile”. In these cases, I sought to make use of the available information, and coded the aspect variable as “unidentified ballistic missile”, “unidentified cruise missile” or “unidentified short-range missile”. This way, it is at least possible to calculate how many cruise, ballistic or short range missiles were test-fired by the country of interest in a given period of time.

Missile Test Count: The third variable records how many times the missile of interest was test-fired in the day of observation. There are two reasons why I created this variable. First,

---

40 Cruise missiles are basically drone-type missiles that can self-navigate to reach a target. The notoriety of the cruise missiles is their capability in dodging radars thanks to the flexibility in their navigation (Doughterty 2015). Tomahawks are historically one of the most notable cruise missiles that helped the US achieve a decisive victor over Iraq in the First Gulf War (Missile Defense Project - Tomahawk). Indian-Russian co-manufactured BrahMos is another example of cruise missile (Missile Defense Project - BrahMos). A discerning aspect of cruise missiles is their speed, and BrahMos is reported to travel faster than the American Tomahawk (Thai News Service 2008). The testing of BrahMos by India was heavily opposed by Pakistan. Yet, Pakistan managed to obtain a used Tomahawk during the US occupation of Afghanistan in 2001. From this missile, Pakistan was able to reverse-engineer and build its own cruise missile (UPI.com 2011). On the other hand, ballistic missiles have a pre-determined trajectory. They are launched vertically (sometimes with a certain angle). They reach an apogee, and then start their descent. Ballistic missiles are more like modern day catapults or mortar cannons.

41 Some of the observations for countries like Iran, Pakistan, North Korea and Soviets have several coded as the type of missile completely unidentified. The only reason these tests are included are because they are confirmed by countries other than Iran, Pakistan, North Korea or Soviet Republic respectively.
if we investigate what explains the variance in the number of missiles that countries test-launch, this is basically the most reasonable outcome variable. Moreover, what makes countries test-fire 10 missiles in a day may be different than what makes them test-fire only one missile in a day.

Second, the number of missile tests should be important in assessing the impact of missile tests on conflict escalation among countries. A country that tests a certain missile 5 times in a row during the same day may be more intimidating to observers than had the country test-fired this missile only once. This deterrent effect might even be stronger if the country manages to test-launch multiple ICBM or ICCMs in a day.\textsuperscript{42}

Coding the number of times a missile was test-fired in a day presents several challenges. For MT and EA-extracted cases, recording how many times a missile was test-fired was straightforward. Both websites typically present this information with much desired clarity. However, in the case of exclusively Lexis-Nexis-coded incidents, there was the risk of the relevant news source not specifying the number of times a missile was test-fired. For these, I coded the number of times as 1 if the news piece simply wrote “Country X tested its missile”. If there was evidence that the missile was test-fired more than once, I coded the variable as 2.\textsuperscript{43}

\textbf{Test location:} This variable marks the location from which the missile of interest was test-launched. A possibility is that the closer a missile’s test location is to an adversary, the

\textsuperscript{42} In my sample, only the USSR and the US were able to test-launch multiple ICBMs in the same day
\textsuperscript{43} It is possible that the missile in question was test-fired more than two times. However, we cannot arbitrarily figure out a number. In addition, the most frequent number besides 1 for the missile test numbers is 2. Thus, by coding 2, I was able to not contaminate the distribution of the variable of interest. Additionally, I come across this uncertainty in the number of missile-launches very rarely. Moreover, any information that one of the sources did not specify would be found in another source that I used.
more provoked or intimidated this adversary may become. For example, a missile that India
test-launches from the Pokhran Desert Test Range may be more threatening to Pakistan
than a test from Chandipur-On-Sea which is in East India. Similarly, the test-launch of a
Soviet missile from Barents Sea may be more concerning to Scandinavian countries than
the test-launch of the same missile from Pacific Ocean. However, the latter test may
actually concern the US more than the test in Barents Sea. In brief, determining the test
launch sites may help us get a better understanding of which missile tests are more likely
to trigger opponents to take hostile actions.

To designate the location of the test, I created four different variables. The first two
variables mark the location of the test in string format, as extracted from the original source.
Of these two variables, the first one shows the broader region where the test is conducted.
This could be a city, a region, or part of a sea/ocean. The second variable records the name
of the actual test facility where the missile was tested. The last two variables code the
longitude and the latitude of the location where the weapon is tested. This information is
present directly in the WNE database for nuclear explosion tests (Yang et al). For the
remaining weapon tests, I extracted the coordinates of the relevant test facility using
longlat.net or Encyclopedia Astronomica.

**Missile Commissioning Date:** The next variable I created indicates the month, day, and
year when the missile in question was commissioned for service. The reason why I

---

44 A test from Chandipur-on-Sea may be relatively more concerning to China than Pakistan-all else being constant
45 Lexis-Nexis extracted cases on Pakistani, Iranian and North Korean tests are rarely specific about the exact
   location of the test. Hence, there are missing “test location” values for these three countries.
46 The exact address is https://www.latlong.net/convert-address-to-lat-long.html. This website is accurate to
   extract information on American locations. I relied on alternative resources for non-American cases.
47 See http://www.astronautix.com/.
recorded this variable is because a country may have different motivations behind testing officially combat-ready missiles compared to those that are not. For example, several incidents highlight the possibility that countries may test-fire commissioned missiles to later sell those missiles to potential clients (see BBC 1998). A similar incentive may not be applicable to test-firing non-commissioned missiles. Alternatively, the public test of a combat-ready missile may intimidate adversaries differently than a missile that is not commissioned and that requires more work before being accepted into service.

In order to record the commissioning date of the tested missile in question, I relied on MT and EA – both of which detail the service history of the missile of interest. For the majority of the observations, the exact day of the commissioning date is missing. Also, for many missiles, the commissioning month of the missile is missing. In that regard, the most reliable way to use this date variable is to compare the year of the commissioning and the year when the missile was test-fired.48

**First Test:** This variable is a binary one that is coded as 1 if the missile in question was test-fired for the first time at the observed date, and 0 otherwise. A possibility is that when a missile is test-fired for the first time, given the availability of funds, more tests of that missile may transpire soon. Therefore, an important factor that may indicate why a certain missile is frequently test-fired at a certain time interval may be that it was recently test-fired for the first time. Thus, accounting for the “first test” binary variable is crucial if the goal is to understand the reasons behind missile tests.

---

48 Another problem that I came across was that for some tests in EA, the operational capability date instead of the commissioning date were presented. If I failed to find any other source to present information on the commissioning date, I ended up using the operational capability date of the missile. Naturally this date is often slightly earlier than commissioning date. However, an operationally capable missile can still be deployed and used -if needed- in combat.
Identifying the first test of the variable was challenging and required me to follow two different strategies. First, I extracted information from MT and/or EA to record the first time a missile was tested. Hence, an observation is coded as 1 if it pertains to the first known public test of the missile, and 0 otherwise. Second, if this information was missing, and if Lexis Nexis was not helpful in revealing the needed information either, I recorded the first observation that a weapon of interest appeared in the dataset as its first test.49

**Nuclear Capability:** The last variable I created indicates whether or not the missile in question is nuclear capable. For almost every observation in the dataset, at least one of the three main resources (MT, EA and Lexis Nexis) had information on the nuclear capabilities of the missile in question. Hence, missing values for this variable are no more than a handful.50 In general, a nuclear-capable missile is designed such that it can carry nuclear as well as non-nuclear warheads. Several incidents indicate that a country test-firing a nuclear-capable missile with a mock warhead is particularly concerning. For example, neither Iran, nor Israel are witnessed to have conducted any nuclear detonation tests. However, both countries have missiles that can theoretically carry nuclear warheads. For example, Iran’s various tests of nuclear-capable missiles met great criticism from both Israel and the United States (BBC 2000). Hence, if we are to determine if missile tests provoke adversaries, accounting for whether or not the missile in question is nuclear-capable is crucial.

---

49 The exception for this rule is that if the first observation is extracted from a Lexis-Nexis based news source which claims that the weapon was publicly tested before (but not specified when and where).
50 It is sometimes the case that a weapon at its early development stages may not be nuclear capable. However, the weapon can later on be upgraded – under the same missile name to carry a nuclear warhead. These cases are very rare in the dataset. The issue is that I was not able to find the exact time a missile switched from non-nuclear capable to nuclear capable. For these rare cases, I coded the missile as simply nuclear capable.
One potential variable that I did not record would indicate whether the test was a success or not. The reason I forgo coding this variable is that report of success or failure can be, and often are, biased. This is particularly an issue for North Korea, the USSR, Pakistan and Iran which often report their weapon tests as successful. However, the success of these tests may be discredited by other nations. In fact, a report by Japanese Business Wire (2000) indicates that even the US Air Force may have been over-reporting the success of its missile interceptor tests. In general, the success or failure of many missile tests have been subject to much debate. To avoid loss of objectivity, I do not code a success indicator. Nevertheless, the success or a failure of a missile test admittedly can affect to what degree observers are provoked or intimidated. Hence, such a variable would be useful in assessing the impact of missile tests on global security if countries did not misrepresent the success of their tests.

Overall, I recorded a total of 11,817 observations (i.e. weapon test days). 5,603 are American observations. For the Soviet Union (1949-1991), I recorded 3,650 observations. Beginning in 1992, Russia has 435 observations. In the case of China, I recorded 200 observations. In the case of India, I recorded a total of 316 observations. Pakistan has 104 observations. For North Korea, I recorded 79 observations. Iran has 171 observations. For France, I recorded 306 observations. The United Kingdom has 157 observations. Iraq has 98 observations. Lastly, Israel has 68 observations which is the least in the dataset. Note that the number of observations is not equivalent to the total number of missiles or missile-related systems a country test-fired. For this information, I plotted a map in Figure 1. Countries in Figure 1 are shaded based on their total number of missiles test-fired. In
addition, the black dots represent the identified test ranges where these launches were conducted.\textsuperscript{51}

\textsuperscript{51} For more detailed information missile test statistics by country, the reader can refer to Table A1 in the Appendix.
Figure 1: Public Missile Tests by Country

Weapon tests by Country
Dots indicate test locations

Missile Tests
- 7064
- 4482
- 342
- 309
- 245
- 223
- 208
- 152
- 120
- 116
- 70
Country Profiles

(i) United States:

At the end of World War II, the Third Reich rocket scientist Werner Von Braun who engineered the world’s first liquid propellant missile\(^{52}\) was hired by the US.\(^{53}\) He and his team started developing missiles in New Mexico’s White Sands Range. Some of the earliest American short-range ballistic missiles such as Hermes A-2, Corporal-E and MGR-1 were test-fired in the late 1940s at this range. Most of these missiles were manufactured based on Von Braun’s seminal V-2 designs (EA-V2, EA-USA, EA-White Sands). Around the same years, many cruise missiles such as Matador, Snark, and MX-775 were test-fired from the Cape Canaveral Range in Florida (EA–Cape Canaveral, Werrell 1985).\(^{54}\) Among these cruise missiles, MX-775 was the first intercontinental missile test-fired by the US, which took place in 1952 (ibid).

During the 1950s, Cape Canaveral Test Range was the most frequently used range to test-fire missiles. Some of the earliest American ICBMs such as Atlas A, B, C, D, and Titan I were test-launched in this base (EA–Cape Canaveral).\(^{55}\) Towards the end of the 1950s, an increasing number of ICBMs were also test fired from the Vandenberg Air Force Base in California. In the coming decades, Vandenberg became one of the busiest – perhaps the busiest – missile test ranges in the US.

---

52 Von Braun designed V-2 as the first liquid propellant missile. However, the theory of a liquid propellant missile was developed in the early 1900s by the American scientist Robert H. Goddard.
53 For those who are familiar with the relevant historical period, the verb “hiring” may come off as too soft in the case of how German scientists started to work for the US.
54 Interestingly, a near-majority of the American tests in the late 1940s and early 1950s are cruise missiles.
55 According to the data I collected however, America’s first ICBM test was conducted at Edwards Airforce Base in Florida (EA–Edwards)
In 1959, the United States test-fired its first Minuteman ICBM from Edwards Air Force Base, Florida (EA – Minuteman 1A, EA - Edwards). The Minuteman family of missiles have a notable place in the American ballistic missile history. With its elegant architecture, Minuteman missiles almost completely replaced their liquid propellant predecessors such as Titan, and Atlas. Until the introduction of the MX (e.g. Peacekeeper) missiles to match Soviets’ SS-18 (Satan) missile, Minuteman constituted the backbone of American ICBMs. The majority of the Minuteman missiles (I, II and III) were test-launched from Vandenberg Air Force Base. It took on average 30 minutes for these missiles to hit their mock target in Kwajalein Test Range in Marshall Islands (see AP 1982, AP 1984).

Minuteman’s global fame was put into question when Soviets tested their SS-18 ICBMs – codenamed Satan. These Soviet missiles could carry up to 10 warheads and were nuclear capable (Encyclopedia Astraunomica – SS-18). As a response to SS-18, in the early 1980s, the Reagan Administration sought to push a bill to start the development of the MX (Peacekeeper) missiles. President Reagan pushed heavily in the Senate to secure funding for the MX missiles since the development entailed tremendous cost. In fact, many policy analysts and senators saw MX as de-balancing and an unnecessary addition to the American missile arsenal (The NY Times 1983). Ultimately, the MX was first test-fired in June 1983 from Vandenberg, California -home of the Minuteman missiles (The NY Times

56 The introduction of Minuteman did not cause the US to drop testing its previous missiles such as Jupiter, Thor and Titan. These latter earlier missiles were able to carry warheads with considerably large blast radius (including nuclear ones). Thus, decommissioning these missiles immediately to make room for Minuteman series would cause significant reduction in American fire power compared to the Soviets (The Washington Post 1979).
1983). Most of the MX tests were declared to be successful and potentially showed the American supremacy over Soviet missiles.

After the Cold War, the US shifted its attention from developing new ICBMs to testing anti-missile systems. There may be two reasons why this is the case. First, countries like North Korea, China and Iran, which opposed American interests in different parts of the world, developed their own ballistic missile industries and posed a great threat to the allies of the US. Consequently, the US sought to supply its allies such as South Korea, Taiwan and Israel with missile shields capable of intercepting ballistic missiles from their neighbors. For example, US Missile Defense Agency frequently conducted THAAD missile interceptor tests since the late 90s (MDA 2018). THAAD systems were later exported to Israel and South Korea. Besides THAAD, Patriot class anti-missile systems (PAC-2 and PAC-3) were often tested in White Sands Test Range in New Mexico. Lastly, GMD anti-missile systems were tested in Vandenberg.

The second reason why we observe more anti-missile/missile interceptor tests from the 1990s on is the delay in the launching of the Star Wars program that President Reagan pushed for in the 1980s. While the Star Wars system was initiated by the Reagan Administration, due to high costs, the first tests had to wait for a decade. Unfortunately, the first two Star Wars tests in 1991 and 1992 were reported to be failures which decreased support for continued funding. Additionally, environmentalists took these tests to courts and succeeded in banning them twice (NYTimes 1991, The AP 1992). Overall, for the Star Wars tests, ERIS, GMD or GDI missile interceptors were used to intercept Minutemen missiles launched from Vandenberg. The warning system for these tests were based on space satellites (ibid).
In terms of naval weapons, during the Cold War, the major challenge for the US was the Soviet nuclear Typhoon submarines, and the nuclear-capable ICBMs that these submarines could launch. In that regard, the US developed naval missile families such as Polaris and Poseidon which were typically test-fired near Cape Canaveral, Florida from 1959 and on (EA – Cape Canaveral). However, it was the ship-launched Tomahawk cruise missiles which showed the supremacy of the American Navy during the first Gulf War in 1991. Tomahawk missiles were test-fired mostly during 1980s, after the US decided to replace Trident missiles.

Figure 2 displays the yearly number of missiles test-fired by the United States between 1949 and 2015. Overall, several patterns in American post-World War II missile tests are worth noting. The first is the early 1960s where the yearly number of missile tests peak. There may be two potential reasons for this peak: first, the Cuban Missile Crisis occurred in 1962. Second, the Vandenberg and Cape Canaveral Air Force Bases were particularly busy with the tests of newly introduced missiles such as Minuteman I series as well as ship-launched missiles such as Polaris. In addition, the total number of nuclear detonation tests (shown with red dashes) peak in early 1960s. From this period and on, despite short term increases, we observe a general decline in the missiles that the US test-fired. Intuitively, one would expect a sharp decline in missile tests after the SALT talks which started in 1969, and after the end of Cold War in 1991. Yet, we see that the decline already started shortly after the Cuban Missile Crisis when the US and the Soviets came to the brink of nuclear war.

In the Post-Cold War era, the decline in the yearly number of missile tests slows down. However, the number of missiles launched in this period never reaches the Cold
War era figures. This presents an interesting issue: The United States fought two wars against Iraq and one against Afghanistan. The country also became involved in various conflicts around the world. Furthermore, new nuclear actors such as North Korea and Pakistan emerged. These countries test-fired some of their most destructive missiles in the post-Cold War era. A possibility is that none of these conflicts were as threatening to the US as the Cuban Missile Crisis. However, at the same time, we know that the US did not develop new ICBMs in the Post-Cold War era. The innovative years of late 1950s and early 1960s for missiles were not repeated in the 1990s. This could be because the Tomahawk missiles proved to be quite effective in achieving decisive victories in the First Gulf War and that the US did not see the immediate need of developing new missiles.

**Figure 2: American Missile Tests (1949-2015)**
(ii) USSR & Russia:

Immediately after World War II, similar to the United States, the Soviet Union hired German scientists. These scientists started to rigorously work on designing and building missiles (EA – Russia). In 1949 the Soviet Union started a four-decade long arms race with the US by detonating its first nuclear device. During these four decades, the Soviets test-fired countless missiles including the world’s first ICBM. At the end of the Cold War, similar to the US, the Russians significantly reduced their test-launches. That said, unlike the US, the Russians started developing new missiles such as Sineva, Topol’ M and Bulava, after a relatively calm decade following the collapse of the USSR. In fact, it is not difficult to come by a foreign policy analysis article that suspects that a new Cold War will start due to Russia’s revived ambition to have the world’s most destructive missile arsenal. Figure 3 displays the Soviet and Russian missile tests.

Figure 3: USSR and Russian Missile Tests
Figure 3 shows that the frequency of Soviet missile and nuclear warheads/devices tests peaked in the early 1960s which coincides with the Cuban Missile Crisis (1962).\textsuperscript{57} A similar peak was observable in Figure 2 in the case of the United States. The only difference is that although there was a decline in Soviet missile tests from 1962 and on, this decline seems to be slower than that of the United States in the same period. Moreover, in the early 1970s, there is a revival period which still does not reach the peak of the early 1960s.\textsuperscript{58} From the 1980s and on, however, despite the Afghan War, the Soviet missile tests experience a sharp decline in frequency. Then, similar to the US, the post-Cold War missile tests of the Russian Federation never reach the frequency of the Cold War tests.

A more in-depth look into the Soviet missile history shows that there are three main Cosmodromes from which a very large number of military-based space rockets and missiles were test-fired. First, between 1949 and mid-1980s, the Kapustin Yar in Astrakhan was scene to test-launches of many famous Soviet ballistic missiles such as R-1, R-1A, R1-2e, R-2, R-5M\textsuperscript{59}, R-12.\textsuperscript{60} Among these, the earliest ballistic missile, R-1 was developed from Von Braun’s V-2 schematics.\textsuperscript{61} Many Soviet R-type missiles also constitute the basis for missile technologies developed in China, India, Iraq, as well as North Korea (and, indirectly, Iran). For example, the Soviets sold R-2 missile schematics to China in the early 1950s which helped the Chinese to develop their Dong Feng missiles (EA – R-2). The short

\textsuperscript{57} Collecting data to produce Figure 3 presented a particular challenge due to the secretive nature of the Soviet weapon development programs. In fact, Nexis Uni and Missile Threat Project did not have much information / news about Soviet tests before 1970s. Thus, relying only on these two sources would give the impression that Soviets tested almost nothing but nuclear devices between 1949 and 1970. It is for this reason that I heavily relied heavily on Encyclopedia Astraunomica to code Soviet weapon tests between 1949 and 1970.

\textsuperscript{58} One could suspect that the SALT talks did not have any visible impact on Soviet missile tests.

\textsuperscript{59} R-5M is used to test nuclear warheads (Encyclopedia Astraunomica – R-5M)

\textsuperscript{60} R-12 IRBM missiles were the ones that Soviets placed in Cuba during Cuban Missile Crisis (Encyclopedia Astraunomica – R-12)

\textsuperscript{61} See Encyclopedia Astraunomica – Kapustin Yar.
range R-17 which we know better under the name *Scud* constituted the core of North Korean and, later, Iraqi missile arsenals (Encyclopedia Astronomica – R17). The Scud missile was first test-launched on February 2\textsuperscript{nd}, 1962 from Kapustin Yar Test Range.

The second Cosmodrome is the Baikonur in Kazakhstan. Baikonour overall was the busiest test-range during the Cold War where Soviets launched many ICBMs. In fact, the very first ICBM in human history was first test-launched from Baikonur (Encyclopedia Astronomica – Baikonur). In addition, one the most feared ICBMs of the Cold War, the SS-18 (Satan) was first test launched from Baikonur in 1971 (Encyclopedia Astronomica - R-36M 15A14). From the early 1980s on, the Soviets conducted substantially more scientific/space-related tests than test lunches of its ICBMs from Baikonur Cosmodrome.\textsuperscript{62}

Last, but not least, the Soviets frequently conducted tests in Plesetsk which is in Arkhangelsk Region. Although frequently used for scientific rocket launches, there are still many notable missiles that were first test fired in Plesetsk. One such missile is Topol’ (Encyclopedia Astronomica – Plesetsk). Topol’ is an ICBM that can be launched from a ground silo or from a submarine. It was first tested in September 1981, and eventually replaced its predecessors UR-100 and UR-100MU. After the fall of the Soviet Union, the Russian Federation continued test-launching Topol missiles from Plesetsk. During the second half of the first decade of the 2000s, Russians developed and tested new versions of Topol’, such as Topol’-M and Rubezh (EA– Topol). According to my data collection, President Putin’s era from the 2000s on saw an increase in Topol test-launches (a pace not

\textsuperscript{62} These are orbital rocket launchers belonging to Molniya, Vostok, Tsiklon, Buran/Energia, Proton and Soyuz family (Encyclopedia Astronomica - Baikonur). It is difficult to suggest that these rockets cannot be also used for military purposes. However, the tests recorded by Encyclopedia Astronomica’s records indicate that these rockets only served scientific / civilian missions. Thus, none are recorded in the present dataset.
observed in the 1990s). However, after 2010, Russia test-launched new SLBMs to replace Topol’ family of missiles such as Bulava, and Sineva.\textsuperscript{63}

There are also several weapon-specific Russian test ranges that were not as busy as Kapustin Yar, Plesetsk or Baikonur. Among these, Sary Shagan in Kazakhstan is one such test range where Soviets and Russians exclusively tested missile interceptors (Encyclopedia Astraunomica – Sary Shagan). S-300 and S-400 are some of the notable missile interceptors that Russia recently sold (or are planning to sell) to various countries. In addition, the Soviet Union and then Russia used a test range in Barents Sea to test-launch some of the Topol-family long or medium range submarine-launched ballistic missiles. Similar to Sary Shagan, Barents Sea Test Range has a lower test-launch traffic than Baikonur, Kapustin Yar and Plesetsk. We also observe a discernible increase in Topol-type missiles such as Layner and Sineva launched from submarines around Barents Sea Range from 2010 and on.\textsuperscript{64}

Overall, it appears that President Putin’s administration is more interested in test-firing naval missiles than land-based missiles – a pattern we also see in the UK and France. This is why, from 2010 and on, we observe more submarine-based Russian missile tests. However, Russia’s greatest handicap in the post-Cold War era was economy. This may be one of the potential reasons why the country kept test-firing Topols – to ensure the longevity of these missiles throughout the 1990s (Missile Defense Project SS-25). It was not until the 2010s that the Russian pursuit of developing new and destructive missiles took a new pace.

\textsuperscript{63} These last two missiles were tested in other test ranges than Plesetsk.
\textsuperscript{64} Unlike the Cold War era, news of these test launches can be found on Nexis Uni and Encyclopedia Astraunomica at the same time.
(iii) China:

From the early 1960s on, China started developing and then test-firing many missiles that drew the attention of the Soviet Union and, later, the United States. As Figure 4 displays, there were two periods when the yearly number of Chinese missile tests peaked. The first was in 1980 immediately after the 1979 Third Indochina War (see Sarkees et al 2010). The second (and the highest) peak we observe was in the mid-1990s and coincides with the Second Taiwan Strait Crisis. These tests were of concern not only to Taiwan, but also to the United States. After this crisis, the yearly number of Chinese tests sharply declined.

Figure 4: Chinese Yearly Missile Tests

China’s most popular family of missiles has been the Dong Feng (DF). The first DF missile was test-fired in 1960 from the country’s busiest test range: Shuang Cheng Tzu in Jiuquan Province (Encyclopedia Astraunomica – Jiquan). The DF missiles were
developed based on Soviet R-2 missile schematics. In fact, Chinese scientists received direct assistance from Soviet scientists to develop the first versions of DF missiles (Encyclopedia Astraunomica – DF 1).

From the early 1960s to 1981, China seems to have only test-fired DF missiles in addition to its various nuclear tests. The DF family of missiles I recorded are DF-1, DF-2, DF-2A, DF-3, DF-4, DF-4 CCS3 and DF-5. The ranges of these missiles vary from IRBM to ICBM. The DF-5 has the longest range (7,500 miles) of these missiles and can carry a nuclear warhead (Encyclopedia Astraunomica - DF-5) Most DF test-launches in this period were conducted in Shuang Cheng Tzu test range.

Another notable observation from China is its heavy and successful investment in space rockets. The DF missile technology was used as a basis to develop orbit launch rockets.\(^{65}\) Although the Soviets and Americans surpassed Chinese rocket technology in the Cold War, Chinese tests presented an unparalleled rate of success. In a report written in Aerospace America (1989), China is recorded as having launched 24 military-based rockets with satellites with only 2 failures – a record ratio superior to that of the US and the Soviets. This report also attributes the success to a Chinese MIT and Caltech scientist Qian Xuesen who oversaw China’s rocket development program at that time (ibid).

Lastly, the years between 2000 and 2015 included particularly notable missile tests that showed China’s interest in developing space-based weapons. For example, in 2007, China launched an IRBM to destroy an old satellite in space.\(^{66}\) This test was a surprise to

---

65 For example, DF-31 which is a long-range missile, was used to develop KT-1 orbital launch rocket (Encyclopedia Astraunomica – KT-1)

66 Contrast this with Russia which conducted its first public anti-satellite test as USSR in June 1982 (The AP 1982a).
the US as well as other potential rivals of China (Facts on File 2007). In January 2014, China tested its first hypersonic glider: WU-14. This glider could potentially launch missiles from space. The Americans reported that they have no weapon that travels faster than WU-14. In addition, the Chinese were ahead of the Russians who test-fired a similar device a year later. (ValueWalk 2015) However, despite many similar innovative tests from 2000 and on, the yearly number of Chinese missile tests never came close to its peaks in late the 1970s and mid-1990s.

(iv) India:

Since its independence in 1947, India experienced considerably conflictual relations with both China and Pakistan. It fought one war with China in 1962, and four wars with Pakistan to this date (see Sarkees et al 2010). Besides the direct interstate wars, India and Pakistan sponsored militant non-state groups against each other due to disputes over Kashmir and Baluchistan regions (Bapat 2012). In brief, India has not enjoyed a moment of peace due to many ongoing conflicts in the shadow of recurring wars in its region.

Unlike the US, USSR, Russia and China, the particularity of India is that once it started its missile development program in 1983\textsuperscript{67}, its yearly number of missile tests always increased (with only a handful of exception periods). This pattern can be clearly observed in Figure 5. The increase in India’s yearly missile tests become more discernible from 1995 – shortly after Pakistan initiated its own missile development program. Interestingly however, the early 2000s and late 1990s which coincide with the Kargil War do not

\textsuperscript{67} See AP International (2003).
coincide with unique peaks in Indian public missile tests. The Kargil War is potentially the most serious conflict India had after Pakistan became nuclear in 1998. However, unlike the potential impact of Cuban Missile Crisis on American and Soviet missile tests, the Kargil War does not seem to have particularly affected Indian missile tests.\(^\text{68}\) Ultimately, the largest number of missile tests appear to have been conducted by India in 2014.

**Figure 5: Yearly number of missile tests by India**

Despite its first nuclear detonation in 1974, India’s first major missile test took place in February 1988 when a Prithvi missile was test-fired.\(^\text{69}\) Prithvi was a short-range missile, and comparable to the Soviet Scud missiles by design. In fact, according to a report, Prithvi missiles have superior destructive capabilities compared to Soviet Scuds. (Agence France-Presse 1996). Prithvi and Prithvi-II missiles were publicly test-fired on 50

\(^{68}\) For example, missile tests conducted by India in 2000 are fewer than those in 1999 according to the data I collected which can be seen in Figure 5.

\(^{69}\) In 1986, India test-fired a short range missile Trishul. However, the Prithvi family of missiles are among the most important ones in India’s arsenal. Thus, it is safe to consider 1988 as the year India had its first major ballistic missile test.
occasions which makes them the most tested family of Indian missiles. India test-launched Prithvi missiles mostly from Bay of Bengal in Chadipur-On-Sea test range.

The next most-tested Indian family of missiles are the Agni. The first version of Agni was a short-range ballistic missile which was publicly tested in May 1989 from Chandipur-on-Sea test range. In 1999, Agni-II was test-fired for the first time. This new missile from the Agni family had a range of 1,500 miles. With this range, India showed its potential to strike any city in Pakistan as well as Beijing in China. Later on, multiple versions of missiles from the Agni family were developed and test-fired publicly. Ultimately, India tested the fifth member of the Agni family, Agni-V, in April 2012. Agni V was nicknamed by the press as “China-Killer”. It had the potential to reach Shanghai as well as Beijing. Unlike its predecessors\(^70\) which were 2-stage missiles, Agni V was a 3-stage missile with an operational range of 3,200 miles (Missile Defense Project– Agni V).

We also observe that India closely cooperated with Russia in developing cruise missiles. One such missile is BrahMos, a short-range missile that was jointly developed by India and Russia.\(^71\) This supersonic cruise missile was often test-launched from a naval vessel. Different versions of BrahMos (Block I, Block II and Block II) were publicly test-fired about 41 times in the dataset. As of 2015, BrahMos technology does not appear to have been matched by Pakistan which often receives assistance from Chinese scientists.

Lastly, India is also is one of the handful of countries in the world that developed and tested its own missile shield system. India’s first missile shield system test took place in November 2006. A test range in Wheeler Island in East India was used to conduct these

---

\(^{70}\) Agni I, Agni II, Agni III and Agni IV

\(^{71}\) BrahMos is possessed by the Indian Army. Therefore, I coded tests of BrahMos as only Indian, and not Russian.
missile shield tests. As missiles to be intercepted, Prithvi-class short range missiles were launched from Chandipur-On-Sea test range (BBC 2006). Only a handful of countries such as the US and China were capable of developing their own missile shield systems around 2006. Ultimately, after launching the Agni V and testing its own missile shield system, India joined the league of countries with superior ballistic capabilities such as the US, Russia and China.

(v) Pakistan:

For the majority of its missile development history, Pakistan had to focus on one single rival: India. In fact, the majority of Pakistan’s missiles were designed to counter missile attacks from India. Most of Pakistan’s public missile tests were covered by various media sources as a direct response to India missile tests. However, economic circumstances and military coups have been major obstacles for Pakistan in its quest to match India’s observably superior missile strike capabilities.

Figure 6 displays the yearly missile test trend of Pakistan. Similar to India, Pakistani missile tests continue to increase from the first year of observation (1988) until 2015. However, this increase is less consistent and slower than that of India. In addition, India, on average, tested twice as many missiles per year than Pakistan. From 1998 on, the increase in Pakistan’s missile tests become more discernable than before. This could be because Pakistan’s indigenous missile development program began in 1998. The program was supposedly backed by Chinese scientists (Missile Defense Project - Pakistan). However, the temporary peak in the number of missile tests in 1998 is not surpassed until

---

72 See [https://www.nti.org/learn/countries/pakistan/delivery-systems/](https://www.nti.org/learn/countries/pakistan/delivery-systems/) and Missile Threat Project - Pakistan
2010. A potential reason for this could be economic hardships as well as the 1999 military coup, both of which could have made it difficult to develop and test new missiles. Nevertheless, there does seem to be a visible increase in Pakistani missile tests from 2010 and on. The average number of yearly missile tests after 2010 was 9, whereas this number was 5 between 1998 and 2009. The global peak in this country’s missile tests is in 2012 with 13 public missile tests.73

Figure 6: Yearly Missile Tests of Pakistan

Pakistan’s main missile family is Hatf. Early versions of this missile family were based on Soviet Scud technology. In fact, the first public missile test that I was able to identify for Pakistan was a Scud (R-17) missile – launched in April 1988. The first Hatf-type missile that was seen as a serious threat by India was test fired in 1998. This was a Hatf V Ghauri missile which was potentially able to penetrate deep into India. Ghauri had

73 As Figure 5 displays, India’s global peak was in 2014 with more than twice of the maximum amount of Pakistan.
a range of 900 miles. Another incident that made Pakistan more threatening to India was the former’s first nuclear explosion test in 1998. In fact, according to several news reports, Pakistan detonated a total of seven nuclear warheads on two days - May 28 and May 30, 1998 (Deutsche Presse-Agentur 2003, NBC News Transcripts 1998). After 1998, Pakistan showed that it was determined to close the missile and nuclear technology gap with India. However, it was not until 2005 that Pakistan managed to test its first nuclear-capable missile, Hatf VII Babur.

Once Pakistan’s missile tests started to become a concern for India, the two countries seemingly engaged in a series of tit-for-tat missile tests. For example, according to the several news sources, the two countries quickly reciprocated each other’s missile tests in 1998, 1999, October 2002 and 2003. During these periods, it was often the case that India would test-launch a missile (or detonate a nuclear device) and Pakistan would reciprocate it with a test of its own. Yet, as I highlighted earlier, Pakistan’s economic constraints and struggle with internal conflicts made it difficult to catch up with India’s often superior missile technologies.

Overall, similar to India, most missile that Pakistan test-launched were short range, with a handful of exceptions. The missile with the longest range in Pakistan’s arsenal is Shaheen III which can potentially reach targets 1,700 miles away. We know that India has

---

74 One claim is that Pakistanis received Chinese assistance in developing the Ghauri missile. Moreover, some argue that Pakistanis relied on Chinese M-11 missiles to develop Hatf V Ghauri (Missile Defense Project – Hatf V).

75 Pakistan’s Hatf VI Shaheen-II is potentially a nuclear-capable missile which was first test-launched before Babur. Yet, in the case of some missiles, it takes time and may tests to determine whether they can carry nuclear warheads or not. According to my survey of the news reports, Shaheen-II missile started to be labeled as nuclear capable only from 2006 and on. Additionally, Shaheen class missiles are ballistic and not cruise. Thus, Babur’s ability to dodge radars – which is not the case in Shaheen – makes the former more concerning to Indian Army. Shaheen’s only superiority over Babur is its longer range.
longer-range missiles, as well as a missile shield technology - something which Pakistan did not possess or test until 2015. Yet, a seeming competition with India pushed Pakistan to develop its own missiles – despite considerable economic burden.

(vi) North Korea:

North Korea has been one of the most media-covered countries when it comes to missile tests. Especially since Kim Jong-Un replaced his father in 2011, a seemingly-increasing number of missiles were test-launched. A look upon Figure 7 shows that this increase actually started in 2006 when the first nuclear detonation test was conducted during Kim Jong-Il’s reign. On the other hand, between 1984 and 2006, North Korea’s yearly number of missile tests are low and do not show any increasing or decreasing pattern. In fact, the post-2006 increase in North Korean missile tests was not consistent. For example, the number of tests in 2007 and 2008 are lower than that in 2006. However, in 2009, North Korea conducted the highest number of missile tests with 28. Following a steep decline in 2010, North Korean missile tests start to continuously increase from 2011 and on (which marks the first year Kim Jong-Un de facto becomes the leader of North Korea).

\(^{76}\) 2009 is also the year when North Korea tests its second nuclear device and receives a major UN sanction as a result.
Before becoming a missile exporter to countries like Iran, Pakistan and Syria, North Korea’s earliest missiles were directly based on Russian Scud schematics. The first known North Korean public missile test (in April 1984) was of a Hwasong-5. Hwasong-5 was essentially a slightly modified version of Soviet Scud B missiles. However, in 1990, North Korea test-fired a Nodong-1 which was a more indigenous missile (EA – Nodong 1). Nodong-1 missiles were later sold to countries like Iran and Pakistan (ibid). In addition to Nodong and Hwasong missiles, North Korea’s other families of well-known missiles are KN and Taepedong. The KN family of missiles (KN-1, KN-2, KN-6, and KN-11) are the most-frequently test-launched ballistic missiles of North Korea according to my dataset. On the other hand, Taepedong-I and Taepedong-II are less often test-launched but are frequently used for space missions.

In terms of tests sites, I was able to identify three locations. These are Kittaeryong Missile Test Site, Tonghae Satellite Launching Ground and Chiha-ri. Kittaeryong is the
site where most of the Nodong-I and Hwagong class missiles were test fired (Encyclopedia Astraunomica - Kittaeryong). On the other hand, Tonghae is the site where North Korea’s longer range and most dangerous missiles such as Taepedong I, II and Unha-3 were test-fired (Encyclopedia Astraunomica - Tonghae). I was able to record only two test-fires from Chiha-ri, both of which are Hwasong-6. 77

(vii) Iran:

Shortly after Iran became an Islamic Republic in 1979, three countries became its arch-rivals. The first was the United States. Not only did the Iranian Islamic Revolution have an anti-American tone, but Iranian revolutionary youth invaded the American Embassy in Tehran and marked the beginning of bitter Iranian-American relations. In addition, following the Revolution, Iran shut down most American-based offices including all the centers the US used to detect Soviet missile telemetry data. The second rival of Iran was the Saddam Regime of Iraq. Between 1980 and 1988, the two countries fought a very costly war. Finally, Iran’s new constitution involved anti-Israeli clauses which created conflict between the two countries. 78

Figure 8 reveals that there is no clear increasing or decreasing pattern in the yearly number of missile tests that Iran publicly conducted. It is rather the case that certain periods involve sharp increases in the frequency of tests. These periods were in 1997, 2006, and, more recently, in 2011. Each time, the missile test numbers experience sudden increases that are not sustained in the long run. In addition, these peak periods coincide with Iranian

---

77 That said, most of the weapon test locations in the case of North Korea were not identified. This country, similar to Pakistan, is notorious for not revealing where a missile was test-fired.
78 One could easily argue that Saudi Arabia also pays close attention to Iranian missile tests. However, I rarely saw any Saudi press reports on Iranian tests.
naval exercises around the Strait of Hormuz, during which Iran displays its missiles and naval vessels to the rest of the world. In 2011, Iran’s missile tests were at a peak with 48. Note that this number is greater than the maximums of China, North Korea, India and Pakistan.

Figure 8: Yearly Missile Tests of Iran

Iran initially received substantial North Korean support to develop its own missile program. Ultimately, Iranian missile development program began around 1985. Facing threats from different nations in the Persian Gulf, it was not surprising that Iran’s weapons program focused on naval defense. In that regard, most of the test-launches I recorded are of ship-launched missiles during naval exercises. Some of the missiles launched during these naval exercises were short range anti-ship missiles such as Qader, Nasr, Noor and

---

79 For example, in October 1998, an Iranian naval exercise involved test-firing a long-range missile. The exercise took place one week after the US deployed its USS Nimitz to the Persian Gulf (Calgary Herald 1997). 80 There are several sources that point to later years as the beginning of Iranian indigenous missile development program. However, according to Encyclopaedia Astrononimica Iran had already test fired its own short-range Shahab-I missiles 8 times in 1985 from Kermanshah Air Base (EA – Shahab I). Following these tests, a test launch of a Fajr-3 is reported to have been conducted in 1990 (EA – Iran)
Ra’ad. Iran also test-fired short-range cruise missiles such as Zafar, and, less frequently, medium range cruise missiles such as Ghadira.

There are several land-based missiles that Iran appears to have tested most of which are short-range. I was able to identify five test range locations for these missiles which are Kermanshah, Khatam-ol-Anbia, Arak, Qom and Imam Khomeini Space Center. Among these, Imam Khomeini Space Center was the busiest test-launch site between the years 1992 and 2015. Iran’s long-range missiles such as Shahab-III, Kavoshgar-5 and Sejjil are frequently test-fired from Imam Khomeini Space Center (Encyclopedia Astraunomica - Semnan). Sejjil has the longest range among Iranian missiles, with 1,242 miles. This missile was also launched for civilian/space research purposes. Interestingly Sejjil missiles, as well as other space rocket launches, were most frequent during President Ahmadinejad’s tenure around 2010 (Encyclopedia Astraunomica – Sejjil).

(viii) Iraq:

Iraq’s missile testing trend was mainly shaped by the wars it fought between 1987 and 2003. As can be seen in Figure 9, there were two periods when the number of missiles that Saddam Hussein’s regime test-launched irregularly peaked: 1990, and 2001. More specifically, from 1987 to 1990, a steady increase in the number of missile tests can be observed. However, in August 1990, Iraq invaded Kuwait which marked the beginning of the First Gulf War. This war caused substantial damage to Iraqi military bases and infrastructure. Many Iraqi missile test ranges were wiped out by American Tomhaw cruise missiles (EA – Iraq). The devastating effects of the Gulf War on Iraq may be the reason we do not observe any Iraqi missile tests between 1992 and 1996. However, from 1997 on, the Saddam Regime appears to have re-started test-firing missiles. This increase also coincides
with the increasing tensions between the US and Iraq. Ultimately, the number of Iraqi missile tests reached its maximum in 2001 with 36 missile tests. The following year, Iraq tested another 32 missiles. However, from 2003 on, I was unable to record any public missile tests. This is likely due to the US occupation of Iraq starting in 2003.

**Figure 9: Yearly missile tests of Iraq**

![Graph showing yearly missile tests of Iraq](image)

Saddam Regime from its first day till its collapse posed serious threat to its neighbors around Mesopotamia. However, a closer look at the Iraqi arsenal indicates that this country’s missiles were never advanced and effective enough to cause significant damage to hostile countries. Moreover, Iraq predominantly test-fired short range ballistic missiles with old liquid propellant engines. Of the 98 test-launches I was able to code, 89 were SRBMs and 9 were IRBMs. Most of the Iraqi SRBMs such as Al Abbas, Al Fatah and Al Samoud were based on old Soviet Scud (R-17) schematics (EA – Iraq). The only two IRBMs I was able to record are Al Hussein and Al Aabed missiles. Al Hussein was another modified R-17. The Iraqi engineers doubled the range of the missile at the cost of reducing its warhead payload (EA – Al Hussein). Al Aabed was test-fired only once and
was declared operationally capable in 1995. However, 7 years later, it was discarded from service (EA – Al Abed).

A major challenge was to find information about Iraqi missile test locations. This challenge was most specifically about determining the coordinates of the test locations. One potential reason, as also highlighted in EA- Iraq is that following the First Gulf War, most of the test locations were destroyed. Therefore, even though the names of the locations were known, their exact coordinates were lost. This is why, the longitude and latitude variables for Iraqi test locations are predominantly missing between the two Gulf Wars. Ultimately, I was able to uncover 5 different test locations which are Tall Afar, Al Anbar, Jebel Himreen, Al Nikheb, and Umm Qasr. Of these ranges, only Al Anbar survived the first Gulf War according to the information I collected.

(ix) Israel:

Israel’s major challenge has been that it is located in a very hostile region. It fought multiple wars against its neighbors since the day of its independence. In addition, Israel has been wrestling with terrorist organizations such as Hamas and Hezbollah which supposedly received funds and other forms of military support from countries like Lebanon and Iran. Amidst all this tension of conflict, Israel managed to create a missile program of its own. However, it is also the country that test-fired the least number of missiles in the sample with 68 in total. This raises the question of why Israel did not test-fire as many missiles as one would intuitively expect from a country that has been under serious threat by most of its neighbors.
Figure 10 displays the yearly missile test launches of Israel. Overall, there is no clear increasing or decreasing pattern in the Israeli missile tests. The median yearly number of Israeli missile tests between 1987 and 2015 is only 2. The maximum number of missile tests – which is 7 – is recorded for the year 1993. In this year, Israel test-fired multiple Arrow-1 missile interceptors as well as one Barak-1 interceptor and a Gabriel SRCM. The second largest number of missile tests is recorded for the year 2003 with 5 missiles. Interestingly, neither 1993, nor 2003 coincided with Israel’s most conflictual periods.

A closer look at Israel’s missile arsenal reveals two important names: Jericho and Arrow. Jericho’s first version was developed in France under the name MD-620 in early 1960s. This SRBM was specifically manufactured for Israel (EA - Jericho). However, the first public missile test of Israel that I was able to record was from 1987 which was a launch of Jericho-2. Jericho-2 was a nuclear-capable MRBM developed by Israel (Missile Defense Project – Jericho II, EA – Jericho II). The third version of Jericho was first test-fired in 2008. This version was also nuclear capable, but an IRBM (MT – Jericho III). On the other hand, according to my data collection, Arrow, which is a missile interceptor, was the most
frequently test-fired weapon of Israel. It had three versions, and 40 percent of the missiles test-launched by Israel were either Arrow-1, Arrow-2 or Arrow-3 between 1987 and 2015. Arrow missile interceptors were manufactured by the US and exported to Israel.\textsuperscript{81} (EA - Arrow). Tests of Arrow and Jericho were primarily conducted in the Palmachim Test Center in Palmachim Beach (EA – Israel). This launch site appears to be the most-frequently used place where Israeli missiles were tested.

(x) The United Kingdom:

In the aftermath of the World War II, the UK joined the club of nuclear countries. Its first nuclear detonation test was conducted in 1952 (WNE). Following the US, USSR, and France, the United Kingdom is the fourth most nuclear test-conducting country in the world (ibid). However, the total number of missiles that it test-fired remains low. Of the 207 total tests I recorded, only 83 are actual missile test-launches and the remaining are nuclear tests. That said, the post-World War II history of England was not short of conflicts. From 1945 to 1970s, the British had to deal with the independence movements of many of its African colonies. Then, despite economic hardships, it fought a war against Argentina in 1982. In addition, England joined many UN-led as well as US-led operations across the world.

\textsuperscript{81} The first Arrow is a complete American design by Boeing company (EA – Arrow). However, Israel contributed to the design and manufacturing of Arrow II and especially Arrow III (Missile Defense Project – Arrow III). In addition, Israel has its own indigenous missile interceptors such as Barak-1 which appear to have been less frequently test-fired than Arrow.
As Figure 11 displays, England had two peak periods of missile tests. The first notable peak happened in 1982 which coincided with the Falklands War. During this year, the British tested 18 missiles and detonated one nuclear device. Five years after the Falkland Wars concluded, the UK recorded its peak number of missile tests at 22. Beyond the years 1982 and 1987, we do not observe a sudden increase or decrease in the UKs missile tests. Between 2010 and 2014, there appears to be an increase in the number of missile tests. Nonetheless, the UK never reached the number of missile tests it conducted in 1987 again.

The data I collected for Britain shows that the majority of the missiles the country test-fired are naval-based ballistic or cruise missiles. Before the end of 1980s, the most-tested missile was the Polaris A-3. This was an American-manufactured long-range cruise missile. Polaris was also capable of carrying nuclear warheads (EA– England, EA- Polaris A-3). In early 1990s, the United Kingdom purchased Trident missiles from the United States (Press Association 1991). Between 1989 and 2012, the UK frequently test-fired
Trident and Trident II ICCMs from nuclear submarines.\textsuperscript{82} Another notable cruise missile imported from the US was the Tomahawk. Observing the success of Tomahawk missiles in the First Gulf War, the John Major Cabinet ordered the purchase of sea and land-based versions of these missiles in 1996. Shortly after this year, Britain conducted a series of Tomahawk missile tests.\textsuperscript{83}

Lastly, Britain also test-fired surface-to-air, air-to-air missiles as well as missile interceptors. These missiles typically had shorter operational ranges and were test-launched less frequently than naval missiles. Typically, tests of these missiles were conducted in Woomera - Southern Australia and Aberforth – Wales. Of these missiles, Brimstone was one of the most well-known short range air-to-air missiles developed by the Royal Air Force and sold to other countries such as Saudi Arabia (Missile Defense Project – Brimstone).

Overall, the United Kingdom’s naval missiles were predominantly imported from the US. On the other hand, the country also developed its own short range missiles. Interestingly, tests of these indigenous missiles were much less frequent than the longer range naval missiles. An important detail is that the UK also manufactured civilian sounding rockets in the 1960s and 1970s (EA – Britain). In fact, despite its many militarized conflict involvements, the UK managed to focus on launching more civilian rockets than test-firing military missiles (\textit{ibid}). One reason could be the United Kingdom’s strategic partnership with the US. After all, thanks to this partnership, Britain had exclusive

\textsuperscript{82} England’s early Trident tests were conducted with the US, in Cape Canaveral – Florida. Later Trident missile tests were conducted around British Isles, and South Sardinia. For these latter tests, British nuclear submarines were used.  
\textsuperscript{83} England is the only country other than the US that possesses Tomahawk missiles (Missile Defense Project-Tomahawk)
access to some of the world’s most destructive naval missiles such as Tomahawk, Polaris and Trident.

(xi) France:

Similar to the UK, France had to deal with the independence wars of its colonies in the aftermath of World War II. Algeria’s independence was one of the longest and costliest wars that France fought during this period. Besides these wars, the French also engaged in many overseas military operations in countries such as Ivory Coast, Libya and Haiti. France also participated in UN-led operations around the world. Needless to say, France had a conflict-ridden record in the post-World War II era. Given this record, it makes intuitive sense that the French were particularly ambitious in developing and testing their indigenous missile technologies as well as nuclear devices.

**Figure 12: Yearly Missile Tests of France**

![Yearly Weapon Test Numbers of France](image)

Figure 12 displays France’s yearly missile tests. There are two initial notable observations. First, France tested more missiles up until 1992 than after 1992. However, second, as the red line in Figure 12 indicates, a large portion of these tests are nuclear – a
patter which we can also observe in the UK case. That said, actual missile tests before 1970s were more frequent relative to nuclear device/warhead tests. Overall, the maximum number of missile launches was recorded in the year 1967 with 14. The peak in 1967 was not a sudden one. The increasing trend apparently started in 1963 which is 1 year after the end of Algerian Independence War. Another notable increase in missile tests happened between 1975 and 1978 (from 2 to 13 missile tests). The decline from 1978 on was slow-paced until 1989. From 1990 on, the French missile tests sharply declined, and never reached its 1978 or 1967 levels.

A closer look at France’s missile arsenal reveals two main attributes of this country. First, similar to England, the French heavily invested in naval-based cruise and ballistic missiles. Second, unlike England, France is a major missile manufacturing and exporting country. For example, it manufactured Jericho missiles for Israel – as highlighted earlier (EA – Jericho). Additionally, France was at the heart of the European Union’s attempts to develop its own arsenal independent of the US. In that regard, French scientists often collaborated with their German and British counterparts. One of the well-known products of such collaboration is air-to-air Exocet missiles (EA -Exocet). Exocet is also one of the most exported missiles in the world (Missile Defense Project– Exocet). Another example is Storm Shadow which is a short-range cruise missile (Missile Defense Project– Apache EP).

There are 3 major ranges where the majority of the French missile tests were conducted. Two of these sites are not in the French mainland. The first was Hammaguira

---

84 Strom Shadow is part of the Apache EP family manufactured by France. It is used by the UK, Italy, Greece, Saudi Arabia, United Arab Emirates (Missile Threat – Apache EP).
Test Center which is located in Algeria. France conducted many civilian and military missile tests in this range between 1947 and 1967. Earlier versions of the MSBS family such as SE4400, and SSBS S112 were often launched in Hammaguira. Following the independence of Algeria, France had to abandon this test range. The second test-range outside the French mainland was Ile du Levant. After abandoning Algeria, France moved most of its Hammaguira operations to Ile du Levant. This range was predominantly used for civilian sounding rocket tests. That said, we also observe tests of IRBM's such as SSBS M112 launched from Ile du Levant between 1959 and 1968.

The third test range that frequently appears in the present dataset is DGM (Landes) Test Center in Biscarrosse (mainland France). The French have been testing missiles in Landes since 1965. Most of the MSBS family of submarine-launched cruise missiles were deployed to and test-launched in DGM Test Center after France abandoned Algeria. The MSBS family involves intermediate range missiles (M013, SSBS S03, and SSBS S112) that were frequently test-fired between 1966 and 1979. More recent versions of the MSBS family missiles have longer ranges. Among these, M4 was frequently test-fired during the 1980s and has a range of approximately 2,500 miles. More recently, in 2006, the French Navy test-fired their longest-range missile, M51 which has a range of 6,000 miles, and which can carry nuclear warheads. Overall, the MSBS Family has been France’s most-test-launched indigenous missiles according to the data I collected.85

85 Besides these three locations, I was able to identify several other test ranges where France test-fired missiles much less frequently. One of these is the DGA Missile Test Center in Cazaux where France test-fired short-range Meteor missiles three times according to a report by Progressive Media – Company News (2012)
Discussion:

In this paper, I presented a new dataset on publicly conducted missile tests by 11 countries between 1949 and 2015. Two questions motivated my data collection on missile tests. First, why do countries test-fire their missiles publicly? Second, are public missile tests associated with the militarization of countries’ ongoing disputes? Both questions require separate and detailed statistical investigations to find answers. However, to conduct any such statistical work, we first need data. I presented a dataset that fulfills this purpose. Interestingly, no previous data collection attempt of this type had been made.

After describing the coding procedures for the creation of my dataset, I presented an overview of the 11 countries’ missile testing record. Each country has their unique history of international conflict, weapon manufacturing strategies, economic limitations, and military capabilities. However, beyond these differences, there seems to be one relatively common aspect: countries’ public missile tests coincide with periods of conflict. In other words, when countries unusually increase the number of missiles they test-launch in a given year, this often coincides with their involvement in a crucial – and often militarized – international conflict. As described earlier, both the US and USSR missile test peaks are observed around the Cuban Missile Crisis time. Also, Iraq’s two wars clearly define its missile test peaks. Similarly, China test-fired its missiles most frequently during the Taiwanese Strait Crisis. Overall, the co-occurrence of peaks in a country’s missile test numbers and its involvement in militarized interstate conflicts can potentially be generalized across the present sample.86

86 Israel might have been an exception
The case of the American and Russian missile tests also raises another question. Does one country perceive its adversaries’ missile tests as a threat and then reciprocates in kind. As described earlier, both American and Russian missile tests peaked around the Cuban Missile Crisis time. Shortly after this crisis, these countries’ missile tests similarly started to decline, and never reached their early-1960s level. In fact, as Figure 13 displays, in a very general manner, the Russian and American missile tests appear to have comparably similar trends both during and after the Cold War. Further statistical inquiries highlight that both countries’ number of missile tests strongly correlate in any given year. In addition, further analyses indicate that we can look at one of the two countries’ missile tests to predict the other within 95 percent confidence level – ceteris paribus.

Figure 13: A Comparison of the American and Russian Missile Tests (1949-2015)

To assess that, I created Russia’s and the US’ yearly number of missile tests between 1949 and 2010 as two different variables. Pearson’s $r$ between these variables is 0.80 with $t(65) = 10.42$ and $p<0.001$.

I used ‘tseries’ package for RStudio to calculate the potential cross-correlation measures between the US and Soviet/Russian missile tests (Trapletti and Hornik 2018). I created a plot that indicates the cross-correlation -ACF- between Russia and the US’ missile tests at different lags (with 10 year being the maximum lag assessed in the analysis).
Another case in which two country’s missile tests seem to follow a similar trend is India and Pakistan. In the previous section, I highlighted the possibility that the increase in India’s missile tests from late 1990s and on may be related to Pakistan’s decision to initiate its own missile development and nuclear programs. As can be observed in the left plot of Figure 14, India and Pakistan’s missile test trends show a somewhat similar pattern from 1998 to mid-2000s. Yet, India’s peak number of missile tests happened in 2004. It seems that Pakistan did not react to this peak with tests of its own. This temporary unrelatedness of both countries’ missile tests could be due to their ceasefire agreement in Kashmir in 2003 (BBC.com 2019). On the other hand, further analyses on the overall Pakistan and India missile tests between 1988 and 2015 show a strong yearly correlation. However, tests conducted by one of these countries do significantly predict those of the other within 95 percent confidence level.

Weak cross-correlation between India and Pakistan’s missile tests may be driven by the first decade of 2000s which includes India’s peak number of missile tests in 2004. Then, the question is what accounts for India’s peak number of missile tests in 2004. Typically, China comes to mind as another rival of India. However, the right plot in Figure 14 shows that India and China’s missile test trends do not appear to follow a discernibly similar pattern throughout 2000s. This may be due to the fact that the Indian Prime Minister made a historic visit to China in 2003. (CNN.com 2010). Such a historic visit may potentially turn both India and China more optimistic about not having to face a military

89 1988 is -as mentioned early on- the year when Pakistan tested its first ballistic missile publicly. Since we are comparing India and Pakistan, I do not include India’s earlier missile tests. Overall, Pearson’s r between Indian and Pakistan’s missile tests is 0.80 with (26) = 6.77 and 95% confidence level [0.61, 0.90]
90 See Figure 16 in the Appendix for the relevant Cross-Correlation Plot.
91 China’s most discernible response to threat in terms of missile tests still appears to be the aftermath of the Second Taiwanese Strait Crisis.
conflict in the near future.\textsuperscript{92} That said, I fail to find any statistically significant association between Chinese and Indian missile tests in general.\textsuperscript{93}

**Figure 14: Yearly Missile Test Trends of India vs its Adversaries**

![Figure 14: Yearly Missile Test Trends of India vs its Adversaries](image)

Ultimately, I identify two possible reasons behind India’s missile test record in 2004. First, this peak may have been a response to a threat that transpired prior to 2004. Missile tests are often expensive to conduct. Hence, the decision to test-launch a missile may not be easily taken without prior careful fiscal planning. Accordingly, missile launches that we observed in a certain year may be based on fiscal decisions taken the previous year. In that regard, India’s peak performance in 2004 may be related to level of threat it perceived in 2003 or even before.

The second possibility could be that countries do not always respond to military threats by test-firing missiles. In the cases of Israel and England, a challenge was to attribute the missile testing trend peaks with periods of pressing foreign military threats.

\textsuperscript{92} In fact, more recently, the US president Donald Trump held two summits with North Korea’s leader Kim Jong-Un in 2019. Since these summits, news of North Korea test-firing missiles has been very uncommon.

\textsuperscript{93} This lack of association is presented in Figure 17 by the cross-correlation plot based on China and India’s missile tests from 1985 and on.
Israel is a country that has never enjoyed a year of complete peace among hostile neighbors. However, not only are its missile test numbers the lowest in the sample, but they also do not seem to show a trend that is associable with a particular conflict. Similarly, while the United Kingdom seems to have test-fired a great many missiles around the Falklands War, a similar peak is observed in the late 1980s long after the war is concluded. Hence, the association between missile tests and perception of military threats may be limited or conditional on factors that are, as of yet, unknown.

All in all, it appears that missile tests are affected by the military threats that countries are exposed to. However, a cursory glance at the time series shows that this “response-to-threat” argument has limitations. Naturally, the next step is to conduct serious statistical analyses. To this day, no such analysis has been conducted. Thanks to the dataset I presented in this paper, such analyses are now possible. The results of subsequent detailed statistical analyses will contribute to our existing understanding of missile tests and supplement these individual country reviews.

References for Paper I:
Aerospace America (1989) “China holds course for space program” Published: September, 1989. Web: Nexis Uni


Calgary Herald (1997) “Iran missile test launch puts Israel within range” Published: November 2, 1997. Web: Nexis-Uni


Encyclopedia Astraunomica – Arrow. Web: http://www.astronautix.com/a/arrow.html


Encyclopedia Astronomica – Germany. Web: http://www.astronautix.com/g/germany.html
Encyclopedia Astronomica – KT-1. Web: http://www.astronautix.com/k/kt-1.html
Encyclopedia Astronomica – Minuteman 1A. Web: http://www.astronautix.com/m/minuteman1a.html
Encyclopedia Astronomica – MSBS. Web: http://www.astronautix.com/m/msbs.html


The NYTimes (1991). **Published:** 1/30/1991 **Web:** Nexis Uni


The NYTimes (1983) Published: June 18, 1983. **Web:** Nexis-Uni


Press Association (1991) “Britain to Buy First Trident Missiles” **Published:** November 21, 1991. **Web:** Nexis Uni

Progressive Media – Company News (2012)


Schelling, Thomas. C. (1960) *The Strategy of Conflict*

Schelling, T. C. (1966) *Arms and Influence* Yale University Press


State.Gov “Treaty Between the United States of America and The Union of Soviet Socialist Republics on the Limitation of Strategic Offensive Arms (SALT II)”. **Web:** [https://www.state.gov/t/ism/5195.htm#treaty](https://www.state.gov/t/ism/5195.htm#treaty)


APPENDIX for Paper I

(1) Missile Aspects by Countries (1949-2015)

Table 1: Aspects of the Missiles Tested by Country

<table>
<thead>
<tr>
<th>Missiles for space</th>
<th>The US</th>
<th>The UK</th>
<th>France</th>
<th>USSR/Russia</th>
<th>Iran</th>
<th>Iraq</th>
<th>Israel</th>
<th>China</th>
<th>North Korea</th>
<th>India</th>
<th>Pakistan</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICBM</td>
<td>1,554</td>
<td>9</td>
<td>-</td>
<td>1067</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>27</td>
<td>-</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>ICCM</td>
<td>127</td>
<td>4</td>
<td>10</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IRBM</td>
<td>346</td>
<td>6</td>
<td>22</td>
<td>565</td>
<td>32</td>
<td>9</td>
<td>3</td>
<td>54</td>
<td>12</td>
<td>11</td>
<td>-</td>
</tr>
<tr>
<td>IRCM</td>
<td>453</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Interceptors</td>
<td>360</td>
<td>7</td>
<td>1</td>
<td>74</td>
<td>17</td>
<td>-</td>
<td>29</td>
<td>6</td>
<td>-</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>MRBM</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>11</td>
<td>-</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>MRCM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nuclear tests</td>
<td>910</td>
<td>45</td>
<td>191</td>
<td>607</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>45</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Missile for space</td>
<td>463</td>
<td>6</td>
<td>9</td>
<td>1442</td>
<td>14</td>
<td>-</td>
<td>6</td>
<td>35</td>
<td>4</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>SRBM</td>
<td>711</td>
<td>-</td>
<td>15</td>
<td>156</td>
<td>48</td>
<td>89</td>
<td>17</td>
<td>5</td>
<td>42</td>
<td>89</td>
<td>49</td>
</tr>
<tr>
<td>SRCM</td>
<td>18</td>
<td>2</td>
<td>-</td>
<td>8</td>
<td>22</td>
<td>-</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>45</td>
<td>15</td>
</tr>
<tr>
<td>Other short range</td>
<td>196</td>
<td>22</td>
<td>24</td>
<td>-</td>
<td>19</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>-</td>
<td>118</td>
<td>8</td>
</tr>
<tr>
<td>Other naval missiles</td>
<td>465</td>
<td>48</td>
<td>32</td>
<td>125</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>11</td>
<td>3</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>Unidentified aspect</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>29</td>
<td>4</td>
<td>-</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Total observations</td>
<td>5,063</td>
<td>157</td>
<td>306</td>
<td>4,085</td>
<td>171</td>
<td>98</td>
<td>68</td>
<td>200</td>
<td>79</td>
<td>316</td>
<td>104</td>
</tr>
</tbody>
</table>
(2) Missile Names by Country

The following is the list of missiles tested by each country according to my data collection. Note that the list does not contain all the missiles that the countries possess, but the ones that were test-fired publicly and thus were recorded in the dataset. The aspects of the missiles are presented in parentheses, and the missile names are presented in italics. If the aspect is “short range”, this includes air-to-air, ship-to-air, or anti-tank missiles that are neither ballistic nor cruise. The French Exocet is in this category. The “military space” aspect designates rockets of ICBM family, or rockets that use boosters from ICBMs or IRBMs that are launched to space for military missions. The “ICBM space” aspect represents missiles that are used as rockets as well as ICBMs. The “interceptor” aspect involves SAMs (surface-to-air missiles), and missile interceptors in general. Thus, if a missile is a SAM, it is not in the “short range”, but in the “interceptor” category.

The US

**Delta 7420-10C** (military space), **Delta 790-10C** (military space), **Delta 7920-10** (military space), **Delta 7920-10C** (military space), **Delta 7925** (military space), **ERCS Blue Scout Jr** (military space), **ERIS** (interceptor), **Excalibur Target System** (interceptor), **GBI** (Interceptor), **GBI BV-Plus** (interceptor), **GMD** (interceptor), **GMI** (Interceptor), **Goose** (ICCM), **HEDI** (interceptor), **HIBEX** (interceptor), **HOE** (Interceptor), **HVAR** (short range), **Harpoon** (naval ballistic), **Hera** (ICBM), **Hermes A-2** (SRBM), **Hermes A-3A** (SRBM) **Hermes A-3B** (SRBM), **Hound Dog** (IRCM), **Jupiter** (IRBM), **Lance** (SRBM), **Lark** (short range), **MGR-I** (SRBM), **MX** (ICBM), **MX-775** (ICCM), **Mace** (ICRM), **Matador** (IRCM), **Midgetman** (ICBM), **Minotaur IV Lite** (military space), **Minuteman IA** (ICBM), **Minuteman IB** (ICBM), **Minuteman II** (ICBM), **Minuteman III** (ICBM), **Navaho** (IRCM), **Navaho G-26** (IRCM), **Navaho X-10** (IRCM), **Nike Hercules** (Interceptor), **Nike Zeus A** (Interceptor), **PAC-2** (Interceptor), **PAC-3/Patriot** (Interceptor), **Pegasus** (military space), **Pegasus H** (military space), **Pegasus XL** (military space), **Pershing 1A** (IRBM), **Pershing 2** (IRBM), **Polaris A1** (naval ballistic), **Polaris A2** (naval ballistic), **Polaris A2E** (naval ballistic), **Polaris A3** (naval ballistic), **Poseidon C3** (naval ballistic), **Quail** (short range), **R-11 / Scud B** (SRBM), **Redstone** (SRBM), **SM-2-IV** (interceptor), **SM-3-IB** (interceptor), **Sergeant** (SRBM), **Skybolt** (SRBM), **Snark** (ICCM), **Spartan** (Interceptor), **Sprint** (Interceptor), **THAAD** (Interceptor), **Talos** (short range), **Taurus 1110** (military space), **Thor** (IRBM), **Thor Ablestar** (IRBM, space), **Thor Agena A** (IRBM, space), **Thor Agena B** (IRBM, space), **Thor Agena D** (IRBM, space), **Thor DM-18A** (IRBM, space), **Thor Delta M** (IRBM, space), **Thorad Agena D** (IRBM, space), **Titan 23B** (ICBM, space), **Titan 24B** (ICBM, space), **Titan 34B** (ICBM, space), **Titan 34D** (ICBM, space), **Titan 34D/Transtage** (ICBM, space). **Titan 401A/Centaur** (ICBM, space), **Titan 401B/Centaur** (ICBM, space), **Titan 402A/IUS** (ICBM, space), **Titan 402B/IUS** (ICBM, space), **Titan 403A** (ICBM, space), **Titan 403B** (ICBM, space), **Titan 404A** (ICBM, space), **Titan 404B** (ICBM, space), **Titan 405A** (ICBM, space), **Titan 405B** (ICBM, space), **Titan I** (ICBM), **Titan II** (ICBM), **Titan IIIC** (ICBM), **Titan III** (ICBM), **Titan IIIB** (ICBM), **Titan IIIIC** (ICBM), **Titan IIIID** (ICBM), **Tomahawk** (IRCM), **Tomahawk Airframe** (IRCM), **Tomahawk Anti-Ship** (IRCM), **Tomahawk Land Attack** (IRCM), **Tomahawk Block IV** (IRCM), **Trident** (naval ballistic), **Trident C-4** (naval ballistic), **Trident II D5** (naval ballistic), **V-2** (SRBM)

**The UK**

**Blue Streak** (IRBM), **Blue Water** (short range), **Brimstone** (short range), **Delta 7925-9.5** (military space), **Harpoon** (naval ballistic), **Hellfire** (short range), **M270B1** (military space), **MBDA ASRAAM** (short range), **Meteor** (short range), **PAAMS/Sea Viper** (Interceptor), **PGM-3** (short range), **Paveway IV** (short range), **Polaris** (naval ballistic), **Polaris A3** (naval ballistic), **SA-6, Sea Dart** (short range), **Delta 7920** (military space), **Delta 7920-10C** (military space), **Delta 7925** (military space), **ERCS Blue Scout Jr** (military space), **ERIS** (interceptor), **Excalibur Target System** (interceptor), **GBI** (Interceptor), **GBI BV-Plus** (interceptor), **GMD** (interceptor), **GMI** (Interceptor), **Goose** (ICCM), **HEDI** (interceptor), **HIBEX** (interceptor), **HOE** (Interceptor), **HVAR** (short range), **Harpoon** (naval ballistic), **Hera** (ICBM), **Hermes A-2** (SRBM), **Hermes A-3A** (SRBM) **Hermes A-3B** (SRBM), **Hound Dog** (IRCM), **Jupiter** (IRBM), **Lance** (SRBM), **Lark** (short range), **MGR-I** (SRBM), **MX** (ICBM), **MX-775** (ICCM), **Mace** (ICRM), **Matador** (IRCM), **Midgetman** (ICBM), **Minotaur IV Lite** (military space), **Minuteman IA** (ICBM), **Minuteman IB** (ICBM), **Minuteman II** (ICBM), **Minuteman III** (ICBM), **Navaho** (IRCM), **Navaho G-26** (IRCM), **Navaho X-10** (IRCM), **Nike Hercules** (Interceptor), **Nike Zeus A** (Interceptor), **PAC-2** (Interceptor), **PAC-3/Patriot** (Interceptor), **Pegasus** (military space), **Pegasus H** (military space), **Pegasus XL** (military space), **Pershing 1A** (IRBM), **Pershing 2** (IRBM), **Polaris A1** (naval ballistic), **Polaris A2** (naval ballistic), **Polaris A2E** (naval ballistic), **Polaris A3** (naval ballistic), **Poseidon C3** (naval ballistic), **Quail** (short range), **R-11 / Scud B** (SRBM), **Redstone** (SRBM), **SM-2-IV** (interceptor), **SM-3-IB** (interceptor), **Sergeant** (SRBM), **Skybolt** (SRBM), **Snark** (ICCM), **Spartan** (Interceptor), **Sprint** (Interceptor), **THAAD** (Interceptor), **Talos** (short range), **Taurus 1110** (military space), **Thor** (IRBM), **Thor Ablestar** (IRBM, space), **Thor Agena A** (IRBM, space), **Thor Agena B** (IRBM, space), **Thor Agena D** (IRBM, space), **Thor DM-18A** (IRBM, space), **Thor Delta M** (IRBM, space), **Thorad Agena D** (IRBM, space), **Titan 23B** (ICBM, space), **Titan 24B** (ICBM, space), **Titan 34B** (ICBM, space), **Titan 34D** (ICBM, space), **Titan 34D/Transtage** (ICBM, space). **Titan 401A/Centaur** (ICBM, space), **Titan 401B/Centaur** (ICBM, space), **Titan 402A/IUS** (ICBM, space), **Titan 402B/IUS** (ICBM, space), **Titan 403A** (ICBM, space), **Titan 403B** (ICBM, space), **Titan 404A** (ICBM, space), **Titan 404B** (ICBM, space), **Titan 405A** (ICBM, space), **Titan 405B** (ICBM, space), **Titan I** (ICBM), **Titan II** (ICBM), **Titan IIIC** (ICBM), **Titan III** (ICBM), **Titan IIIB** (ICBM), **Titan IIIIC** (ICBM), **Titan IIIID** (ICBM), **Tomahawk** (IRCM), **Tomahawk Airframe** (IRCM), **Tomahawk Anti-Ship** (IRCM), **Tomahawk Land Attack** (IRCM), **Tomahawk Block IV** (IRCM), **Trident** (naval ballistic), **Trident C-4** (naval ballistic), **Trident II D5** (naval ballistic), **V-2** (SRBM)
Sea Eagle (short range), Sea Skua (short range), Sea Wolf (short range), Storm Shadow (ICCM), Thor Delta M (IRBM, space), Tomahawk Land Attack (IRCM), Tomahawk Block IV (IRCM), Trident 2 (naval ballistic), Trident II D5 (naval ballistic)

France

ASMP-A (IRCM), Ariane 5ECA (military space), Ariane 40 (military space), Ariane 5GS (military space), Ariane 5Gp (military space), Ariane 5GS (military space), Aster (interceptor), SM-39 Exocet (short range), Exocet MM40 (short range), Hades (SRBM), Jericho (SRBM), MSBS M012 (IRBM), MSBS M013 (IRBM), MSBS M112 (IRBM), MSBS M20 (IRBM), MSBS M21 (IRBM), MSBS M4 (SRBM), MSBS M45 (IRBM), MSBS M51 (ICBM), MSBS SSBS S01 (IRBM), MSBS SSBS S02 (naval ballistic), MSBS SSBS S03 (naval ballistic), MSBS SSBS S112 (naval ballistic), SE4400 (short range), Storm Shadow / SCALP EG (ICCM)

The USSR/Russia

51T6 (Interceptor), 53T6 (interceptor), A-350R (interceptor), A-350Zh (Interceptor), Bulava (naval ballistic), Energia (military space), Grad, Iskander-E (SRBM), Iskander-M (SRBM), K65M-R (military space), Kalibr (naval cruise), Kh-35 (military space), Kosmos 11K63 (military space), Kosmos 11K65 (military space), Kosmos 11K65M (military space), Kosmos 63S1 (military space), Kosmos 63S3 (military space), Kosmos 65S3 (military space), Kosmos K 1163 (military space), Kosmos K11K63 (military space), Kosmos11K65M (military space), Kuryer (ICBM), Layner (naval ballistic), MR-UR-100 (ICBM), MR-UR-100 15A16 (ICBM), MR-UR-100U 15A16 (ICBM), Molniya 8K78 (ICBM, space), Molniya 8K78M (ICBM, space), Molniya 8K78SP (ICBM, space), Moskit (naval cruise), Nudol (military space), Oka (IRBM), Pantsir (SRBM), Proton-K/17S40 (ICBM, space), Proton-K/DM (ICBM, space), Proton-K/DM-2 (ICBM, space), Proton-M/Briz-M (ICBM, space), Proton-U-PVB (ICBM, space), Proton/Briz K/M (ICBM, space), R-1 (SRBM), R-2 (SRBM), R-2A (SRBM), R-21 (SRBM), R-11 (naval ballistic), R-11FM (naval ballistic), R-11M (naval ballistic), R-12 (IRBM), R-12U (IRBM), R-13 (naval ballistic), R-14 (IRBM), R-14U (IRBM), R-16 (ICBM), R-16U (ICBM), R-17/ Scud (SRBM), R-IV (SRBM), R-2 (IRBM), R-21(naval ballistic), R-2A (IRBM), R-2R (IRBM), R-31 (naval ballistic), R-36 8K67 (ICBM), R-36 8K67M (ICBM), R-36 8K67MA(ICBM), R-36 8K67P (ICBM), R-36 8K67PM (ICBM), R-36 M2 15A18M (ICBM), R-360 8K69 (ICBM), R-36M 15A14 (ICBM), R-36M2 (ICBM), R-36M2 15A18M (ICBM), R-36MU 15A18 (ICBM), R-5A (IRBM), R-5M (IRBM), R-7 (ICBM), R-7A (ICBM), R-9A (ICBM), R-12e (SRBM), RT-1 (MRBM), RT-15 (IRBM), RT-2 (ICBM), RT-20P (ICBM), RT-23 15Zh44 (ICBM), RT-23 15Zh52 (ICBM), RT-23U 15Zh60 (ICBM), RT-23U
15Zh61 (ICBM), RT-2P (ICBM), Rif (Interceptor), Rokot (military space), Rubezh (ICBM), S-125 (Interceptor), S-225 (Interceptor), S-300 Favorit (Interceptor), S-400 (Interceptor), SS-18 (ICBM), SS-20 (IRBM), SS-21 Scarab (IRBM), SS-21 Scarab A (IRBM), SS-21 Scarab B (IRBM), SS-N-18 Volna (naval ballistic), SSC-8 (IRCM), SSN-20 (Interceptor), Shamil (naval ballistic), Sineva (naval ballistic), Soyuz 11A510 (ICBM, space), Soyuz 11A511M (ICBM, space), Soyuz-2-IA (ICBM, space), Soyuz-2-JB (ICBM, space), Soyuz-U (ICBM, space), Soyuz-U-PVB (ICBM, space), Soyuz-U2 (ICBM, space), Sputnik 11A59 (military space), TORMI (Interceptor), Temp-2S (ICBM), Tochka (IRBM), Topol’ (naval ballistic), Topol’-M (naval ballistic), Topol’-M 15Zh55 (naval ballistic), Tsiklon-2 (ICBM space), Tsiklon-2A (ICBM space), Tsiklon-3 (ICBM space), UR-100 (ICBM), UR-100K (ICBM), UR-100M (ICBM), UR-100N (ICBM), UR-100NU 15A35 (ICBM), UR-100NU 15A35P (ICBM), UR-100NU 15A35S (ICBM), UR-200 (ICBM), UR-300 Proton (ICBM), V-1000 (Interceptor), Voskhod 11A57 (military space), Voskhod 8A92 (military space), Voskhod 11A57 (military space), Voskhod 65S3 (military space), Voskhod 8A92M (military space), Voskhod 11A57 (military space), Vostok 11A57 (military space), Vostok 8A92 (military space), Vostok 8A92M (military space), Vostok 8K72K (military space), Vysota (naval ballistic), Yars (ICBM), Yars 15Zh65M (ICBM), Yu-71 (military space), Zenit-2 (military space)

Iran

AGM-65 Maverick (interceptor), Ajdar (SRBM), Arrow (interceptor), Bina (short range), C-701 (SRCM), C-802 (SRCM), Dehlaviyeh (short range), Emad (IRBM), Fadjr-3 (short range), Fadjr-5 (short range), Fakour (short range), Fateh-110 (SRBM), Fateh-313 (SRBM), Ghadr-1 (IRBM), Hag ( interceptor), Kavoshgar-1 (military space), Kavoshgar-3 (military space), Kavoshgar-2 (military space), Kavoshgar-5 (military space), Kavoshgar-6 (military space), Kavoshgar-K110 (military space), Kowsar (SRCM), Mersad (interceptor), Misag-1 (SRBM), Nasr (SRCM), Nazeat-10 (short range), Noor (SRCM), Qader (SRCM), Qiam-1 (SRBM), Ra’ad (naval cruise), S-200 (Interceptor), S-300 (Interceptor), SA-6 (short range), Safir (military space), Safir-2 (military space), Sayyad-1 (short range), Sayyad-2 (short range), Sejjil (MRBM), Shahab-1 (SRBM), Shahab-2 (SRBM), Shahab-3 (IRBM), Shahab-4 (IRBM), Shahin/Mershad (interceptor), Shalamcheh/Mersad (interceptor), TOW (short range), Talaash 3 (interceptor), Tondar 69 (SRBM), Toopan (short range), Tor-M1 (interceptor), Zelzal (short range)
Iraq


Israel


China


North Korea


India

(SRBM), R-73 (short range), Sagarika (naval ballistic), Shourya (naval ballistic), Trishul (short range)

Pakistan

Ariane 5ECA (military space), Baktar-Shikan (short range), Exocet AM-39 (short range), Exocet SM-39 (short range), FM 90 (interceptor), Ghauri II (MRBM), Ghauri III (MRBM), H-4 (short range), Hatf I (SRBM), Hatf 1B (SRBM), Hatf II Abdali (SRBM), Hatf III Ghaznavi (SRBM), Hatf IV Shaheen-I (SRBM), Hatf IV Shaheen-1A (SRBM), Hatf IX Nasr (SRBM), Hatf V Ghauri (MRBM), Hatf VI Shaheen-2 (MRBM), Hatf VII Babur (SRCM), Hatf VIII Ra'ad (SRCM), Scud (SRBM), Shaheen-3 (MRBM), Zenit-2 (military space)

(3) Cross-Correlation Plots

Figure 15: US vs USSR/Russia

Cross-Correlation between the US and USSR/Russian Public Missile Tests
Figure 16: India vs Pakistan

Cross-Correlation between the India and Pakistan Public Missile Tests

Figure 17: India vs China

Cross-Correlation between the India and China's Public Missile Tests
Paper II: Why do countries test missiles and missile-related technologies?

When North Korea drastically increased its missile tests in late 2016, many policy experts questioned what motivated such changes. Was it North Korea’s desire to intimidate the US and South Korea who were often conducting joint military maneuvers at the time? Was it simply due to the unpredictable nature of the North Korean leader Kim Jong-Un? Was it because the US was deploying THAAD missiles to South Korea? Alternatively, was it because North Korea was purposefully defying the UN sanctions? While many potential explanations were set forth, none appeared uniquely capable of fully accounting for the drastic increase in North Korean missile tests.

The North Korean case is not unique. Countries often test missiles or missile-related technologies publicly. At certain times, these tests are conducted at an alarmingly high frequency. Chinese missile tests were very frequent in the 1990s around the time the Taiwan Strait Crisis took place. Shortly after the conclusion of the crisis, Chinese missile tests became less frequent. However, since approximately 2010, China has been test-firing more missiles some of which focus specifically on space-based strike capabilities.94 Similarly, Iran’s missile tests grew considerably during President Ahmadinejad’s tenure when the country’s relations with the US were particularly tense. Finally, India and Pakistan frequently test-fire missiles in a tit-for-tat fashion since Pakistan started its own missile development program in early 1990s. In fact, there have been numerous cases when

---

94 We know that China tested a supersonic space glider at least twice. This glider, called Wu-14 is technically capable of mounting nuclear strikes from space (ValueWalk 2015). In addition, China, in 2009 test-fired a medium range missile that destroyed an old satellite which was in a higher orbit that most American surveillance satellites (Facts on File 2007).
India tested a missiles and Pakistan responded in kind with a test of its own. The question is what factors account for these changing trends in publicly conducted missile tests.

In this paper, I develop and empirically tests several hypotheses that shed light on why countries sometimes test-fire more missiles than other times. A survey of the extend literature reveal three general – yet empirically untested – ideas that can account for variance in countries’ missile testing. First, increase in missile tests may be attributed to the introduction of new weapon development programs. Second, countries may increase their public missile tests as a reaction to recent or ongoing military threats. Finally, countries may test-fire more missiles to counter economic duress created by sanctions. Countries might pursue this strategy to market their already existing arsenal to potential buyers and generate additional income by exporting missiles.

To test these three different expectations, I first present a dataset on missile tests conducted by North Korea, Iran, India, Pakistan, China, Israel, Iraq, the USSR/Russia, the US, the UK and France between 1949 and 2010. Second, I conduct a series of negative binomial regressions to evaluate the three sets of arguments. These statistical analyses seem to place greater weight on the first and the second explanations. Specifically, countries appear to test-fire missiles more often when they are faced with military threats and when they are developing new missiles that require further testing before commissioning for service. However, I fail to find any evidence to support the economic sanctions argument. It seems that economic sanctions have no positive or negative impact on missile tests, even though many policy-makers seem to believe that this policy tool can be used to reduce missile tests by rogue countries.
The paper’s contribution to our understanding of international security is two-folds. First, it presents the first systematic empirical work that addresses the motivations behind missile tests. Although the policy field is full of qualitative reports on individual countries such as North Korea, no quantitative analysis has been conducted to assess competing (or complementary) claims about what makes countries test-fire more or fewer missiles. Second, the paper presents clear policy implications regarding what measures can be taken to reduce a target country’s missile tests. Notably, sanctions remain ineffective according to the present findings. However, avoiding dragging countries into militarized interstate disputes may be a potential way to decrease their incentives to test-fire missiles. Further analyses, however, show that if one wages a war against the country of interest, this has a consistent impact on reducing this country’s tests of new missiles – which, in the long term, also significantly reduces its overall missile tests.

In what follows, I first develop three sets of empirical expectations regarding how frequently countries test missiles. Second, I present the research design to statistically assess these empirical expectations. Third, I summarize the statistical results and discuss the implications. Finally, I conclude the paper with a discussion of the policy implications of the findings.

**What do we know of missile tests so far?**

The set of actions that the present study focuses on are test-launching missiles or missile-interceptors. Most of these tests are conducted publicly. The test-launched weapons range from ground, ship or air-based ballistic missiles to cruise missiles and missile interceptors. None of these launches take place during wars. In other words, when we
observe countries test-launching their missiles, we have to assume that they are not directly attacking their adversaries. What we actually observe are rehearsals of actual missile strike or missile defense missions outside the context of war.

While they do not take place in the context of war, missile tests are far from trivial or innocent incidents in international relations. Almost every missile test-launch by North Korea is criticized by the US as well as South Korea, Japan and even China and Russia. When Iran test-fires cruise missiles during naval exercises in the Strait of Hormuz, US Navy forces in the region and Israel are among the first to criticize. These cases suggest that that policy-makers do not see peaceful intentions behind missile tests. In fact, during most of 1970s, the US and the USSR sought to monitor and reduce each other’s missile tests though the SALT I and SALT II meetings. Similarly, since 2006, the UN has repeatedly imposed sanctions on North Korea partly to prevent the latter from test-firing missiles.

Thus far, the field of international relations has largely overlooked what explains the changing frequency with which countries conduct public missile tests. One reason for this lack of attention could be that we already are equipped with the appropriate theoretical tools to understand these tests. For example, a frequent argument in policy-focused articles or news pieces is that countries try to intimidate their adversaries with their missile tests (see Schilling and Kan 2015: 9). Intimidation is a well-studied topic in international relations (Hutt and Russett 1984, Jervis 1982, Schelling 1960, Signorino and Tarar 2006). Hence, scholars of international politics may believe that we already have a theoretical framework to understand why countries test-fire missiles and that we do not need to think of alternative theories. On the other hand, it could be that missile tests are a necessary
component of developing new missiles. Thus, they may occur independently from particular incidents that threaten the missile-testing countries. How can we compare these two arguments?

Even though we may be equipped with several arguments to understand why countries test-fire missiles, we cannot know whether these arguments are valid without testing them against data. Unfortunately, as I highlighted earlier, neither a systematic data collection effort nor any empirical test has been conducted to understand the factors that explain missile tests. The only exception is CSIS’ Missile Defense Project which collected chronological data on Iranian and North Korean missile tests (Missile Defense Project – Iran, Missile Defense Project – North Korea). Nevertheless, this data has never been used to conduct statistical analyses.

Accordingly, two task lies ahead if we seek to improve our understanding of the frequency of missile tests. The first task is to derive what arguments can be used to generate hypotheses to help us understand what factors typically make countries increase or decrease their publicly-conducted missile tests at certain times. The second task is to empirically assess these hypotheses. In what follows, I present three main sets of arguments from which I derive several testable hypotheses. Finally, I detail an empirical research design suitable for evaluating these hypotheses.

**Argument 1 – Missile Tests as a Necessary Part of Developing New Missiles.**

Countries spend millions or even billions of dollars to develop missiles. For example, the initial cost for the US government to develop the MX (Peacekeeper) ICBM
was around $11 Billion. Only a few years after the initiation of the MX program, the costs increased to $20 Billion (Encyclopedia Astraunomica - Peacekeeper). The development cost of the French MSBS M51 exceeded $4 Billion (Global Security – M5/M51). Naturally, neither politicians nor voters would want such large sums to be wasted. What if the M51 malfunctioned during an actual combat situation and hit a French naval vessel? What if the same class of missile exploded in a submarine launch pad? What if an MX, a nuclear capable ICBM that can carry multiple warheads, exploded in the launch pad during wartime and caused great damage at Vandenberg Air Force Base?95

In order not to avoid wasting billions of dollars on malfunctioning missiles, countries often spend millions of additional dollars testing the new missiles they develop. In fact, we often observe years between the day a missile program is approved for development and the day the missile is deemed operationally capable. By test-firing these missiles, countries get feedback about a missile’s accuracy, range, navigational capabilities, and the stability of its engines. Based on successful or failed tests, countries re-consider what to adjust in the schematics of the missiles and how long it will take to commission the missile. Tests (especially failed ones) can also reveal that a missile should never be introduced to service. Overall, test-firing a missile is almost a necessary component before commissioning it.

If missile tests are a necessary part of developing a missile and making it operationally capable, we should empirically observe that a country increases its missile tests shortly after it initiates new missile-development programs. Unfortunately, we do not always know precisely when a missile program is introduced, since many countries such

---

95 Vandenberg is one of the American busiest Air Force Bases where many ICBM tests have been conducted since 1960s (Encyclopedia Astraunomica – Vandenberg)
as Russia and North Korea seek to keep such information secret. However, if we see a missile being test-fired for the first time, this could mean that the missile is in its early development stages. Thus, more tests of this missile may be conducted soon. In brief, an increase in a country’s missile tests at a certain time period may be positively associated with how many new missiles were developed or tested for the first time recently.

To introduce a formal hypothesis derived from this first argument, we need to be more specific about the time interval when the missile tests take place. For the present study, I use one year as the appropriate time interval. For most countries, a fiscal period is approximately one year.\textsuperscript{96} Thus, missile tests in a given year may reflect the decisions taken at the end of the previous (fiscal) year. In the present case, this first argument is supported if an increase in missile tests in a year is positively associated with the number of new missiles test-fired for the first time the year before. Accordingly, the first hypothesis can be stated as:

**Hypothesis 1:** In a given year, a country should test fire more missiles the more missiles were tested for the first time the previous year

**Argument 2 – Missile Tests to Respond to Military Threats**

Countries employ different foreign policy tools to respond to threats from their adversaries. Some of the foreign policies designed to counter foreign threats are the initiation of preventive wars (Fearon 1995, Powell 1999, 2006, Waltz 1979), forming defensive alliances (Smith 1995, Vasquez 1987, Waltz 1979), increasing military

\textsuperscript{96} In fact, one of the most expensive missile tests in Soviet Russia were part of 5-year spending plans (Encyclopedia Astraunomica – Russia)
expenditures (Morgan and Palmer 1997), conducting joint military exercises (Johnson et al 2015), initiating nuclear alertness (Schelling 1960, 1966), displaying force through actions such as jet passes and naval vessel maneuvers. Some of these foreign policy tools such as wars serve as means to wipe out an existing threat. On the other hand, other tools such as alliance formation, nuclear alertness or shows of force seek to warn adversaries that escalating an ongoing conflict will have lethal consequences.

Several incidents suggest that increases in a country’s missile tests may be related to its drive to intimidate hostile foreign powers that pose a military threat. In other words, one possibility is that countries sometimes test-fire missiles to warn their adversaries about the dangers of future hostile military encounters. For example, shortly after the US deployed 12 ships and 20,000 troops to the Persian Gulf, in 1996 Iran conducted a series of cruise missile tests in the Strait of Hormuz (USA Today 1996). These tests would one of many that coincided with the increase in American military activity in the Persian Gulf (see BBC 1997). The Russian Federation’s submarine-launched Topol-M tests in 2007 were interpreted as a direct response to the US plan to deploy anti-missile systems to Poland and other Eastern European countries (API 2007). A general look at Iraq under Saddam’s regime shows that the majority of its missile test-fires took place shortly before the first and second Gulf Wars (Encyclopedia Astraunomica – Iraq). If these incidents are part of a general pattern, then one factor that contributes to periodic increases in a country’s missile tests is its perception of foreign military threats.

The question is whether we can empirically identify periods of foreign threat and whether these periods correspond to increases in a country’s missile tests. In the present
study, based on multiple missile test narratives\textsuperscript{97}, I focus on three broad types of events which may be interpreted by countries as foreign military threats: militarized interstate disputes (MIDs), adversaries’ missile tests, and joint military exercises. If countries test-fire missiles in order to intimidate adversaries that pose a threat, these three broad types of events should lead to an increase in a country’s missile tests. In what follows, I detail how these three events can represent foreign military threats to a country.

The first type of foreign threat is a militarized interstate dispute. A MID is a phase of interstate conflict that involves at least one of the participants issuing a threat to use military force, displaying force, or using military force short of war. Display of force involves shows of force, maneuvers, troop and/or ship deployments which are often seen as producing substantial levels of threat\textsuperscript{98}. Use of military force short of war involve causing harm against an opponent which may lead to physical and even humanitarian losses.\textsuperscript{99} It is also possible that these actions can end up in wars (Jones et al 1996, Palmer et al 2015). Thus, if a country was involved in a MID recently, this means that an adversary recently issued a direct threat to use force, made a show of force, or actually used force (which might even have ended up in a war). In brief, MID is a level of conflict among nations where the use of military is now on the table in addition to (or instead of) diplomacy.

After a country experiences a MID, chances are high that the underlying political issue is not completely resolved. Hence, more MIDs around the same political issue may

\textsuperscript{97} The first paper of the dissertation presents a collection effort on missile tests. Narratives of missile tests were obtained/created based on this data collection project. The paper also concludes with a discussion of how countries’ increase in missile tests appear to periodically coincide with some of its military crises with others.

\textsuperscript{98} For a detailed analysis on how troop deployment can constitute a military threat, see Slantchev (2011)

\textsuperscript{99} The most frequent type of military action observed in MIDs is use of force short of war (Palmer et al 2015)
follow in the future. Several scholarly works show that the probability of two countries facing each other in a MID increases if they already had previous MIDs. In other words, MIDs often beget future MIDs. (Colaresi et al 2008, Klein et al 2006, Maoz and Mor 1996). Thus, a country’s recent MID involvement generally (if not always) leads to subsequent MID involvements around the same political issue. As a result, we can see that the associated foreign threat persists even after an initial dispute has ended. From this, we can speculate that one potential reason behind an increase in a country’s missile tests may very well be its recent involvement in MIDs.

More specifically, if a country experiences a MID in a given year, its decision-makers may choose to allocate additional funds to test more missiles for the next fiscal year. For example, following a military encounter in Kashmir with Pakistan in year t, the Indian legislators may decide to test-fire more missiles to intimidate Pakistan from continuing aggression. Their decision may be to increase funds for missile tests in year t+1. Hence, the impact of a military threat in a given year may potentially be observed as an increase in missile tests the next year. Accordingly, the second hypothesis can be stated as:

**Hypothesis 2:** In a given year, a country should test-fire more missiles publicly the more MIDs it was involved in the previous year.

In addition to MIDs, joint military exercises (JMEs) conducted by adversaries can constitute a military threat to a country. One of the reasons why countries conduct JMEs is to improve coordination among their armed forces. For example, NATO often conducts JMEs to accommodate its new members and to facilitate military cooperation with existing members (Wallander 2000). The US conducts joint military exercises with distant allies such as Thailand and South Korea to not only maintain good relations, but also to bolster
inter-army cooperation. This coordination is designed to ultimately make these forces more effective in addressing threats in East and Southeast Asia (see PACOM 2019). Therefore, the more countries conduct JMEs together, the better they coordinate and, as a result, their union becomes a more serious threat than an uneasy and uncertain alliance (see also Johnson et al 2015).

Having adversaries that often engage in peacetime military coordination activities with others should theoretically present a threat to a country. Anecdotal evidence shows that countries rarely welcome, and often protest, when their adversaries conduct JMEs. Russia often states its opposition to the US and its allies conducting JMEs in the Black Sea, the Pacific, or any region too close to the Russian sphere of political influence (CNN 2018, NY Times 2008). North Korea also condemns almost every joint US-South Korea exercise that takes place close to the Korean Peninsula (DW 2019).

From these details, we can, therefore, expect that the more a country’s adversaries conduct JMEs together, the greater of a military threat these adversaries should constitute. Consequently, this country may sometimes choose to address threats from foreign JMEs by publicly test-firing missiles. As a result, incidences of JMEs by its adversaries should be associated with an increase in a country’s subsequent missile tests. Similar to the two previous hypotheses, there should be at least one year between the incidence of JMEs by adversaries and an increase in a country’s publicly-conducted missile tests. The reason, again, is that fiscal decisions that affect missile tests are often made on a yearly basis.\textsuperscript{100} The third hypothesis, then, can be stated as:

\textsuperscript{100} In fact, this could be the reason why d’Orazio (2013) finds almost no immediate North Korean reaction to any US-South Korea joint military exercises. It could be that it takes a fiscal year for North Korea to plan funding for missile tests that respond to these exercises.
Hypothesis 3: *In a given year, a country test-fires more missiles publicly the more joint military exercises its adversaries conducted the previous year.*

Last but not least, there are many instances which lead us to suspect that countries feel threatened by their adversaries’ own missile tests and respond in kind. For example, shortly after India detonated two nuclear devices in 1998 (Shakti-I and Shakti-II), Pakistan conducted a series of missile tests as well as a nuclear warhead detonation test. Several news sources interpreted Pakistan’s motives as a message to India that Pakistan should not be taken lightly (The AP International 1998). Indeed, there are many instances in which Pakistan and India appear to test-fire missiles consecutively. Similarly, we also see that the Soviets and the Americans during the Cold War sought to reciprocate each other’s test launches.\(^{101}\)

If countries see the missile tests of their adversaries as a foreign threat, then one way to address this threat may be to respond with missile tests of their own. Then, observing increases in adversaries’ missile test-launches, a country may decide for the next fiscal term – typically a year – to also increase its own missile tests. Ultimately, we may even end up observing a missile test race pattern among countries. Hence, the last hypothesis for the missile tests as a response to military threat can be stated as:

**Hypothesis 4:** *In a given year, a country should test-fire more missiles publicly the more missiles its adversaries test-fired in the previous year.*

---

\(^{101}\) See Paper I of this Dissertation
Argument 3 – Missile Tests as Means to Cope with Sanctions

As illustrated earlier, missile tests are expensive, and require serious budgeting before taking place. By intuition, countries with inadequate funds should not be able to test as many missiles as they need. Ultimately, if countries do not have funds to test-fire their missiles, they also should be less likely to develop and maintain most of their ballistic capabilities.

Economic sanctions to prevent North Korea or Iran from developing nuclear capable weapons were designed with this logic: deny funds and they will not be able to afford test-firing their weapons. After test-firing its first nuclear warhead in 2006, North Korea received a total of 3 major sanctions from the UN. These episodes started in 2006\textsuperscript{102}, 2009\textsuperscript{103}, 2013\textsuperscript{104}. The common goal of these sanctions was to prevent North Korea from developing nuclear capabilities and test-firing missiles that can potentially carry nuclear warheads. Similarly, Iran received UN sanctions that sought to curb its nuclear weapon development programs.

Despite this motivating logic, the case of North Korea demonstrated that sanctions to curb missile development programs may actually have a counter effect. Despite economic pressure from the UN sanctions, North Korea did not stop testing missiles. In fact, North Korea’s missile tests increased from 2011 on (Missile Defense Project – North Korea, Paper I). The question is why would North Korea keep conducting missile tests despite these severe economic pressures and US-backed UN sanctions? Moreover, is North

\begin{footnotes}
\item[103] See https://www.un.org/ga/search/view_doc.asp?symbol=S/RES/1874%282009%29
\item[104] See https://undocs.org/S/RES/2087(2013)
\end{footnotes}
Korea’s increasing tests following sanctions generalizable across countries that are under similar economic pressures? Are sanctions ineffective against missile tests?

Recently, several scholars investigated how countries that receive economic sanctions ultimately learn to adapt to the economic pressures from these sanctions (Naghavi and Pignataro 2015). One such strategy is to find new trade partners to overcome with the sanctions. Several studies show that sanctions against countries that can find new trade partners tend to be less effective than against countries that fail to find alternative trade partners (Peksen and Peterson 2016, Whang et al 2013).

Several instances suggest that countries can conduct missile tests as part of a plan to export missile technology. The United States test-fired many Trident missiles before selling them to the UK’s in the 1980s and 1990s (The Independent 1989, US Newswire 1994). There are also instances where the Soviets, Russians, Pakistanis, and Iranians test-fired missiles before selling to potential customers (see BBC 1998). Indeed, Iran often exports missiles to Syria and Hezbollah (Missile Defense Project – Syria, MT-Hezbollah). China is a potential missile (or missile schematics) exporter to Pakistan. Additionally, Russia openly exports to and develops missiles with India (Encyclopedia Astraunomica - BrahMos). Thus, if economic sanctions do not decrease missile tests, this could be because the target of these sanctions may be working to export some of its missiles in order to generate new income.

Overall, these instances suggest that economic sanctions may have a positive impact on the number of missile tests that the target country conducts. Intuitively, these

---

105 See the missile overview of Pakistan in Paper I as well as http://armstrade.sipri.org/armstrade/html/export_values.php
missiles are likely to have already been commissioned. This is because customers often may not be interested in missiles that are not tested enough to be declared as operationally-capable. As highlighted earlier in this section, a missile that is not commissioned yet has the risk of being discarded should its tests result in too many failures. This is why, for example, North Korea may have more trouble marketing a not-commissioned Taepedong-III than a decades-old Nodong-I. Although the former has much greater range and destructive capabilities, the latter is already operationally capable and has entered into service for North Korea (Missile Defense Project – Nodong I). In addition, North Korea has still been test-firing Nodong-I missiles despite the UN sanctions (Paper I).

Empirically, after countries receive economic sanctions, we may observe a small but significant increase in the number of commissioned missiles that they test-fire. These tests are not necessarily “tests” for the test-firing country itself, but rather a demonstration for the potential clients. Finally, if a country is under economic sanctions in a given year, the expected positive impact on tests of already-commissioned missiles should be observed in the subsequent fiscal term. Thus, we can state the final hypothesis as:

**Hypothesis 5:** In a given year, a country should test-fire more commissioned missiles publicly if it was in an ongoing sanction regime in the last year that if it were not.

**Research Design (Sample and Variables)**

(i) **Missile Tests:** To assess the hypotheses, I collected data on missile and missile-related tests by the United States, USSR, Russia, France, England, Iran, Iraq, Israel, China, North Korea, India and Pakistan between 1949 and 2015. These tests involve the publicly disclosed launchings of ballistic and cruise missiles of various ranges as well as missile
interceptors and orbital launch vehicles that use a military rocket. Therefore, all the tests that I collected are related to the projectile capabilities of the countries of interest. I chose these 11 countries for two reasons. First, these countries developed their own domestic missile technologies and exported them to their allies. Thus, each of these countries generally test-fired more missiles than almost any other country in the world. Second, missile tests by these countries often became a concern for their potential adversaries in the aftermath of Second World War (hence the year coverage of the sample). Consequently, news reports abound detailing these countries test-firing a missile and receiving demarche from various other countries.\textsuperscript{106}

To collect the relevant data, I relied on three main sources: CSIS' (Center for Strategic and International Studies) Missile Defense Project,\textsuperscript{107} the Lexis-Nexis News Database, and Mark Wade's online database "Encyclopedia Astronautica".\textsuperscript{108} Using these three sources, I recorded tests of missiles or missile-related systems that are publicly conducted by the countries of interest.\textsuperscript{109} Furthermore, I sought to ensure that all the recorded missile tests of the countries of interest were witnessed either physically or through radar technologies by the rest of the world.

Initially, I recorded the necessary information for missile tests such that each observation represented a weapon tested at a certain day by the country of interest. For each of these observations, I coded the type of the missile of interest\textsuperscript{110}, how many test-

\textsuperscript{106} For a more detailed explanation behind the countries of choice, refer to Paper I (Public Missile Tests – A New Dataset) of this dissertation
\textsuperscript{107} See https://missilethreat.csis.org/
\textsuperscript{108} See http://www.astronautix.com/
\textsuperscript{109} See Paper I of the dissertation for an in depth description of the coding procedure.
\textsuperscript{110} The weapon types are missile interceptors such as Russian S-300, American PAC-3/Patriot, Israeli Barak; Short Range Ballistic Missiles (SRBM) such as R-17/Scud, Jericho 1 or Iraq’s Al Hussein; Intermediate Range Ballistic Missiles (IRBM) such as Thor, Jupiter, Pershing or Russian R-12, Medium Range Ballistic
firings of the same weapon were made, whether the weapon was tested for the first time or whether the weapon was already commissioned for service during the time of testing, and whether the missile was nuclear capable. Therefore, if two different weapons were tested during the same day, this led to two different observations being coded in the original form of the dataset. On the other hand, if a certain missile was test-fired more than once during the same day, it is coded as one observation.

Due to the nature of the hypotheses, I collapsed the original data to a country-year form. Therefore, I obtained how many missiles a country tested in a year, how many of these were tested for the first time, how many were tested after being commissioned, and what types of weapons were tested (and how frequently each type was tested). Figures 18, 19 and 20 display the yearly weapon testing trends by countries between 1949 and 2015. The dashed-blue curve for each country designates the number of tested missiles that I was able to identify as already commissioned on the day of testing (which I use for the hypotheses pertaining to the impact of economic sanctions). Note that the y-axis (i.e. number of missile tests) for each country’s plot is differently scaled. The difference is based on the maximum number of missiles each country test-fired in a year.

Missiles (MRBM) such as Soviet RT-1 or Iran’s Shahab 3, Intercontinental Range Ballistic Missiles (ICBM) such as Soviet Voevod or American Minuteman missiles; military rockets (orbital launch vehicles, anti-satellite weapons and space gliders) such as American Atlas Agena, Chinese CZ-6, Wu-14, or Soviet Energia; nuclear warhead tests; Short Range Cruise Missiles (SRCM) such as Israel’s Gabriel, Intermediate Range Cruise Missiles (IRCM) such as US’ Navaho X-10, or Hound Dog; Medium Range Cruise Missiles (MRCM) and Intercontinental Range Cruise Missiles (IRCM) such as Tomahawk.

111 The reader can refer back to the previous section for why I test the hypotheses in a yearly sample.
Figure 18: Yearly Missile Tests (1)

Figure 19: Yearly Missile Tests (2)
Ultimately, I obtain a sample of 736 observations between 1949-2015. Each observation is a country year, and each country has 67 observations. As presented in Figure 1, there are two outcome variables. The first is the total number of tests conducted by a country of interest in a given year. The second – which is only used for sanction-related hypotheses – is the total number of tests of already-commissioned missile conducted by a country of interest in a given year. The distributions of the missile tests variables (all tests, and commissioned ones) are displayed in Figure 7 in the Appendix of this paper.

(ii) New weapon acquisition/development: Hypothesis 1 investigates the possibility that testing missiles are necessary for the completion of a missile development program. Thus, we should see more missile tests as countries develop new missiles. Accordingly, I created a variable indicating how many new missiles were tested for the first time at t-1 (where t is the year of observation). Paper I of this dissertation details how I identified whether the
observed test of a missile was its first one. Of the 726 observations, 510 do not involve any new missiles tested at t-1. There are, however, 120 observations where there is at least 1 missile tested for the first time at t-1. The maximum number of new missiles tested at t-1 is 8 which happened in 1959 in the US. I also include the distribution of this variable in Figure 7 of the Appendix.

(iii) Militarized Interstate Dispute Involvement: To record a country’s involvement in a militarized interstate dispute, I used Dyadic MID v3.1 Dataset collected by Maoz et al (2019). This dataset clarifies a country’s role in a MID better than alternative versions such as Palmer et al (2015). Based on the information from Maoz et al (2019), I created 9 different variables. Each of these variables specify what sort of a role the country of interest plays in a MID.

First, I recorded a variable indicating the number of MIDs the country of interest was involved in, one year before (t-1) the observed time of its missile tests (t). Of the 688 country-year observations, 97 (14.2%) have no MID at t-1. 170 country years have one MID (24.8%), and 126 have 2 MIDs (18.5%) at t-1. The maximum number of MIDs in a year is 12 from the US in 1983 (observed as US-1984). The average for this variable is 2.5 and the median is 2. Since there seems to be enough variance accounting for 2 or more MID in a year, I did not dichotomize this variable.

112 Indeed, according to the data I collected, the US tested 202 missiles in 1959, and 298 in 1960. However, it tested 309 missiles in 1961. If this can be generalized for all the other years of the US, and for all the other countries, statistical measures should be such that new tests of weapons impact more than the immediate subsequent years.

113 The reason I use dyadic MIDs is because this dataset presents more accurate information about which country attacked which country as well as countries’ roles in a MID than Palmer et al’s (2015) participant level MID dataset.

114 Any variable I created using the MID dataset covers years 1949 and 2010, hence there are at least more the 50 observations dropped when I used MID variables in the statistical models.
The second and third variables I created separately count the times the country of interest was the primary initiator and primary target of a MID. It is possible that the more a country is targeted by others, the more it feels threatened. However, the more a country initiates a MID, the threat it perceives may be weaker. Of all the observations, 252 country years have no MIDs initiated at t-1 by the country of interest. 1 MID was initiated in 215 observations (31.52%), and 2 in 105 observations (15.4%). The maximum number of MIDs primarily initiated by a country is 9 (observation: Iran-1987). The median MIDs initiated is 1, and the mean 1.2. For the number of times a country was the primary target of MID in a year, there are 299 cases where the variable is 0 (43.8%). In 205 cases, a country was the primary target in a year only once (30.1%). The maximum value for this variable is 7 (US-1959), the median 1, and the mean is also 1.

The fourth and fifth variables are two dummy variables to account for joiners in a MID. The level of threat for MID joiners may potentially be weaker than primary initiator and target countries. Hence, the first binary variable is coded as 1 if the country of interest was a joiner in a MID on the side of the original initiator, and 0 otherwise. 70 out of 682 observations are coded as 1 (10.3%). Second, I created another binary variable marked as 1 if the country joined a MID on the side of the primary target. This latter variable is 1 in 112 observations (16.4 %). The reason I dichotomized these two indicators is because countries were rarely joiners more than once in a year.

Sixth and seventh, I sought to only account for MIDs that involve use of force or war (fatal MID, hereby). One reason is that many MIDs tend to involve border skirmishes

---

115 Morgan and Palmer (2000) also indicate that MID initiation may be less driven by a country’s need to maximize its security than its need to change the status quo in international politics.
or fishery versus coast guard encounters of little consequence to a country’s national security. Accordingly, I created two binary variables. The first is coded 1 if the country at least was involved in MID that had use of force or war by at least one of the participants (550 observations – 80.6%). The second is coded as 1 if the use of force or war was the highest action in a MID against the country of interest (511 observations, 74.9%).

Finally, the eighth and ninth variables are two binary variables that specifically capture war involvement of the country of interest. The first one is coded 1 in 82 observations (12.02%) if the country was involved in a war the previous year. The second variable is coded as 1 if the country of question was involved in a war the previous year where it was the primary target (21 observations, 3.08 %). These figures indicate that we have a sample that is far more war-prone compared to all countries in the MID dataset: in the aftermath of World War II, only 1.6 percent of all the MIDs recorded by Palmer et al (2015) escalated to war.

**(v) Adversaries’ Non-Violent Military Actions:** I define adversaries as countries that frequently antagonize each other in militarized interstate disputes. Therefore, in a given time, a country may have a higher expectation of finding itself in a MID against an adversary than a non-adversary. In that regard, I relied on the Dyadic MID dataset (Maoz et al 2019). For the present 11 countries of interest, I rank-ordered which other countries they were most paired with in a MID between 1949 and 2010. I subsequently coded those that were present in the simple majority (greater than 50%) of these MIDs as adversaries.

---

116 This definition is similar to Maoz and Mor’s (1993) rivalry definition. However, unlike these authors, I do not place a time restriction.
Overall, this approach may potentially be less arbitrary than alternative measures one can think of.\textsuperscript{117}

For each country in the present sample, no more than six countries constituted their top MID adversaries based on my measurement. This is in line with the findings from previous scholarly work. Just because a MID between two countries ends does not mean that they will not encounter each other in future MIDs. In fact, statistically speaking, two countries that fight each other in a MID are more likely to fight in subsequent MIDs (Maoz and Mor 1993).\textsuperscript{118} Thus, in the present sample, countries potentially are more likely to face their adversaries in a MID than any other country.\textsuperscript{119} Relatedly, adversaries’ actions such

\textsuperscript{117} For example, Strategic Rivalry criteria that Colaresi et al (2008) presented can be one way to detect a country’s adversaries. However, the method to determine a strategic rival is does not consider countries’ previous military encounters. Furthermore, countries may see each other as strategic rivals, yet it is possible that they may rarely or even never attack each other. Under these circumstances, the basis for strategic rivalry becomes unclear and arbitrary. To remedy this, the authors engage in an archival research to understand whether the countries suspected as strategic rivals pay attention to each other. This approach is difficult to replicate and too dependent on the authors’ arbitrary coding decisions (see ibid: Chapter II). Another approach is to simply pick those countries that are geographically contiguous with the countries of interest. After all, scholars suggest that contiguity is a factor that significantly increases the risk of conflict and even war among nations (see Bremer 1992). This criterion seems to work well in the present dataset for India and Pakistan, as well as Israel and its Arabic neighbors. However, for the rest, there are issues: can we consider Belgium, Switzerland, Italy and Germany as France’s adversaries in the post-World War II era? Similarly, can we consider Canada and Mexico as adversaries that the US would be concerned with in the same era? Can we argue that Eastern European countries with their own Communist regimes during the Cold War would be the USSR’s adversaries? Would China be North Korea’s adversary? Overall, relying on geographic contiguity may lead us to pick many adversaries that the countries of interest are neighbors yet also friendly with.

\textsuperscript{118} Admittedly however, the theory behind why this is the case is not well-developed by scholars who point to this statistical association. In fact, we can indirectly infer more theoretically-driven work that how countries militarily in a MID can be enough for them to cease hostilities for a long time (see Blainey 1973, Smith and Stam 2004, Wagner 2000)

\textsuperscript{119} The following countries were determined as potential/actual adversaries: For the US’ the adversaries are USSR/Russia (57 MIDs), China (52 MIDs), North Korea (47 MIDs), Iran after 1979 (21 MIDs) and Iraq before 2004 (21 MIDs). For the USSR/Russia, the adversaries are the US, Japan (35 MIDs), China (23 MIDs), Georgia (20 MIDs) after 1991, and England (15 MIDs). For China, the adversaries are the US, Taiwan (36 MIDs), India (750 MIDs), USSR/Russia (23 MIDs), South Korea (22 MIDs) and Vietnam (19 MIDs). For North Korea, the adversaries are South Korea (58 MIDs), and the US (47 MIDs). For India, the adversaries are recorded as Pakistan (58 MIDs), and China (27 MIDs). For Pakistan, India is the only adversary making up 54 percent of its MIDs. For France, the adversaries are Libya (11 MIDs), Yugoslavia/Serbia (9 MIDs), Iraq before 2004 (8 MIDs), USSR/Russia (5 MIDs) Tunisia (5 MIDs), and China (5 MIDs). For the UK, the adversaries are Iraq before 2004 (21 MIDs), USSR/Russia (15 MIDs), China (12 MIDs), Yemen (11 MIDs) and Yugoslavia/Serbia (10 MIDs). For Iran, the adversaries are Iraq (17 MIDs) and the USSR (16 MIDs)
as joint military exercises or missile tests may potentially be seen as threatening or as a sign of more MID\(\text{s}\) to come in the near future.

After designating adversaries for the countries of interest, I created two sets of covariates in accordance with Hypotheses III and IV. The first set focuses on the public missile tests and the second on the joint military exercises conducted by the adversaries. For missile tests, I created three variables that capture the missile tests conducted by the adversaries of the countries of interest. Accordingly, I first recorded the yearly number of all types of missile tests that the adversaries conducted at \(t-1\). The median for this variable is 33, the mean is 75.2 and the maximum is 557 missiles launched by adversaries. The countries that have the US as their adversary skew the distribution of this variable towards higher values.

Second, I created a variable indicating the tests of missiles with extended range by adversaries at \(t-1\). After all, many of the short-range missiles that countries test hardly reach crucial cities of their potential adversaries (with the exception of North Korea). Thus, tests of such short-range missiles may not be as threatening as longer-range missiles. In that regard, the second variable I created involved missiles that can be in the form of MRBM, before 1979. In the aftermath of the Islamic Revolution, the country’s adversaries are Iraq (35 MID\(\text{s}\)), the US (21 MID\(\text{s}\)) and Turkey (10 MID\(\text{s}\)). Finally, for Iraq, the adversaries are Iran (50 MID\(\text{s}\)), Kuwait (28 MID\(\text{s}\)), Turkey (25 MID\(\text{s}\)), the US before 2004 (21 MID\(\text{s}\)) and the UK (21 MID\(\text{s}\)) – again before 2004. For Israel, the adversaries are Syria (84 MID\(\text{s}\)), and Egypt (67 MID\(\text{s}\)) until Camp David 1978. In addition, in the case of Israel I went beyond my measurement strategies and added post-revolutionary Iran. Had I not done that, it would be the case that Israel did not have adversaries which tested missiles or conducted joint military exercises. Nevertheless, it is possible that Israel does not actually expect to have any direct MID\(\text{s}\) with Iran. Thus, this particular case is the most arbitrary adversary designation.
IRBM, ICBM, MRCM, IRCM or ICCM. The average for this variable is 27.3, the median is 12 and the maximum is 249 missiles tests conducted by adversaries at t-1.

Third, I recorded tests of missiles that have extended range and are capable of carrying nuclear warheads. Tests of nuclear capable missiles pose great concern for the world. That said, countries in general separately test nuclear warheads and missiles and rarely test-fire missiles with actual nuclear warheads. Nevertheless, as long as a missile is known to be nuclear-capable, its test-launch without a nuclear warhead can be concerning to observers. For example, North Korea, Iran, and Pakistan have never test-fired missiles with live nuclear warheads. However, they often test-fired missiles that are (potentially) capable of carrying nuclear warheads and received severe criticisms from the rest of the world for doing so. Ultimately, I entertained the possibility that countries may feel more threatened by adversaries that test nuclear capable and longer-range missiles. The variable in question has a mean of 20.3, a median of 8 and a maximum of 202 missile tests by adversaries at t-1.120

The second set of variables I coded capture joint military exercises (JMEs) conducted by the adversaries of the countries of interest. To gather data on JMEs, I used d’Orazio’s (2013) dataset which records all the JMEs conducted by nations between 1974 and 2015.121 D’Orazio’s dataset in its raw form is at the JME level. Thus, I reformatted it

120 For the adversaries’ missile tests variable, I only included adversaries for which I collected already collected data on. So, for example if Libya test-fired a missile, I did not record it as an additional adversary test for France. Similarly, a Turkish missile tests was not recorded in the case of Iran or Iraq. That said, countries besides the 11 ones I collected data on rarely test-fire their own missiles. Neither Libya, nor Turkey are known to test-fire their own indigenous or imported missiles (see Encyclopedia Astraunomica’s sections for Turkey and Libya).
121 D’Orazio’s original dataset on JMEs is available online at https://www.vitodorazio.com/data.html. The earlier version of the dataset was used by d’Orazio in its 2013 dissertation. Yet, since then he updated year coverage the dataset to 2015.
to a JME-participant level. Next, based on the adversary criteria described in the beginning of this subsection, I recorded how often the adversaries of the 11 countries of interest conducted JMEs in a given year.

A non-trivial detail about JMEs is that they happen in different formats and are not always the archetypical full-scale multinational naval wargames such as RIMPAC (Rim of Pacific Exercises). Indeed, JMEs can be in the form of Field Training Exercises (FTX) which involves practical training of military assets/vehicles and personnel and which is the closest activity to war/combat rehearsal. At the same time, JMEs can also be in the form of Command Post Exercises (CPX). The CPX type JMEs focus on training military staff for combat readiness. While JMEs can be exclusively in the form of FTX or CPX, they can also be rigorous and involve both forms. Intuitively, when a country’s adversaries conduct exercises that involve both CPX and FTX, such exercises may be the most intimidating ones.

To capture these dynamics, I created two binary variables regarding the JMEs. The first is coded as 1 if the country’s adversaries conducted at least one JME at t-1.\textsuperscript{122} 332 of the 451 observations (73.61\%) for this variable is coded as 1. The second binary variable captures a more involved form of JMEs, namely those that involve both CPX and FTX. Out of 451 cases for which I was able to gather information, the adversaries of the countries in question conducted such JMEs 75 times (16.7\%). That said, for many JMEs, d’Orazio (2013) was not able to identify the training format of the JME in question. Thus, the 1’s of this latter binary variable records the JMEs identifiable as involving both CPX and FTX.

\textsuperscript{122} I dichotomized the JME variables because most of the time, the countries typically have 1 JME of their adversaries at t-1 (unless their adversary is the US).
(vi) Economic Sanctions: Hypothesis 5 suggests that imposed economic sanctions will have a positive effect on subsequent tests of commissioned missiles. With this in mind, I created four separate variables using the Threat and Impositions of Economic Sanctions (TIES) v4.0 Dataset by Morgan et al (2013). The TIES dataset records sanction episodes between 1945 and 2005 which involve threats, and/or impositions of economic sanctions. Since the relevant hypotheses focus on the actual economic pressure caused by sanctions, I record economic sanctions where the countries of interest were primary targets and which involved an actual imposition (regardless of whether sanction senders issued a threat or not).

The first variable I created indicates whether the country of interest had an economic sanctions episode at t-1. Of the 633 observations, 309 involved a sanction episode at t-1 (which is 48.9%). Second, I created another binary variable indicating whether the country of interest received an economically severe or major sanction by any sender. Since, for various reasons, there are many sanctions that have trivial economic effects on a country\textsuperscript{123}, I sought to focus on the ones that place substantial economic burden on the target country. In doing so, I relied on a variable that Morgan et al 2013) created which shows an estimated cost of the imposed sanction to the target country. Of the 633 country-year observations, 61 (i.e. 9.6%) has a severe or major sanction at t-1.

(vii) Additional Control Variables: I also included several control variables which can potentially affect how frequently a country tests missile. First, I created a binary variable showing whether the observed year is during the Cold War (i.e. 1949-1991). 471 of the 264

\textsuperscript{123} These sanctions could target a specific economic sector, the leadership or a political party of a country. So, such sanctions may have little or no impact on the general welfare of this countries.
observations took place during the Cold War. The arms and space race between the US and the USSR during the Cold War era might have a marginal impact on how frequently they test-fired missiles.\textsuperscript{124}

Second, I created another binary variable indicating whether there was a change in leadership at t-1 in the country of interest. To operationalize this variable, I used Goemans et al.’s (2009) Archigos v. 4.1 project which records leaders of countries between 1875 and 2015. The leader-level dataset includes information on dates when leaders started their tenure. Using this information, for each observation, I record whether there was a new leader at t-1. There are a handful of cases in which there were more than one leadership changes. Yet, since this was not a large number, I contended with a binary variable.

Based on the anecdotal evidence in the case of North Korea (and more specifically Kim Jong-Un), I expect that new leaders may be more interested in trying to intimidate potential aggressors through missile tests than experienced leaders. The reason may be that there is often greater uncertainty about new leaders’ skills, and so other countries may try to test these skills through escalatory military actions (see Wolford 2007). Hence, a possibility is that new leaders respond to such tactics through showing off with missile tests. Overall, in 161 of the 736 country years, I was able to identify at least one leadership change at t-1.

Third, I suspected that greater economic wealth should increase how frequently countries test-fire missiles. After all, the US is the wealthiest country in the sample which is often followed by the USSR. These two countries overall are the most-frequent missile

\textsuperscript{124} Indeed, we see such a pattern in Figure 1.
testers. To see whether the impact of wealth can be generalized across all the countries in the sample, I recorded the real GDP values for all the countries using Gleditsch’s (2002) most recent version of Expanded Trade and GDP dataset. I lagged the GDP variable by one year and took its natural logarithm.

Fourth, I created a variable which indicates the percentage change in the military expenditures of the country in question from t-2 to t-1. I expected that greater military spending may also mean greater budget for missile development programs, hence greater subsequent missile tests. To collect data on military spending, I used Greig et al’s (2017) most recent version of the National Material Capabilities Dataset – originally presented by Singer et al (1972). Like all the other variables, I lagged this military spending variable. Overall, I was able to code change in military spending for 661 observations since the latest year for which the data was gathered is 2012. The mean for this variable is 0.11, and the median is 0.07.

Fifth, using SIPRI’s dataset on Military Imports, for each country I calculated the percentage change in the dollar value of the missile imports over total weapon imports from year t-2 to t-1. In 254 of the 732 observations, I identified no change in a country’s missile imports from t-2 to t-1 (i.e. the variable is 0). 249 observations have a decrease in the missile imports from t-2 to t-1 (i.e the variable is strictly smaller than zero). For the remainder of the observations, there is an increase in the share of missile imports over total weapon imports. The mean for this variable is 0.34 and the median is 0.

Lastly, I created a variable to capture the level of democracy of the country of interest. I expected that more liberal polities may be less in favor of developing and testing missiles. Alternatively, it could be that the greater voter coalition an executive needs to
stay in power, the less funds it may choose to channel from social spending to developing and testing missiles. In other words, holding all else constant, greater levels of democracy may be inversely related to subsequent missile tests. To measure democracy, I relied on the most recent version of the Polity IV dataset (Marshall et al 2016). The Polity II variable in this dataset presents a composite measure of the institutional level of democracy ranging between -10 and 10. Greater values of this variable indicate greater levels of democracy for the country of interest. Roughly-speaking, 43 percent of the observations have a Polity score less than 0 – which is, for many studies, considered undemocratic. The median for this variable is 5, and the mean is 1.37. Table A1 in the Appendix presents summary statistics for all the variables (except the missile tests) employed in the empirical analyses.

**Statistical Analyses**

Since the outcome of interest is the number of missile tests in a given year, I estimated Negative Binomial models to test the present hypotheses. The reason why I employed Negative Binomial specification is due to the considerable overdispersion in the outcome variable. To explore this aspect of the data, I first ran a simple Poisson model with only the missile tests variable. Then I tested the null hypothesis that the overdispersion in the model is less than or equal to 1. The overdispersion parameter is 10.01, and I was able to reject the null with p<0.001, z=7.89.\(^\text{125}\) This lends greater validity for discarding Poisson estimation and resorting to Negative Binomial regression instead. For each model, I clustered the standard errors by countries. The expected rate of incidence (i.e. count of missile tests), \(\lambda_t\) is modeled as

\[
E[\lambda_t] = e^{X_t\beta},
\]

\(^{125}\) I used Kleiber et al's (2008) `AER` Package in RStudio to conduct the overdispersion test.
\[ X\beta = \alpha + \beta_1*\text{main treatment}_{t-1} + \beta_2*\ln(\text{GDP})_{t-1} + \beta_3*\text{Polity II}_{t-1} + \beta_3*\text{New leader}_{t-1} + \beta_4*\text{% Change mil. spend.}_{t-1} + \beta_5*\text{New missiles}_{t-1} + \beta_6*\text{% Change missile Imp.}_{t-1} + \beta_7*\text{Cold war}_{t} \]

The estimated models are structured such that the treatment variable for Hypothesis 1, which is \( \text{New missiles}_{t-1} \), is always included in the analyses. I included the treatment variables for the remaining hypotheses separately and investigated what values \( \beta_1 \) takes as the impact of \( \text{New missiles}_{t-1} \) on Missile tests at \( t \) is partialled out. This way, I can investigate Hypotheses 2, 3, 4, and 5 by controlling for a very conventional explanation for why countries test-fire missiles. Table 2 displays the results for the Hypothesis 2.

According to Hypothesis 1, more new missiles test fired at \( t-1 \) should be positively associated with more missile tests in the observed year \( t \). This expectation is supported across all 9 models estimated in Table 2. Moreover, the coefficients on the \( \text{new missiles} \) variable at \( t-1 \) are strictly greater than zero within their 95% confidence margins. This is in line with what one would conventionally expect. The more newly developed missiles are test fired at time \( t-1 \), the more these missiles should test-fired further in subsequent time periods until they are considered operationally capable and accepted into service.

There is nothing significantly out of the ordinary about the finding that missiles are test-fired because these tests are a necessary phase before commissioning these missiles. The question is whether there are additional factors that account for the variance in yearly missile tests. Accordingly, Hypothesis 2 expects that missile tests conducted at year \( t \) may be positively related by the militarized interstate disputes (MIDs) that the country of interest involved in at \( t-1 \). Models 1 through 9 employ 9 different types of MID involvements as the treatment variables (in addition to new missile tests at \( t-1 \)).
Overall, estimates in Table 2 suggest that only some types of MID involvement are statistically significantly, and positively associated with public missile tests. First, in Model 1, the general MID involvement variable has a positive and statistically significant association with missile testing within the 95 percent confidence level. This finding already shows that there is some support for Hypothesis 2.

However, upon further investigation into how the country of interest was involved in the MID, the findings change. In Model 2, the number of MIDs initiated by the country at t-1 have no statistically significant impact on its missile tests at year t, although the direction of estimate happens to be typically positive. On the other hand, as Model 3 shows, if we account for the number of times the country of interest was the primary target of a MID, this variable appears to have a statistically significant and positive association with the missile tests in the subsequent year. Overall, it seems that the statistical significance in direct MID involvement is particularly driven by if the country of interest was targeted by a MID and did not initiate the MID itself.

If we consider joiners to MIDs in Models 4 and 5, we observe different results based on whether the country of interest joined the MID on the side of the primary initiator, or the primary target. In Model 4, the binary variable recording whether the country of interest joined any MID on the side of the initiator has a negative coefficient. However, this effect is statistically insignificant. On the other hand, in Model 5, the impact of joining any MID on the side of the primary target has a positive and statistically significant effect on missile testing the subsequent year.

Models 6 and 7 focus on another subset of MIDs as the main treatment variable, namely MIDs that involve uses of force (i.e. fatal MIDs) in Models 6 and 7. Model 6
investigates whether any involvement in a fatal MID is positively associated with the number of missile testing in the subsequent year. Model 7 focuses on fatal MIDs where the country of interest was attacked. Both fatal MID variables have statistically significant and positive impacts on the number of missile tests, which is expected if we buy into the argument laid out for Hypothesis 2. Fatal MIDs are not mere MIDs that only consist of a threat to use force or a display of force. One or two sides actually attack each other in various forms. Thus, such MIDs may constitute a substantial level of military threat which may translate into an increase in the involving countries’ subsequent missile tests.
### Table 2: Negative Binomial Estimates of Missile Tests (Hypothesis 2)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MID (t-1)</td>
<td>0.172***</td>
<td>(0.032)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MID initiated (t-1)</td>
<td>0.047</td>
<td>(0.061)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MID initiated against (t-1)</td>
<td>0.308***</td>
<td>(0.062)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joined MID initiator (dummy, t-1)</td>
<td>-0.250</td>
<td>(0.240)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joined MID target (dummy, t-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.774***</td>
<td>(0.179)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatal MID (dummy, t-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.598***</td>
<td>(0.194)</td>
<td></td>
</tr>
<tr>
<td>MID, force against (dummy, t-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.462**</td>
<td>(0.180)</td>
</tr>
<tr>
<td>War (dummy, t-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.272</td>
<td>(0.220)</td>
</tr>
<tr>
<td>Primary target in war (dummy, t-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.677</td>
</tr>
<tr>
<td>GDP(logged, t-1)</td>
<td>0.887***</td>
<td>(0.076)</td>
<td>1.060***</td>
<td>(0.065)</td>
<td>1.089***</td>
<td>(0.076)</td>
<td>1.091***</td>
<td>(0.078)</td>
<td>1.016***</td>
</tr>
<tr>
<td>Poiety 2 (t-1)</td>
<td>-0.037***</td>
<td>(0.011)</td>
<td>-0.041***</td>
<td>(0.010)</td>
<td>-0.043***</td>
<td>(0.010)</td>
<td>-0.044***</td>
<td>(0.010)</td>
<td>-0.046***</td>
</tr>
<tr>
<td>New leader (t-1)</td>
<td>0.076</td>
<td>(0.105)</td>
<td>0.034</td>
<td>(0.180)</td>
<td>0.171</td>
<td>(0.180)</td>
<td>0.010</td>
<td>(0.180)</td>
<td>0.012</td>
</tr>
<tr>
<td>% Change mil. spend (t-1)</td>
<td>-0.497</td>
<td>(0.266)</td>
<td>-0.526</td>
<td>(0.273)</td>
<td>-0.513</td>
<td>(0.273)</td>
<td>-0.538</td>
<td>(0.273)</td>
<td>-0.567</td>
</tr>
<tr>
<td>New missiles (t-1)</td>
<td>0.817***</td>
<td>(0.057)</td>
<td>0.683***</td>
<td>(0.054)</td>
<td>0.664***</td>
<td>(0.054)</td>
<td>0.707***</td>
<td>(0.054)</td>
<td>0.672***</td>
</tr>
<tr>
<td>% Change missile Imp. (t-1)</td>
<td>-0.082</td>
<td>(0.035)</td>
<td>-0.001</td>
<td>(0.048)</td>
<td>-0.001</td>
<td>(0.048)</td>
<td>0.000</td>
<td>(0.048)</td>
<td>0.000</td>
</tr>
<tr>
<td>Cold War (t)</td>
<td>0.116</td>
<td>(0.215)</td>
<td>0.038</td>
<td>(0.239)</td>
<td>0.213</td>
<td>(0.190)</td>
<td>0.054</td>
<td>(0.227)</td>
<td>-0.112</td>
</tr>
<tr>
<td>Intercept</td>
<td>-25.780***</td>
<td>(2.171)</td>
<td>-27.298***</td>
<td>(2.118)</td>
<td>-28.497***</td>
<td>(1.865)</td>
<td>-28.073***</td>
<td>(2.191)</td>
<td>-26.092***</td>
</tr>
</tbody>
</table>

| Observations | 641 | 641 | 641 | 641 | 641 | 641 | 641 | 641 | 641 |
| theta            | 0.477*** (0.035) | 0.447*** (0.032) | 0.482*** (0.035) | 0.449*** (0.032) | 0.465*** (0.034) | 0.458*** (0.034) | 0.434*** (0.033) | 0.446*** (0.032) | 0.447*** (0.032) |

Note: p<0.1, **p<0.05, ***p<0.01

Robust standard errors in parentheses, clustered by country
When MIDs that escalated to war are taken into account in Models 8 and 9, I find no statistically significant result in a particular direction. Moreover, any involvement in a MID that ended up in war at t-1 has a positive but statistically insignificant impact on the count of missile tests at t – as seen in Model 8. That said, if the country was the primary target in a MID that ended up in a war, this has a negative but statistically insignificant association with subsequent missile tests. One reason could be that wars generally are the ultimate means for countries to achieve their goals and solve their problems (Clausewitz 1832). Therefore, once a war is fought and especially if it ends in a decisive victory of one side, the conflict which led to the war in the first place might have been resolved (Blainey 1973, Smith and Stam 2004). Hence a military threat after a war may be irrelevant – especially after a decisive victory. This may be why war has no clear impact on subsequent missile tests by the belligerents. Alternatively, the results may simply be driven by the rarity of MIDs that end up in wars in the present sample.

To better make sense of the statistically significant results, I calculated the predicted counts of missile tests across simulated values of treatments for Models 1, 3, 6 and 7 of Table 2.\footnote{In doing so, I sampled 1,000 \( \beta \) estimates from a multivariate normal distribution with means extracted from the actual model of interest, and the variance from the robust variance-covariance matrix again from the actual estimated model. Then, I calculated the predicted missile count \( E[\lambda_t] = e^{X\beta} \).} for Models 1 and 3, I simulated 11 values of general MID involvement (from 0 to 10 with 1 increments). I fixed the remaining control variables to their means (if continuous) or medians (if categorical). For each plot, I calculated not only the predicted mean count of missiles, but also the 95% upper
and lower boundary predictions. The shaded lines for the predictions using Models 1 and 3 represent 95% confidence boundaries.

As can be observe in Figure 21, the substantive impact of being targeted by a MID (Model 3) is greater than generally being involved in a MID (Model 1). Moreover, for a country to test-fire at least one more missile, it typically has to be involved in 3 additional MIDs – holding all else constant. If the country is involved in, for instance, 10 MIDs (as initiator, target, or joiner), this leads to typically 10 missile tests in the subsequent year. However, according to the Model 3, if the country was the primary target in 5 MIDs, this would increase the number of missiles tested the next year by 10. In other words, the impact of being the primary target of a MID is typically twice as strong as simply being involved in a MID.

**Figure 21 Substantive Effects of Various Forms of MID involvement on Missile Tests**
The substantive effects for fatal MID involvement (Model 6) and being the primary target of a fatal MID (Model 7) at t-1 on missile tests at t appear not to be strong – according to the estimates in Figure 4. Moreover, being involved in a fatal MID at t-1, increases the number of expected missile test the next year by approximately 2. However, the 95% upper and lower bounds (red lines in Figure 4) of the two simulated values of fatal MID involvement (i.e. 0 and 1) weakly intersect. Thus, there is a small chance that the impact of being involved in a fatal MID will not substantively increase the expected missile test count the next year.

In the case of Model 7, we observe that being the target of a fatal MID at t-1 approximately increases predicted missile test count at t by 1. Again, the 95% margins for both simulations have several estimates in common which weakens the confidence in the impact of being targeted by a fatal MID.

In terms of control variables for Models 1 through 9 of Table 2, we mostly observe intuitive results. First, in all the models, greater GDP at t-1 is positively and significantly associated with greater missile tests at t. This finding is intuitive since missile tests are expensive, and a country needs to allocate large sums to be able to conduct these tests. Second, the Polity II score has a negative and statistically significant effect on missile tests. This finding may be driven by Israel, France and England as well as Russia which conducted far fewer missile tests after the collapse of the USSR. The United States may be the only full democracy that test-fires more missiles than what we typically observe. Third, the impact of leadership change at year t-1 is typically positive, yet statistically insignificant within 95% confidence level. Fourth, increase in military spending from year t-2 to
t-1 leads to a decrease in missile tests at year t. This effect is statistically significant, and counter-intuitive. Fifth, an increase in the share of missile-related imports typically has a negative impact on missile tests in all the models. Nevertheless, this variable’s impact is only statistically significant in Models 1, 3, 5 and 6. This finding is not counter-intuitive and actually indirectly supports the first hypothesis. The more missiles a country imports, the fewer missiles of its own it would have to develop and subsequently test-fire.\footnote{South Korea is a good example to make sense of this finding. The reason why I did not include South Korea in the sample of missile testing countries is because I was unable to find enough relevant data for this country. Although South Koreans are recently developing ballistic technologies of their own, most of their missiles and missile defense systems (such as THAAD) are imported from other countries such as the US.} Lastly, I find that the binary variable for Cold War has a positive, though statistically insignificant, impact on missile tests for all the models.

Missile tests and joint military exercises by adversaries can be military threats that serve as alternatives to MID involvement. Accordingly, Hypotheses 3 and 4 anticipate that such activities at time t-1 should increase a country’s missile tests at t where t is the observed year. Table 3 presents the results of analyses that investigate these hypotheses with various measurements of adversaries’ missile tests and JMEs. Unfortunately, none of the analyses reveal statistically significant results.

Models 1 through 3 in Table 3 address the third hypothesis with various specifications of missiles that a country’s adversaries test. However, these models fail to find statistically significant effects within the 95% confidence level. Model 1 investigates all types of missiles test-fired by adversaries. Although the
coefficient of this variable is estimated to be greater than zero ($\beta_1=0.001$), the standard error of the estimate is too large ($se_{\beta_1}=0.001$), which suggests that – within 95% confidence levels – there is no discernable impact of adversaries’ missile tests. A similar pattern is observed in Model 2, where I only consider adversaries’ tests of missiles with intermediate or longer ranges. A positive yet statistically insignificant coefficient is estimated. For the third model, I also fail to find any statistically significant result. Moreover, its adversaries’ tests of intermediate or greater range missiles that are also nuclear capable do not have a consistently positive relationship with the countries’ missile tests at t. Overall, I fail to find support for Hypothesis 3.

Models 4 and 5 investigate whether JMEs conducted by adversaries at t-1 increase a country’s missile tests at t. In both models, the coefficients are statistically insignificant, although typically positive. Moreover, Model 4 shows that any JME that a country’s adversaries conduct the previous year has a typically positive impact ($\beta_1=0.17$) on the country’s missile tests. However, this coefficient estimate also has a large standard error ($se_{\beta_1}=0.24$), which suggests that the impact of JMEs can be either zero or even negative within 95% confidence level. In the case of JMEs that involve both Field Training Exercises (FTX) and Command Post Exercises (CPX), I still find a typically positive coefficient estimate. However, this estimate is quite small (0.001) whereas its standard error relatively large (0.22). Overall, I again fail to find support for Hypothesis 4.
Table 3: Negative Binomial Estimates of Missile Tests (Hypotheses 3 - 4)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adversaries' Tests (lagged)</td>
<td>0.091</td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ad's Tests w/ Short Range (lagged)</td>
<td>0.002</td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ad's Tests w/ Short Range + Nuclear (lagged)</td>
<td>0.001</td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adversaries' JMEs (lagged)</td>
<td></td>
<td></td>
<td>0.168</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.242)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ad's FTX and CPX (lagged)</td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
<td>(0.219)</td>
</tr>
<tr>
<td>GDP(logged, lagged)</td>
<td>1.062***</td>
<td>1.067***</td>
<td>1.077***</td>
<td>0.767***</td>
<td>0.773***</td>
</tr>
<tr>
<td>(0.079)</td>
<td>(0.078)</td>
<td>(0.077)</td>
<td>(0.072)</td>
<td>(0.068)</td>
<td></td>
</tr>
<tr>
<td>Policy 2 (lagged)</td>
<td>-0.040***</td>
<td>-0.040***</td>
<td>-0.042***</td>
<td>-0.035***</td>
<td>-0.035***</td>
</tr>
<tr>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>New Leader (lagged)</td>
<td>0.026</td>
<td>0.018</td>
<td>0.033</td>
<td>0.133</td>
<td>0.102</td>
</tr>
<tr>
<td>(0.179)</td>
<td>(0.176)</td>
<td>(0.178)</td>
<td>(0.166)</td>
<td>(0.165)</td>
<td></td>
</tr>
<tr>
<td>%s Change Mil. Spend. (lagged)</td>
<td>-0.521*</td>
<td>-0.518*</td>
<td>-0.530*</td>
<td>-0.304</td>
<td>-0.309</td>
</tr>
<tr>
<td>(0.286)</td>
<td>(0.285)</td>
<td>(0.280)</td>
<td>(0.209)</td>
<td>(0.210)</td>
<td></td>
</tr>
<tr>
<td>New Missiles (lagged)</td>
<td>0.692***</td>
<td>0.657***</td>
<td>0.694***</td>
<td>0.399***</td>
<td>0.406***</td>
</tr>
<tr>
<td>(0.060)</td>
<td>(0.059)</td>
<td>(0.059)</td>
<td>(0.065)</td>
<td>(0.065)</td>
<td></td>
</tr>
<tr>
<td>%s Change Missile Imp. (lagged).</td>
<td>-0.066*</td>
<td>-0.066*</td>
<td>-0.064</td>
<td>-0.046</td>
<td>-0.037</td>
</tr>
<tr>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.049)</td>
<td>(0.049)</td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>Cold War</td>
<td>-0.032</td>
<td>-0.019</td>
<td>0.029</td>
<td>0.214</td>
<td>0.173</td>
</tr>
<tr>
<td>(0.244)</td>
<td>(0.237)</td>
<td>(0.233)</td>
<td>(0.209)</td>
<td>(0.224)</td>
<td></td>
</tr>
<tr>
<td>(2.283)</td>
<td>(2.249)</td>
<td>(2.237)</td>
<td>(2.027)</td>
<td>(1.966)</td>
<td></td>
</tr>
</tbody>
</table>

Observations: 641, 641, 641, 388, 388
Log Likelihood: -1,702.719, -1,702.769, -1,703.163, -1,120.916, -1,121.260
Theta: 0.445*** (0.032), 0.445*** (0.032), 0.444*** (0.032), 0.613*** (0.034), 0.611*** (0.054)
Akaike Inf Crit.: 3,423.439, 3,423.539, 3,424.325, 2,259.832, 2,260.519

Note: *p<0.1, **p<0.05, ***p<0.01

The last hypothesis investigates whether a country under economic sanctions test fires more missiles that are already in service. Accordingly, Table 4 presents 4 models that investigate Hypothesis 5. Overall, there is no statistically significant support for the Hypothesis 5. Economic sanctions do not appear to significantly increase missile tests. Additionally, they do not appear to significantly decrease missile tests either.
Table 4: Negative Binomial Estimates of Already Commissioned Missile Tests (Hypothesis 5)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Commissioned Missiles</th>
<th>New Missiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Sanction Episode (lagged, dummy)</td>
<td>0.174 (0.213)</td>
<td>0.141 (0.155)</td>
</tr>
<tr>
<td>Severe/Major Sanction (lagged, dummy)</td>
<td>-0.409 (0.364)</td>
<td>0.003 (0.336)</td>
</tr>
<tr>
<td>GDP(logged, logged)</td>
<td>1.542*** (0.122)</td>
<td>1.552*** (0.125)</td>
</tr>
<tr>
<td>Polity 2 (lagged)</td>
<td>-0.053*** (0.014)</td>
<td>-0.054*** (0.014)</td>
</tr>
<tr>
<td>New Leader (lagged)</td>
<td>0.122 (0.265)</td>
<td>0.129 (0.263)</td>
</tr>
<tr>
<td>% Change Mil. Spend (lagged)</td>
<td>-0.514 (0.464)</td>
<td>-0.533 (0.462)</td>
</tr>
<tr>
<td>New Missiles (lagged)</td>
<td>0.635*** (0.069)</td>
<td>0.621*** (0.070)</td>
</tr>
<tr>
<td>% Change Missile Imp (lagged)</td>
<td>-0.057 (0.041)</td>
<td>-0.048 (0.041)</td>
</tr>
<tr>
<td>Cold War</td>
<td>0.075 (0.254)</td>
<td>0.110 (0.253)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-41.772*** (3.481)</td>
<td>-41.942*** (3.588)</td>
</tr>
</tbody>
</table>

Observations | 595 | 595 | 595 | 595
Log Likelihood | -997.374 | -997.185 | -484.951 | -485.399
theta | 0.302*** (0.030) | 0.303*** (0.030) | 1.522*** (0.385) | 1.512*** (0.382)
Akaike Inf. Crit. | 2,012.748 | 2,012.370 | 987.903 | 988.797

Note: *p<0.1; **p<0.05; ***p<0.01
Panel-corrected standard errors estimated and provided in parentheses

Models 1 and 2 in Table 4 are estimated with commissioned missile tests at time t as the outcome variable. In Model 1, the main treatment is any sanction episode at time t-1. This variable has a positive coefficient that is statistically insignificant. In Model 2, the main treatment variable is whether or not the country of interest had a severe or major sanction episode at time t-1. The impact of this variable appears to be typically negative yet is, again, statistically insignificant.
Models 3 and 4 estimate the count of new types of missiles test-fired at time $t$. It seems that both any sanction episode (Model 3), and severe/major sanction episode (Model 4) have a typically positive, yet statistically insignificant impact. Overall, these results do not provide support for Hypothesis 5.

**Robustness Checks**

**Crises instead of MIDs:** An alternative measure of military threats which may serve a role similar to MIDs is international crises. According to Brecher (1996), an international crisis is an event during which countries have a time pressure to respond to a political conflict with their adversaries. More importantly, during a crisis, war as a response to solve the pressing conflict is eminent and is thus significantly more likely than non-crisis situations. Many scholars of warfare used the framework of “crisis” to evaluate whether countries solved their differences through war or negotiations (Fearon 1995, Morgan 1984, Smith and Stam 2004). Therefore, empirically accounting for international crises can help us capture situations with a high dosage of military threat. Consequently, in line with Hypothesis 2, we should expect more missile tests following international crises.

To measure international crises, I used International Crisis Behavior (ICB) dataset and recorded the crisis behavior of the countries of interest between 1949 and 2015 (Brecher et al 2017). A crisis can involve military or non-military response by one of the participating nations. Thus, I created a binary variable indicating whether the country was in a crisis with another at $t-1$. Then I created two additional binary variables. The first accounts for crises that involved a violent
or nonviolent military response by any of the crisis participants. The last binary variable indicates whether the country in question was involved in a crisis at time t-1 which had a violent military response (i.e. with casualties). Using these three variables as the main treatment, and control variables, I estimated three additional negative binomial models. As before, I clustered the standard errors by countries. Table 5 displays the results.

I find that involvement in any crisis (Model 1, Table 5) or in a military crisis (Model 2, Table 5) at time t-1 has a statistically significant positive effect on missile tests at t within 95% confidence level. However, the impact of being involved in a fatal crisis is not statistically significantly positive at 95% confidence level. Overall, these results show that crisis involvement at time t-1 seems to typically increase the number of missile tests at t – which may be interpreted as an indirect support for Hypothesis 2.

---

128 To record this information, I relied on ICB Dataset’s variable with the default name **MAJRES** which records all the possible responses types to crises by countries. The response can be non-material/diplomatic, military or economic or a combination these responses (Brecher et al 2017). In the present sample, 142 observations had military crises (19.6%) at t-1

129 Only 92 (12.7%) of the 726 observations had such violent military crisis at t-1.
Table 5: Negative Binomial Estimates of Missile Tests

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Yearly Missile Tests (i)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Crisis (dummy, t-1)</td>
<td>0.424***</td>
</tr>
<tr>
<td></td>
<td>(0.152)</td>
</tr>
<tr>
<td>Military Crisis (dummy, t-1)</td>
<td>0.447***</td>
</tr>
<tr>
<td></td>
<td>(0.165)</td>
</tr>
<tr>
<td>Fatal Military Crisis (dummy, t-1)</td>
<td>0.225</td>
</tr>
<tr>
<td></td>
<td>(0.186)</td>
</tr>
<tr>
<td>GDP(logged, t-1)</td>
<td>1.055***</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
</tr>
<tr>
<td>Polity 2 (t-1)</td>
<td>-0.043***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
</tr>
<tr>
<td>New Leader (t-1)</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>(0.182)</td>
</tr>
<tr>
<td>% Change Mil. Spend (t-1)</td>
<td>-0.499*</td>
</tr>
<tr>
<td></td>
<td>(0.302)</td>
</tr>
<tr>
<td>New Missiles (t-1)</td>
<td>0.693***</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
</tr>
<tr>
<td>% Change Missile Imp. (t-1)</td>
<td>-0.058</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
</tr>
<tr>
<td>Cold War (t)</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>(0.230)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-27.142***</td>
</tr>
<tr>
<td></td>
<td>(2.209)</td>
</tr>
</tbody>
</table>

Observations: 641
Log Likelihood: -1,699.220
θ: 0.453*** (0.033)
Akaike Inf. Crit.: 3,416.439

* p<0.1; ** p<0.05; *** p<0.01

Robust standard errors in parentheses, clustered by country.

Country Sensitivity: A likely issue about the presently estimated models is whether the results they generate are driven by the observations of one particular country. After all, the sample consists of 11 countries. Some of the notable findings might be the result of one or a specific set of countries. Accordingly, I ran a series of country sensitivity tests for Models 1 and 3 of Table 2 where I found statistically
and substantively strong results backing Hypothesis 2. In doing so, I re-estimated each model 11 times on different samples where the observations of one country are dropped. This approach allowed me to observe to what degree the estimates of interest changed when one of the 11 countries were dropped from the sample. Next, I plotted the coefficient (i.e \( \beta \)) estimates with their 95% confidence intervals extracted from these 11 models. Figure 22 displays this coefficient plot.

**Figure 22: Country Sensitivity Analyses for Models 1 and 3 of Table 2**

The left-hand side of Figure 22 displays the country sensitivity results for Model 1 of Table 2. We observe that the estimate of MID involvement at t-1 is slightly sensitive to dropping the US observations. Moreover, a small minority of the estimates within the 95% confidence interval are smaller than or equal to zero if we drop the US observations. Nevertheless, the impact of MID involvement is typically and predominantly greater than zero even when we drop the US observations. On the other hand, the findings of Model 3 appear to be slightly more sensitive to dropping the US observations. Even though the impact of being targeted
in a MID typically and predominantly increases the expected number of missile
tests in the subsequent year, there is a slight chance that this impact is smaller than
or equal to zero within 95% confidence level. We also observe that dropping none
of the remaining countries change the impact of MID involvement or being targeted
in a MID.

I also conduct a similar sensitivity analysis in the case of Model 1 of Table
2 to evaluate the country-sensitivity of the “new missiles (t-1)” estimate pertaining
to Hypothesis 1. Figure 23 presents the estimates of new missile tests (t-1) for each
model where observations of one country were dropped. It seems that the findings
for the “new missiles” variable is not sensitive to any particular country’s
observations.

**Figure 23: Country Sensitivity Analyses for Model 1 of Table 2 (New
Missiles)**

![Model 1 - New Missiles (t-1) Sensitivity](image)

Zero-Inflated Negative Binomial Regression: Given the overwhelming number
of zeros in missile tests, a reasonable suspicion is whether there is an exclusive
process that determines the change from zero to one in this variable. One way to tackle with this problem is to use zero-inflated negative binomial model. More specifically, this approach models the change from 0 to 1 in the missile test count as a logit model as well as a negative binomial model. The rest of the values are still estimated through negative binomial specification. Models 1 and 3 of Table 2 are re-estimated using Zero-Inflated Negative Binomial approach. For both models, I use logged GDP as the variable to account for the zero inflation. I find that the results do not change with this alternative estimation method. In addition, I conduct a likelihood-ratio test presented by Vuong (1989) between the negative binomial and zero-inflated negative binomial models. I fail to reject the null that the two parametric specifications are statistically distinguishable. Therefore, I choose not to report the results of the zero-inflated negative binomial regressions.

**Alternative outcome variables:** The models I estimated in this paper use all types of missiles as the outcome variables. What has not yet been investigated here is how the present covariates of interest impact specific subset of missiles being test-fired. After all, a treatment that is found to inconsistently (statistically insignificantly) affect all types of missile tests may actually have a consistent effect on certain types of missiles. While this would not change the fact that I failed to reject the null hypothesis in this present paper, it would still show us the limitations of the initial findings. For example, what if initiating MIDs has actually a statistically significant and positive impact on new missiles test-fired at t? This

---

130 I use ‘pscl’ package in RStudio to estimate the negative binomial regressions as well as Vuong’s likelihood ratio test (Jackman 2007)

131 The economic sanctions-based models are an exception where the outcome variable of interest is commissioned missiles
finding would be noteworthy since the initial analyses here showed that MID initiation has no consistent impact on all the missile tests.

Accordingly, I re-estimated all 16 models in Tables 2, 3 and 4 using different types of missile tests. These are (i) number of missiles test-fired for the first time, (ii) number of tested ICBMs, (iii) number of commissioned missiles being tested, (iv) number of nuclear explosion tests, and (v) tests of missiles that have greater than short range. In total, I estimate 80 negative binomial models in addition to the original 16. Table 6 presents a summary of these results.

Each cell in Table 6 indicates the impact of the row variable on the alternative outcome variable represented by the columns. I only report the direction of the estimate as well as whether it was statistically significant at the 95% confidence level. Accordingly, MID involvement at t-1 has a statistically significant and positive impact on new missiles, ICBMs, commissioned missile tests and greater than short-range missile tests. Nevertheless, it has a statistically insignificant negative effect on nuclear tests.

On the other hand, although the number of times a country initiated a MID has no consistent impact on the tests of all its missiles, this is not the case if we focus on a specific set of missiles. For example, more initiated MIDs at t-1 has a positive impact that is statistically significant on new missiles, ICBMs, commissioned missiles and missiles with greater than short range. The inconsistency of the impact of MIDs initiated by the country appears to be mainly driven by nuclear tests. Furthermore, more MIDs initiated at t-1 does not have a statistically consistent impact on nuclear warhead tests at t.
The number of times a country was targeted in a MID at t-1 was originally found to have a statistically significant and positive impact on missile tests at t. This finding is consistent across all the subsets of missile tests with the exception of the nuclear warhead tests. Moreover, being targeted at a MID at t-1 has a negative yet statistically insignificant impact on the number of nuclear warheads tests the country conducts at t.

For the rest of the covariates that were initially found to be statistically insignificant, the only variable that has a consistent impact on one of the alternative outcome variables is JMEs with FTX and CPX. More specifically, it seems that this binary variable has a positive impact on the number of ICBMs test fired at time t. The arguments presented in this paper, however, do not directly clarify why the subset of JMEs with FTX and CPX has a positive impact on the subsequent tests of ICBMs and no consistent impact on any other types of missiles. Besides this outcome variable, none of the variables that record sanctions, JMEs and tests of adversaries have a consistent impact on the alternative outcome variables.

Lastly, the total number of new missiles a country tests fires at t is the most consistent outcome variable that is positively affected by MID involvement variables. Moreover, of the 9 different MID involvement types at t-1, 6 has a positive impact that is statistically significant at 95% confidence level. In addition, 8 of the MID involvement variables have at least typically a positive impact on this variable.
### Table 6: Different Missile Types Estimated

<table>
<thead>
<tr>
<th>Main treatment</th>
<th>All Missiles (t)</th>
<th>New Missiles (t)</th>
<th>ICBMs (t)</th>
<th>Commissioned (t)</th>
<th>Nuclear Tests (t)</th>
<th>&gt; Short Range (t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MID involvement (t-1)</td>
<td>sig +</td>
<td>sig +</td>
<td>sig +</td>
<td>sig +</td>
<td>insig -</td>
<td>sig +</td>
</tr>
<tr>
<td>Primary MID initiator (t-1)</td>
<td>insig +</td>
<td>sig +</td>
<td>sig +</td>
<td>sig +</td>
<td>insig +</td>
<td>sig +</td>
</tr>
<tr>
<td>Primary MID target (t-1)</td>
<td>sig +</td>
<td>sig +</td>
<td>sig +</td>
<td>sig +</td>
<td>insig -</td>
<td>sig +</td>
</tr>
<tr>
<td>Joins MID on init's side (t-1)</td>
<td>insig -</td>
<td>insig +</td>
<td>insig +</td>
<td>insig -</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td><strong>H2</strong> Joins MID on target's side (t-1)</td>
<td>sig +</td>
<td>sig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td>Fatal MID involvement (t-1)</td>
<td>sig +</td>
<td>sig +</td>
<td>sig +</td>
<td>sig +</td>
<td>insig -</td>
<td>insig +</td>
</tr>
<tr>
<td>Fatal MID target (t-1)</td>
<td>sig +</td>
<td>sig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig -</td>
<td>insig +</td>
</tr>
<tr>
<td>War involvement (t-1)</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig -</td>
<td>insig +</td>
</tr>
<tr>
<td>War target (t-1)</td>
<td>insig -</td>
<td>sig -</td>
<td>sig -</td>
<td>insig -</td>
<td>sig -</td>
<td>insig -</td>
</tr>
<tr>
<td><strong>H3</strong> Adversaries' tests (t-1)</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td>Adv.s' &gt;Short tests (t-1)</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td>Adv.s' &gt;Short + nuke tests (t-1)</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td><strong>H4</strong> JMEs (t-1)</td>
<td>insig +</td>
<td>insig -</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td>JMEs with FTX and CPX (t-1)</td>
<td>insig +</td>
<td>insig -</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td><strong>H5</strong> Economic sanction (t-1)</td>
<td>insig +</td>
<td>insig +</td>
<td>insig -</td>
<td>insig +</td>
<td>insig +</td>
<td>insig +</td>
</tr>
<tr>
<td>Severe or major sanction (t-1)</td>
<td>insig -</td>
<td>insig +</td>
<td>insig -</td>
<td>insig -</td>
<td>insig -</td>
<td>insig -</td>
</tr>
</tbody>
</table>

**Notes:** (1) Every cell represents a multivariate model where the column variable is outcome, and the row variable is the treatment. Control variables are the same ones used across all the other multivariate models. (2) Sig. refers to coefficient estimates that are statistically significant at 95% confidence level. These are presented in italics. (3) Shaded cells indicate the estimates that directly assess the hypotheses.
Notice that the only variable that appears to significantly decrease the number of missile tests is if the country of interest was the target of a war. This finding has important policy implications. One of the strongest predictors of missile tests at time \( t \) is the number of missiles test fired for the first time (i.e. “new missiles” variable) at \( t-1 \). If waging war on a country has a consistently negative impact (at least with 95% probability) on the number of new missiles it test-fires, then this action could be the one way to reduce the missile tests of a country. The only issue is that wars are very costly policies.\(^{132}\) In what follows, I overview and discuss the findings.

**Conclusion and Discussion**

World War II demonstrated clearly how missile and warheads technology can help a country to significantly weaken or even defeat its opponents. Since the end of this war, missiles and missile-related technologies have become more and more destructive. At the same time, an overwhelming majority of these ballistic and cruise missiles have never been used in active combat situations.\(^{133}\) Instead, we have been seeing news of countries test-launching these weapons publicly. This paper presented the first empirical investigation of several potential reasons behind countries test-firing their missiles.

\(^{132}\) In addition, the negative associations between new missiles at \( t \) and being targeted by war at \( t-1 \) may be model and data-driven.

\(^{133}\) For example, the American Peacekeeper is an ICBM that can carry 10 warheads, and Soviet/Russian Voevod with similar attributes were never used in a war. North Korea tested many nuclear capable missiles and nuclear warheads. None were used against the US, South Korea or Japan. Even North Korea, detonated far more destructive nuclear devices that what the US dropped over Hiroshima and Nagasaki in 1945.
The findings from the empirical analyses suggest that developing new missiles and a strong economy are the strongest predictors of increases in a country’s missile tests. In addition, involvement in previous militarized interstate disputes appears to increase the missile tests that a country conducts. Yet, this latter finding is sensitive to sample specifications. On the other hand, I fail to find evidence to suggest that a country’s missile tests are a response to its adversaries’ missile tests and/or joint military exercises. In addition, economic sanctions appear to have no consistent impact on missile tests. Further robustness checks and alternative covariate specifications did not substantially change these non-findings.

How do we interpret these results? First, this paper counter-intuitively finds that increases in missile tests are not typically driven by a need to respond to military threats. One can come across news of North Korea test-firing missiles as a reaction to a joint military exercise between the US and South Korea, Pakistan test-firing a missile to reciprocate India’s tests, or Russia test-launching its missiles when the US dispatches naval vessels to the Black Sea. However, I find these individual stories to not apply to a more general sample of missile testing countries: adversaries’ joint military exercises or missile tests do not affect a country’s decision to increase the missiles it publicly test-launches. Even though involvement in recent militarized disputes typically increases the number of missiles a country publicly launches, this effect is rather weak not robust.

Second, missile tests seem to be very intuitively driven by countries’ need to develop new missiles or to maintain existing ones. When a newspaper covers a missile test news as “Country X flexed its muscles by test-firing missiles to
intimidate its opponents”, this is not an inaccurate description. Flexing muscles not only is a threatening sight for potential aggressors, but also a means to ensure muscular development. In this paper, flexing muscles seems to be typically driven by maintaining or improving muscular development instead of a need to fend off aggressors. In other words, countries appear to typically test-fire missiles because these tests help with developing new missiles or maintaining old ones. For example, when North Korea test-launched a Hwasong-15 in late 2017, many interpreted it as a reaction to recent US-South Korea joint military exercises. This ad hoc interpretation may be true – for this particular case. Nevertheless, this paper highlights the strong possibility that North Korea’s test was scheduled way before the US-South Korea exercise, with rather the purpose of investigating whether Hwasong-15 is a reliable nuclear-capable ICBM.

Third, strong economy is a robust predictor of increases in missile tests. There is not only an intuitive, but also a counter-intuitive aspect to this finding. Missile tests are very costly. Thus, it makes sense that countries with strong economies can afford these tests. After all, while the US test-launches missiles by thousands, countries like Pakistan, Iran, North Korea and even Israel, France or England seldom test-launch more than twenty missiles in a given year. The counter intuitive part of this story is that we often read news of North Korea or Iran test-launching missiles when they face a foreign threat. Absent a large-N statistical analysis, one can infer that test-launching missiles is the strategy of a weak country to intimidate strong ones. This paper indicates that quite the opposite is to be the
case: if economy is an indicator of a country’s strength, then the stronger a country is the more missiles it is capable of test-launching.

Lastly, sanctions do not seem to consistently affect how many missiles a country can test-fire subsequently. For years, not only talks like SALT-I or SALT-II, but also economic sanctions were sought to diminish missile tests. North Korea received many sanctions in that regard. However, North Korean missile tests did not decrease. I developed a hypothesis that suggested that countries may actually test-fire their commissioned missiles to attract potential clients. Nevertheless, I failed to find support for this hypothesis either. Ultimately, what I find is that sanctions do not affect missile tests, and the role of sanctions in international politics still remain a mystery to be unraveled.\textsuperscript{134}

The question is what policy-makers should do if they desire to halt the missile tests of hostile countries. The first answer is that they should figure out means to prevent these hostile countries from developing new missile systems. The present statistical analyses show that imposing economic sanctions does not have an impact in preventing a country from developing new missiles. That said, the analysis in Table 6 shows that the only covariate that consistently decreases the tests of new missiles is the being the primary target of war. Moreover, if a country was the primary target of war, then the new missiles that it tests significantly drops the next year. Noticeably however, war is a costly option.

\textsuperscript{134} See Morgan (2015) for a more in depth discussion on the state of the literature which quantitatively investigates the role and the impact of economic sanctions.
Another potential, but less effective way to decrease a country’s missile tests could be to try to ensure that it is not involved in militarized disputes. If a country is not frequently targeted in militarized interstate disputes, then there is a small chance that is will test-fire fewer missiles. On the other hand, if the goal is to intimidate the country with joint military exercises or missile tests, this does not seem to work according to my findings. For example, if the US test-fires missiles to frighten North Korea and/or Russia, and prevent their missile tests, there is a non-negligible chance that this may be a futile strategy. Similarly, conducting joint military exercises will probably not prevent North Korea from test-firing missiles. This is not to argue that these strategies are in general ineffective. What I rather find in this paper is that they typically do not intimidate their targets and make them test-fire fewer or no missiles.

**References for Paper II:**


Clausewitz, Carl Von (1832) On War


Global Security – M5/M51, web source: https://www.globalsecurity.org/wmd/world/france/m-5.htm


Missile Defense Project, “Iran” Missile Threat, Center for Strategic and International Studies, https://missilethreat.csis.org/country_tax/iran/?lcp_page=2#lcp_instance_0


Schelling, Thomas. C. (1960) *The Strategy of Conflict*
Schelling, T. C. (1966) *Arms and Influence* Yale University Press


The AP International (1998) *Web: Nexis-Uni*


Vuong, Quang H. (1989) “Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses” *Econometrica* 57(2) 307-333


Appendix for Paper II

Figure 24: Missile Test Variables (Frequency Histograms)
Table 7: Summary Statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIDs (t-1)</td>
<td>682</td>
<td>2.513</td>
<td>2.000</td>
<td>2.059</td>
<td>0.000</td>
<td>12.000</td>
</tr>
<tr>
<td>MIDs initiated (t-1)</td>
<td>682</td>
<td>1.229</td>
<td>1.000</td>
<td>1.368</td>
<td>0.000</td>
<td>9.000</td>
</tr>
<tr>
<td>MIDs initiated against (t-1)</td>
<td>682</td>
<td>1.010</td>
<td>1.000</td>
<td>1.214</td>
<td>0.000</td>
<td>7.000</td>
</tr>
<tr>
<td>Joined MID initiator (dummy, t-1)</td>
<td>682</td>
<td>0.103</td>
<td>0.000</td>
<td>0.304</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Joined MID target (dummy, t-1)</td>
<td>682</td>
<td>0.164</td>
<td>0.000</td>
<td>0.371</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Fatal MID (dummy, t-1)</td>
<td>682</td>
<td>0.806</td>
<td>1.000</td>
<td>0.395</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>MID, force against (dummy, t-1)</td>
<td>682</td>
<td>0.749</td>
<td>1.000</td>
<td>0.434</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>War (dummy, t-1)</td>
<td>682</td>
<td>0.120</td>
<td>0.000</td>
<td>0.325</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Primary target in war (dummy, t-1)</td>
<td>682</td>
<td>0.031</td>
<td>0.000</td>
<td>0.173</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Adversaries' Tests (lagged)</td>
<td>732</td>
<td>75.257</td>
<td>33.000</td>
<td>98.729</td>
<td>0.000</td>
<td>557.000</td>
</tr>
<tr>
<td>Adv's Tests w/ &gt;Short Range (lagged)</td>
<td>732</td>
<td>27.270</td>
<td>12.000</td>
<td>40.582</td>
<td>0.000</td>
<td>249.000</td>
</tr>
<tr>
<td>Adv's Tests w/ &gt;Short Range + Nuclear (lagged)</td>
<td>732</td>
<td>20.339</td>
<td>8.000</td>
<td>31.373</td>
<td>0.000</td>
<td>202.000</td>
</tr>
<tr>
<td>Adversaries' JMEs (lagged)</td>
<td>451</td>
<td>0.736</td>
<td>1.000</td>
<td>0.441</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Adv's FTX and CPX (lagged)</td>
<td>451</td>
<td>0.166</td>
<td>0.000</td>
<td>0.373</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Sanction Episode (lagged, dummy)</td>
<td>633</td>
<td>0.488</td>
<td>0.000</td>
<td>0.500</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Severe/Major Sanction (lagged, dummy)</td>
<td>633</td>
<td>0.096</td>
<td>0.000</td>
<td>0.295</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>GDP(logged, t-1)</td>
<td>671</td>
<td>26.698</td>
<td>26.923</td>
<td>1.757</td>
<td>22.128</td>
<td>30.207</td>
</tr>
<tr>
<td>Polity 2 (t-1)</td>
<td>718</td>
<td>1.368</td>
<td>5.000</td>
<td>8.010</td>
<td>-10.000</td>
<td>10.000</td>
</tr>
<tr>
<td>New leader (t-1)</td>
<td>736</td>
<td>0.219</td>
<td>0.000</td>
<td>0.414</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>% Change mil. spend.(t-1)</td>
<td>661</td>
<td>0.111</td>
<td>0.070</td>
<td>0.425</td>
<td>-0.806</td>
<td>8.478</td>
</tr>
<tr>
<td>New missiles (t-1)</td>
<td>726</td>
<td>0.532</td>
<td>0.000</td>
<td>1.040</td>
<td>0.000</td>
<td>8.000</td>
</tr>
<tr>
<td>% Change missile Imp. (t-1)</td>
<td>732</td>
<td>0.303</td>
<td>0.000</td>
<td>1.970</td>
<td>-1.000</td>
<td>23.000</td>
</tr>
<tr>
<td>Cold War (t)</td>
<td>735</td>
<td>0.641</td>
<td>1.000</td>
<td>0.480</td>
<td>0.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
In 1995, China test-launched a series of missiles in Taiwanese Strait without informing Taiwan. This was the beginning of the Taiwanese Strait Crisis. Shortly after the outbreak of the crisis, the US sent an aircraft carrier (USS Nimitz) to the Taiwanese Strait. Next year, this time Iran conducted tests of its new naval-based missiles in the Strait of Hormuz. It was interpreted by news reporters as a warning act against any potential US invasion from the Persian Gulf. The US decided to redeploy USS Nimitz to the Persian Gulf. The Iranian Mullah Regime claimed that the USS Nimitz presence in the Persian Gulf was driven by Americans’ concerns about the recent Iranian naval-based missile tests (BBC 1997). In both cases, pundits and policy experts were concerned about the risk of violent military clashes to come.

More recently, missile tests of some countries again led the world to be alarmed about the possibility of violent conflicts. In 2017, North Korea’s leader Kim Jong-Un ordered a series of missile tests that were of great concern to policy-makers, and pundits around the world. Most of these missiles were technically capable of reaching America’s allies in the Pacific Ocean. The last recorded test in 2017 was of a new ICBM named Hwasong 15. This nuclear-capable missile was larger than any of the known North Korean missiles and used a new engine technology. Potentially, Hwasong 15 could hit targets in Alaska and Western Canada (CNN 2017, Missile Defense Project – Hwasong 15). Amidst political tension with South Korea, Japan, and the US, North Korea’s test of a Hwasong-15 raised important questions. The most dramatic one of these was whether another Korean War was approaching.
In 2018, the North Korean case took a different turn. The US president Donald Trump had a face-to-face meeting with North Korea’s leader Kim Jong-Un. In February 2019, they had another meeting in Hanoi, Vietnam. In the meantime, from its last Hwasong 15 test in 2017 until May 3rd, 2019, North Korea did not test-fire any missiles publicly. In this period, a Korean war was rarely highlighted in political debates. That said, in May 4th, 2019, news sources reported a guided-missile test-launch conducted by North Korea (BBC 2019). These recent tests started to raise eyebrows again: will North Korea reconsider the option of using deadly military force to tackle with its ongoing conflicts?

The cases of China, Iran, and North Korea illustrate that when countries publicly test-launch their missiles, this raises questions about the risk of future militarized conflicts and wars. The United States’ reaction against the Chinese and possibly Iranian public missile tests by dispatching an aircraft carrier suggests that pundits may not be completely wrong about their concerns. However, we still do not know whether missile tests are typically followed by militarized conflicts and/or wars. To my knowledge, no scholarly work has investigated this question in a scientific fashion to produce generalizable knowledge. In other words, whether or not public missile tests pose an actual danger to the international security is an unanswered question. Yet, the answer to this question may have quite important implications for how policy-makers should approach countries like North Korea, Iran, Russia, India, Pakistan or China when they test-fire their missiles. If their missile tests actually increase the risk of their militarized conflict involvement, then preventive measures should be prioritized. If not, then countries may consider shifting their foreign policy resources to solve problems other than missile tests in the world.
In this paper, I develop three hypotheses and present a series of empirical analyses that investigate whether there is an association between countries’ public missile tests and their subsequent involvement in militarized disputes and wars. In doing so, I present a theoretical argument based on Jervis’ (1976) concept of “security dilemma” which leads us to anticipate a positive association between missile tests and subsequent onset of militarized interstate disputes as well as wars.

I test my expectations using a recently-updated version of the MID Dataset (Maoz et al 2019), and new data that I collected on 11 countries’ publicly-conducted missile tests between 1949 and 2010. I conduct a series of multivariate logit, and Heckman Probit models. Ultimately, I fail to find robust statistical support for my empirical expectations. Moreover, even though missile tests have a small and positive effect on subsequent MID onset, this finding is not robust to sample specifications. In addition, I find no consistent pattern that could in any way suggest that missile tests increase the risk of MID escalation to wars. In fact, in a sample of MIDs without the US observations, missile tests almost consistently decrease the risk of war escalation.

Overall, this paper suggests that public missile tests appear not to have a discernible impact on countries’ militarization of their conflicts. Note that the idea of this non-finding was out there, but it was never tested. Moreover, some scholars and pundits already downplayed the risk that missile tests present in international security. Some others claimed the opposite. This paper is the first one to actually empirically highlight the lack of association between public missile tests and militarized conflicts. In addition, it presents a data collection project on public missile tests that the subsequent studies can rely on to empirically address many further policy-relevant issues in international security.
In what follows, I first present my argument and develop three hypotheses focused on missile tests and conflict. Second, I detail the empirical research design and introduce the missile test data I collected for this study. Third, I summarize the statistical results from a series of logistic regressions, Heckman Probit analyses as well as a series of robustness checks which aim to test my hypotheses. Finally, I discuss the implication of the findings.

**Theoretical Argument**

This section is dedicated to presenting a theoretical argument that explains why public missile tests of a country increase the risk of its involvement in subsequent militarized conflicts. In doing so, I focus on how tests of a country can increase insecurity in the rest of the world, and how this insecurity can affect the risk of militarized conflicts. To develop this argument, I rely on terms such as anarchy and security dilemma. Accordingly, I first elaborate on anarchy in the international system and how it connects to the notion of security dilemma among nations. Second, I explain what role missile tests may potentially play in this setting. Finally, I indicate how my argument leads to expect that missile tests can increase the risk of subsequent militarized disputes and wars.

Anarchy in international relations refers to the absence of an overarching institution or global government that can perfectly enforce rules upon sovereign countries. In this setting, countries are the core units of international politics (Gulick 1955, Waltz 1979). A common assumption is that countries’ goal is to maintain their survival in the international system (Levy and Thompson 2010, Mearsheimer 1995, Waltz 1979). In doing so, they rely on their military capabilities. Moreover, each country in the anarchical international system has a certain level of military power that it can exert upon others to be able to maintain its own survival (ibid).
Ensuring survival is not easy when there is no law-enforcing world government in the international system. Under anarchy, nations are not protected by supranational rule enforcers from invasion. Even though countries build international organizations and/or sign treaties to ensure their survival, these measures typically lack the enforcement capabilities that a sovereign government has over its citizens. In fact, Mearsheimer (1995, 2001) argues that international organizations (IOs) serve the security interests of the countries that create it. Hence, under anarchy, there is always a non-negligible risk for a country to be attacked by others. Consequently, countries have to rely on their own military capabilities to survive, and to fend off aggressors (Jervis 1976, Waltz 1979).

How countries ensure their own survival in an anarchical system varies. Scholars like Mearsheimer (i.e. offensive realists), argue that countries can only ensure their own survival through warfare and invasion (Mearsheimer 1995). Japan’s pre-World War II territorial expansion on Southeast Asia, Germany’s invasion campaigns during World War I and the USSR’s invasion of Eastern Europe during World War II appear to be driven by their belief that expansion ensures survival (Jervis 1976: 63-64). Some other scholars, however, argue that countries form alliances to fend off threats to their security (Senese and Vasquez 2005, Vasquez 1987, Waltz 1979). Lastly, one common way countries ensure their security is to engage in military budget expansions and develop deadly weapons - including missiles and nuclear devices (see Jervis 1976, Kahn 1960, Schelling 1960).

As countries pursue the maintenance of their own survival, a typical consequence is the “security dilemma.” When a country takes measures to ensure its own survival, this decreases all the other countries’ level of security in the anarchical international system. According to Jervis, this is because there is always uncertainty about the intentions of a
country that appears to take security measures. On one hand, it could be that these measures are driven by defensive purposes. On the other hand, these seeming security measures may be conducted in preparation for invasion. Due to the anarchical nature of the interstate system, there is no guarantee that the worst case scenario is not going to happen. Hence, facing such uncertainty about the intentions behind a country’s security measures, other countries have to consider the risk of the worst case scenario: a foreign invasion. The greater weight other countries place on the likelihood of the worst case scenario, the more security measures they take (Jervis 1976, 1985).

Ultimately, a country’s security measures tend to be (mis)perceived as a threat by other nations. Then, the latter take security measures of their own given that their goal is to ensure their own survival. Subsequently, the former becomes less secure due to the latter’s reaction and takes even more security measures. These reciprocal measures spiral among the members of the international system. As it does so, countries find it more difficult to back down from arming themselves further. Accordingly, the dilemma is that the initial security measure of one country ultimately makes itself and the rest of the world less secure (Jervis 1976).

**Missiles and Missile Tests**

Intuitively, building up or acquiring missiles can be a security measure that countries make. If we buy into Jervis’ (1976) argument on security dilemma, missile arsenal expansion of one country should incentivize other countries to take protective measures of their own. Indeed, the missile development race between the US and USSR
may corroborate this argument.\footnote{The US was the first country to develop a nuclear weapon. Yet, Soviets did not take long to manage to detonate their own nuclear device which took place in 1949. In 1952, the Soviets developed and launched the first Intercontinental Ballistic Missile (ICBM) in history – the R-7 (Encyclopedia Astronautica – R-7). It took only a couple of years for the US to finally complete their own first ICBM. Both countries continued to develop deadlier and very expensive missiles in a tit-for-tat fashion until the end of the Cold War. Towards the end of the Cold War, the USSR developed its billion dollar-worth multiple warhead-carrying (MIRV) and nuclear-capable Voevoda ICBM as a response to the ‘Star Wars’ anti-ballistic program pushed by the US’ Reagan administration (Encyclopedia Astronautica – R36 M2). Although many experts thought it was not necessary, the Reagan Administration in the 1980s believed that the Soviet Voevod was a considerable threat to the US. President Reagan did not the Minuteman III missiles as a balancing factor against Soviets’ most-recently developed missiles. This is why he spend years to convince the US Senate and Congress to allocate billions to the MX/Peacekeeper program. Ultimately, the missile was built. MX was an ICBM with the capability of carrying a dozen warheads some of which nuclear capable. One Peacekeeper had the capacity to devastate entire cities at once (Encyclopedia Astronautica – Peacekeeper) Hence, as this case illustrates, development of missiles can greatly contribute to security dilemma among nations.} Similarly, Pakistan and India have been experiencing a missile development race of their own.\footnote{See Paper I of this dissertation for more information on the India-Pakistan missile development race.} Yet, there remains the question of whether the public missile tests can be significantly more threatening than acquisition of missiles.

The test-launch of a missile typically involves the same standard operating procedures that countries use in a combat situation. Typically, during a test day, one missile launch squad is mobilized. Members of this squad first conduct all the necessary checks to ensure that the missile is ready to be launched. Then after a certain countdown, the missile of interest -which is typically worth millions of Dollars- is test-launched. If this missile was launched against an adversary, the exact same procedures would be followed. The only difference would be that in the combat situation, the missile would be launched against targets in a hostile country.

When a country increases the number of missiles it publicly test-fires, this could make it difficult for others to answer why these missiles are not remaining idle in the silos. According to Jervis (1978), when countries have trouble understanding the motivations behind others’ actions, these actions start seeming as threatening. Hence, more frequent
public missile tests of one country can lead others to see this country as more threatening. To illustrate, suppose that the North Koreans increase the number of missiles they publicly test-launch in a certain period of time. The Americans may believe that North Korea has no chance to defeat the seemingly much stronger US-South Korea alliance. Then, North Korea should normally have no incentives to provoke the US. Nevertheless, it still somehow test-launches missiles and risks provoking its much stronger adversaries. The trouble to understand reasons behind North Korea’s missile tests may then cause the American policy-makers to see these tests as a threat.

**First Strike Advantages**

The story so far may lead us to predict that increase in public missile tests of a country will simply lead the rest of the world to take more security precautions. Hence, the potential consequence of these tests may be greater military spending or alliance formation. Yet, what is investigated presently is the risk of militarized conflict: is there a chance that missile tests actually increase the risk of militarized conflict? The answer can be yes, and this is based on the additional aspect of missiles: “the first strike incentives”.

The first strike incentives transpire in cases when two opposing sides are both better off being the first one to strike than to be the second. This is nicely illustrated by Schelling (1960) with a gunslinger duel example. In a duel, both gunslingers have incentives to shoot first. The one that fails to shoot first basically loses the duel (i.e. typically dies). Thus, both gunslingers try to avoid being the second one to shoot.

---

137 In fact, the US army reports are in line with the expectation that North Korea cannot defeat a South Korea – US alliance despite its missiles (Scobell and Sanford 2007).
Missiles also can present countries with incentives to strike first. For simplicity, assume that there are two countries in conflict. One is planning to attack the other and has missiles. The attacker can launch missiles and consequently eradicate vital targets in the other (defender) country. These vital targets could be important infrastructure and military bases that the defender would otherwise use to mount a successful defense against the attacker. Consequently, a successful series of missile launches provides the attacker with an opportunity to dramatically weaken the defender without having to fight a major battle with its armed forces.

On the other hand, the defender has incentives to act as quickly as possible to wipe out the attacker’s military leverage gained by missile strikes. Such pre-emptive strike can be against the attacker’s missile-launching vehicles/vessels, its command posts, or the infrastructure that the attacker relies on to mobilize against the defender. If the attacker is deprived from the means to use its missile strike abilities, then it loses the opportunity to fight a war with low expenses in which victory is rather warranted. Hence the defender has incentives to be the first to strike against the attacker. However, the attacker, then, also has incentives to launch its missiles against the target, before the latter succeeds in its pre-emptive strike. Ultimately, who strikes first gains a crucial strategic advantage, and is better off than striking the second. This simplified scenario illustrates how the presence of missiles in at least one of the opposing sides creates incentives in two to be the first one to strike.

The worst case scenario when a country test-fires its missiles is a war in which being the first one to attack is optimal. Under the anticipation of such a war, countries may feel more insecure the more they delay an offensive strategy. The longer the offense is
delayed the more time the missile-tester may have to mount a successful missile strike. Thus, the security measure as a reaction to missile tests can be an attack against the missile testing country. On the other hand, given the risk to be attacked first, the missile-testing country has incentives to be the first to strike so that its ballistic leverage is not destroyed by others. Hence, the security dilemma that is created by public missile tests is more likely to be associated with the subsequent risk of militarized conflicts than mere arms races.

According to the present argument, public missile tests of a country should increase the likelihood that it will be the initiator or the target of a Militarized Interstate Dispute (MID, hereby). Militarized interstate disputes (MIDs) represent a notable change in how countries manage their ongoing disputes. MIDs include threats to use military force, display of military force, and violent uses of force. Although rarely, some MIDs can even escalate to war (Jones et al 1996). The imminence of war during MIDs make these incidents typically comparable to what Brecher (1996) calls international crises. In that regard, the probability of war onset is generally greater during a MID than the during the absence of a MID. Given the present argument, the heightened sense of threat created due to public missile tests may push countries (including the missile testing one) to initiate MIDs as a way to exploit their first strike advantage. Therefore, the more a country test-fires missile, its likelihood of initiating a MID, and also being the primary target of a MID should increase. The first two hypotheses, can accordingly be stated as:

**Hypothesis 1:** Ceteris paribus, the more public missile tests a country conducts, the more likely it is to subsequently initiate a MID against another

**Hypothesis 2:** Ceteris paribus, the more public missile tests a country conducts, the more likely it is to be the main target in a MID.
The remaining question is what happens once countries start to employ their military forces in a hostile manner against each other. What is the risk of these military hostilities escalating to war? If more public missile tests increase the first-strike advantage for missile testers and their potential adversaries, then MIDs between both sides should be more likely to escalate to wars. As long as incentives to attack first persist after MID initiation, countries may find it disadvantageous to do stay their hand. Thus, we can expect that they will use force against each other to exploit their first strike advantage and/or deny the other to do so until one side is completely incapacitated – which typically can be one way wars happen. Hence, in general, MIDs that follow more public missile tests of one of the participants should be more likely to escalate to wars. This hypothesis can be stated as:

**Hypothesis 3:** *Ceteris paribus, a MID is more likely to escalate to war, the more public missiles one of the MID participants previously test-fired*

In sum, an increase in the public missile tests of a country will lead to the security dilemma situation where other countries as well as the test-conducting country feel less secure. However, I expect that the security measures that every country consequently takes will be dire. This is, as I argued, due to the nature of missiles – they create a “shoot first or die” situation for not only the missile testing country, but to the rest of the world. This is the reason why missile tests may be conducive to MIDs and wars, as pundits tend to expect.

**RESEARCH DESIGN**

(1) **MID Onset:**

*Sample:* The first hypothesis focuses on a country’s unilateral decision to initiate a MID, and the second hypothesis focuses on the probability that the country will be targeted by any MID. As Bremer (1992) and Ray (2005) argue, the likelihood of a militarized interstate
conflict depends on the attributes of the both opposing sides of the conflict. This is why I created a directed-dyad sample. The reason why I use the directed structure is to be able to trace which country in the dyad started the MID. The time unit is a quarter-year between 1949 and 2010. The country (country A hereby) that takes an action against the other (country B) is the one for which I collected data on missile tests. In that regard, there are 11 countries recorded as “country A” in the sample. The countries are selected on the basis of being the prominent missile-testing countries in the world – which I elaborate on further when I describe the missile tests variable. These countries are the US, USSR/Russia, China, France, England, North Korea, Iran, Iraq, Israel, Pakistan and India.

For country B, I first recorded every single other country in the world in the given year. In other words, the first sample I created is 11 country A.s versus all possible countries. This sample has 319,983 observations (i.e. directed-dyad quarter years). Second, I created a relevant dyad sample. This sample includes all the 11 countries as country A versus those that have direct land contiguity with these countries (i.e. A’s land neighbors) as country B. In addition, for US, USSR, England and France -given their major power status- I included not only neighbors but also all other possible countries. For this relevant directed dyad sample, the number of observations is 158,153. The logic behind the relevant dyad sample is to trim the excess 0’s (i.e. non-conflict observations) in country pairings where the conflict is unthinkable (Lemke 1995, Lemke and Reed 2001). For example, it is highly unlikely that Pakistan will initiate a MID against Fiji, or North Korea against

---

138 The reason why I did not use year but quarter year as a time unit has less theoretical reasons than empirical/practical reasons. It is an arbitrary time increment I chose to capture the relatively short-term effects of missile tests on MID incidence. Had I picked year as the time unit, this could have caused several problems. Will a missile test have an effect on the likelihood of a MID in one year? This is rather a question about the long-term effects of missile tests. The present study focuses on what happens as a country happens to test-fire missile in a rather short-term period -which I arbitrarily pick as a quarter year.
Belarus. The presence of these pairings may potentially bias the sample in favor of non-conflict observations.\footnote{139}

Outcome Variables: I created two binary indicators to code a country’s MID involvement using information from Maoz et al’s (2019) Dyadic MID Dataset. This dataset is presented in directed-dyad structure which makes it straightforward to record which country initiated a MID against which other country. The first binary variable I created takes the value of 1 if country A was coded as the primary initiator of any MID against country B in the observed quarterly period. In several cases, country A might have initiated multiple MIDs against B in a quarter year. However, this does not happen frequently, and chances are high that these multiple disputes are related.\footnote{140} Overall, of the 319,983 observations, 612 are recoded as a quarter year when country A was the primary initiator of a MID against country B. In the relevant dyad sample, 482 of the 158,153 relevant dyad observations are cases when A initiates a MID against B.

The second outcome variable is recorded as 1 if the country A was the primary target of a MID initiated by country B, and 0 otherwise. 453 quarter-years are coded as 1 if A was the primary target of MID initiated by B according to Maoz et al’s (2019) operationalization. In the relevant dyad sample, this number is reduced to 368. Note that both outcome variables exclude cases where countries A and B are joiners.

\footnote{139} A critic of the relevant dyads approach could be the following: there are reasons why conflict among some countries are almost unthinkable. For example, distance is an important factor. When we remove these so-called politically irrelevant dyads, we do not necessarily introduce the desired level of randomness in the sample. We in fact potentially remove a significant part of the variation in the conflict onset variable that otherwise needs to be explained. It is for this reason that I conduct my analyses in all-dyads and relevant-dyads samples

\footnote{140} Also, this variable is recorded as 0 if the country A was a joiner in the MID.
**Missile Tests:** To capture countries’ missile tests (country A in the present samples), I collected new data on missile test-launches that are publicly conducted by the 11 countries of interest between 1949 and 2010.\(^{141}\) The general reason why I focused on these particular countries is that they are the most missile-testing countries and attract significant global attention. For example, North Korean missile tests are often covered by many American, Japanese, South Korean, and Chinese media outlets. Iran’s missile tests are closely watched by Israel newspapers such as Haaertz and other Turkish and American media outlets. The same is true for all the countries in the sample.

The 11 countries that I collected data on develop their own indigenous missile technologies. Even though the number of missile-related weapons they import are not low (especially in the case of the UK, India, Pakistan, Iraq and Iran), the countries in the sample tend to reverse engineer what they import and then develop their own weapons. For example, the root of the North Korea’s indigenous missile program is based on Soviets’ Scud-B missiles (Missile Defense Project – Hwasong 5, Encyclopedia Astraunomica – Hwasong 5). The North Koreans reverse engineered Scud missiles and created their own arsenal. Likewise, almost all of Iraq’s Saddam Hussein era missiles were based on Soviet Scud technology (Encyclopedia Astraunomica - Iraq). Similarly, Israel imported its first Jericho engine from France. From the first Jericho schematics, Israeli scientists developed at least two more destructive and ultimately nuclear-capable versions of their own (Missile Defense Project- Jericho 3). In fact, nearly all of the American and Soviet missile technologies are primarily based on the V-2 launch vehicle developed in Nazi Germany by Von Braun (Encyclopedia Astraunomica – USA, Encyclopedia Astraunomica – Russia).

---

\(^{141}\) See the subsection where I describe the sample for which countries are included.
The types of weapons I collected data on are short-range ballistic missiles (SRBM), medium-range ballistic missiles (MRBM), intermediate-range ballistic missiles (IRBM), intercontinental-range ballistic missiles (ICBM), short-range cruise missiles (SRCM), medium-range cruise missiles (MRCM), intermediate-range cruise missiles (IRCM), intercontinental-range cruise missiles (ICCM), other types of short range missiles\textsuperscript{142}, nuclear warheads, missile interceptor systems and space launches with military missiles. I ensured that the tests are not electronic launch simulations or isolated booster/engine tests. All the data I collected are events that involve a public attempt to test-fire a missile by the country of interest.\textsuperscript{143}

I used three main sources to collect information on the missiles tested by the countries of interest. The first source is Center for Strategic and International Studies’ (CSIS) Missile Threat Project.\textsuperscript{144} This source presents an overview of the missiles including their service history and key tests dates for all the countries in the dataset except Iraq. I use information from CSIS and code the days when a particular missile was test-fired by a country. If two different missiles (say, America’s Minuteman III and MX) were test-fired the same day, I record two different observations. If the same missile was test-fired more than once in a day, then this is coded as one observation. However, I also record how many times the missile was test-fired that day. The second source I used was Lexis-

\textsuperscript{142} These short range missiles include air-to-air, ship-to-air, or anti-tank missiles that are neither ballistic nor cruise. The French Exocet, the American AGM-142, Brimstone, AIM-9 or Iranian Fajr are examples for these types of missiles. Note that SAMs are not included in this short range category, but in the missile interceptor category.

\textsuperscript{143} A missile that explodes in the launch pad is still recorded as a missile test. After all, this is an attempt that results in utter failure. The test still has the informational value that I highlighted in the previous “Argument” section. Overall, the events of interest were not discriminated based on whether the test was a success or failure.

\textsuperscript{144} See https://missilethreat.csis.org/missile/
Nexis (Nexis-Uni) which is an online international news archive. I relied on Lexis-Nexis primarily to ensure that the information I had from the other sources refer to publicly-conducted tests. There are also many missile tests for each country that CSIS did not cover, but news sources from Lexis-Nexis did. Finally, I relied on Mark Wade’s Encyclopedia Astraunomica. This is an online source containing extensive information on countries’ missile and scientific rocket launches. Wade’s Encyclopedia Astraunomica was especially helpful in coding the Cold War tests conducted by the countries in the dataset. In fact, absent Encyclopedia Astraunomica, the dataset would look like the US and the Soviets did not test-fire any missiles prior to 1980s; which is not the case at all.

Given the sources, I recorded the following information about the missile tested in the given day: (i) Month/Day/Year of the test, (ii) the name of the tested weapon in question, (iii) the aspect/type of the missile, (iv) the commissioning date of the tested missile, (v) whether or not the missile is/was capable of carrying nuclear warheads (vi) whether or not the missile was test-fired for the first time, and (vii) how many times the missile was test-fired in the observed day. Overall, for 11 countries, I recorded 11,257 test days where a specific missile, military rocket or missile interceptor was test-fired. Paper 1 of this dissertation presents the more detailed coding procedure of the data on missile tests.

---

145 I sought to ensure that the missile test of a country is not only reported by its official news agency, but also by foreign news outlets. Admittedly, this still is not the perfect solution to ensure the “publicness” of the missile test since foreign sources often simply quote what a national news source publishes. However, especially in the case of countries like Iran, North Korea, Russia or China, news sources specify whether the test was confirmed by others. Hence, the advantage of Lexis-Nexis is that it does not only display news of a missile test incident, but also news about the confirmation of the missile test in question. The confirmation news article may be published weeks later than the original incident. Thus, I coded every information manually with a thorough search on news pieces generated by Lexis-Nexis.

146 http://www.astronautix.com/

147 For many cases, I was not able to record an exact “day” of the missile test in question. This may simply be due to the lack of information, or the country of interest attempting to blur the exact time of the test. However, for the overwhelming majority of the tests the “year” and the “month” were coded.
In line with the current sample, I collapsed all the information from the missile dataset on quarter-year level for each country. Then I created six variables. The first variable records the total number of missile tests conducted by country A in a quarter-year period. These tests include any types of missiles, missile interceptor systems, military rockets, and nuclear warhead explosions. The second variable I created records the total number of missiles that are test-fired for the first time in the observed quarter year. As a third variable, I recorded the total number of tests launches of already-commissioned missiles. Lastly, I created three variables that respectively recorded the total number of all missile tests, new missile tests, and commissioned missile tests conducted within 6 months. For all the analyses, I lagged each of these six variables by one quarter year. Table I summarizes these variables’ distribution by country.

148 The three hypotheses I developed are based on the association between conducting missile tests and a country’s improving first-strike capabilities. However, if most of these missiles are tested for the first time, chances that they will immediately create a first-strike advantage is low. That said, if the missiles are already commissioned for service, their test-launches may be mostly related to maintenance/life-extension purposes. Then the impact of testing commissioned missiles on improving a country’s missile strike capabilities may be stronger than that of brand new missiles.
Table 8: Summary Stats of Missile Test Variables by Country (N=36,777 by country)

<table>
<thead>
<tr>
<th>Country</th>
<th>All Tests (prev. quarter)</th>
<th>All Tests (prev. 6 months)</th>
<th>New Tests (prev. quarter)</th>
<th>New Tests (prev. 6 months)</th>
<th>Commissioned Mis. Tests (prev. quarter)</th>
<th>Commissioned Mis. Tests (prev. 6 months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>SD</td>
<td>Max</td>
<td>Median</td>
<td>SD</td>
<td>Max</td>
</tr>
<tr>
<td>The US</td>
<td>14</td>
<td>23.228</td>
<td>107</td>
<td>32</td>
<td>45.060</td>
<td>194</td>
</tr>
<tr>
<td>The UK</td>
<td>0</td>
<td>1.886</td>
<td>18</td>
<td>1</td>
<td>2.812</td>
<td>22</td>
</tr>
<tr>
<td>France</td>
<td>0</td>
<td>1.604</td>
<td>7</td>
<td>2</td>
<td>2.524</td>
<td>11</td>
</tr>
<tr>
<td>USSR/Russia</td>
<td>12</td>
<td>14.192</td>
<td>104</td>
<td>23</td>
<td>26.247</td>
<td>163</td>
</tr>
<tr>
<td>Iran</td>
<td>0</td>
<td>1.999</td>
<td>15</td>
<td>0</td>
<td>3.131</td>
<td>18</td>
</tr>
<tr>
<td>Iraq</td>
<td>0</td>
<td>1.875</td>
<td>14</td>
<td>0</td>
<td>3.343</td>
<td>22</td>
</tr>
<tr>
<td>Israel</td>
<td>0</td>
<td>0.679</td>
<td>4</td>
<td>0</td>
<td>1.038</td>
<td>6</td>
</tr>
<tr>
<td>China</td>
<td>0</td>
<td>1.559</td>
<td>12</td>
<td>0</td>
<td>2.651</td>
<td>23</td>
</tr>
<tr>
<td>North Korea</td>
<td>0</td>
<td>1.701</td>
<td>17</td>
<td>0</td>
<td>2.767</td>
<td>23</td>
</tr>
<tr>
<td>India</td>
<td>0</td>
<td>2.094</td>
<td>11</td>
<td>0</td>
<td>3.655</td>
<td>17</td>
</tr>
<tr>
<td>Pakistan</td>
<td>0</td>
<td>1.091</td>
<td>8</td>
<td>0</td>
<td>1.613</td>
<td>8</td>
</tr>
</tbody>
</table>

Note: SD refers to Standard deviation
Control Variables: For each model that I estimated to test Hypotheses 1 and 2, I included 12 control variables. These variables, according to different studies, are suspected to positively or negatively affect tensions among countries and thus the likelihood that they will find themselves in a militarized interstate conflict. The goal is to see if missile tests have an independent impact on the likelihood of A’s MID involvement even when we account for alternative reasons. Nevertheless, I also reported Bivariate statistical results in the Appendix.

(i) Number of Previous MIDs: using information from Maoz et al (2019), I added a variable that recorded the number of previous MIDs that took place prior to the observed quarter year between A and B. Previous research suggests that the more times two countries met each other previously in a MID, the more likely they are to get in a MID against each other again (Colaresi et al 2008). Thus, I expect a positive association between the total number of previous MIDs between A and B, and the probability that these countries initiate a MID against each other in a given quarter year. The maximum number of previous MIDs is 83, the median is 0 and the mean is 0.07. Israel vs Syria dyad in the third quarter of 2007 reached the maximum number of previous MIDs. It must be noted that 319,491 (i.e. 99.66%) observations do not have any previous MID.

(ii) CINC A/CINC B: relying on the most recent version of the National Material Capabilities dataset (Greig et al 2016), I created a continuous variable indicating the relative material military capabilities of the countries between country A and country B. This dataset includes a composite indicator called “CINC” that consists of macro indicators of a country such as population, military personnel, military
spending and energy consumption that can translate into its material capabilities (ibid). I divided the CINC score of Country A in the dyad for the year when the MID is observed by that of Country B. As the relative CINC score increases, this indicates that A appears to be materially stronger than B. The smaller this variable is, then the material strength advantage is to Country B. When the variable approaches 1, this indicates some sort of material capability parity. Some studies suggest that countries with similar material capabilities risk finding themselves in militarized conflicts more often than countries with capability parity (Bremer 1992). Some other scholars theoretically expect, and empirically find the discrepancy in material capabilities is more likely to trigger MIDs (Fearon 1994, 2002, Colaresi et al 2008)

(iii) Joint Democracy (dummy): I created a binary variable indicating whether both countries A and B are democracies. Many studies empirically find that democracies are less likely to initiate a MID against each other (Bremer 1992, Bueno de Mesquita et al 1999, 2003, Maoz and Russet 1993, Russet and Oneal 2001). To measure joint democracy, I relied on the Polity II score from the Polity IV dataset compiled by Marshall et al (2017). The variable I used is coded as 1 if the Polity II scores of both countries are greater than or equal to 6, and 0 if otherwise. Of all the observations, 87,579 (i.e. 24%) are joint democracy dyads in the observed quarter year.

(iv) Contiguity: relying on Hensel’s (2017) Direct Contiguity v.3.20 dataset, I created a variable coded as 1 if countries A and B share a border by land, and 0 otherwise. Empirical results from previous research suggests that contiguity is an
important predictor of MID onset between two nations (Bremer 1992, Ray 2005, Reed and Chiba 2009, Senese 2005). 14,606 directed dyads (i.e. 4%) in the sample are coded as contiguous by land.149

(v) GDPs of A and of B: based on the most recent version of Gleditsch’s (2002) Extended Trade and GDP Dataset, I recorded the GDPs of both countries of the dyad in the dataset. I lagged the variables by one year, and also took their natural logarithm. Many studies highlight different mechanisms through which the economic wealth of nations affect the likelihood of their militarized conflict involvement. On one hand, several studies show that wealth has a pacifying impact on conflict (Bremer 1992, Mousseau et al 2003). On the other hand, some scholars indicate that countries with greater material wealth are more likely to militarize their disputes (Morgan and Palmer 1997, Palmer and Morgan 2006). In any case, I have reason to believe that GDPs of countries A and B will significantly affect their likelihood of MID involvement either positively or negatively.

(vi) % Changes in military spending of A and of B: for both countries A and B in the dataset, I created a variable indicating the percentage change in military spending from year t-2 to t-1 where t is the observed year of the quarter year dyad. I relied on National Material Capabilities Dataset’s most recent version to extract change in military spending (Greig et al 2016). Palmer and Morgan (2006) expect that increases in military spending should positively covary with initiating MIDs. In that regard, in models where I estimate A’s initiation of MID, I only included

---

149 Admittedly, there are different measures of contiguity use in the literature such as contiguity by sea that I did not employ. For simplicity, and because contiguity is not a variable of primary theoretical interest, I do not include alternative contiguity measures in the statistical models.
A’s change in military spending. For models where the likelihood of B initiating a MID is estimated, I controlled for changes in B’s military spending.

(vii) **Sharing Alliance**: I created a binary variable indicating whether countries A and B share any alliance treaty, as designated by Leeds et al’s (2002) Alliance Treaty Obligations Project (ATOP) dataset. Several studies, with various empirical approaches, show that countries that share an alliance treaty are less likely to initiate MIDs against each other (Kinsella and Russett 2002, Reed 2000). Of all the observations, 55,524 (15.10%) share any alliance treaty, which may be a defense, offense or a neutrality treaty. Of the dyads in which A and B share any alliance treaty, only 234 (0.4%) are the cases where country A initiates a MID against B, and 157 (0.03%) are the cases when A is the primary target of B in a MID.

(viii) **New Leader in A and in B**: for both countries in the dyad, I created a binary variable indicating whether or not the country of interest had a leadership change within the last six months prior to the observed directed dyad quarter-year. I used Goemans et al’s (2009) most recent version of the ARCHIGOS Dataset. Luckily, this dataset presents the exact dates when an executive in a country starts her/his tenure, which enabled me to record leadership changes happening in quarter-year or six-month periods. For models where I estimate A’s initiation of MID, I only included change in leadership in country A within six months prior to the observed quarter year (coded as 1 in 15,346 or 4.17% of the observations). For models where the likelihood of B initiating a MID is estimated, I controlled for leadership change in B (coded as 1 in 1,110 or 0.34% of the observations).
Extant literature presents several interesting findings that motivate the inclusion of leadership variables in the present analyses: Wolford (2007) shows that new leaders are more likely to be tested by MIDs or hostile military strategies in general while their skills are still uncertain to foreign adversaries. Gelpi and Grieco (2001) presents a series of empirical analyses that support this idea among democracies. However, empirically, Bak and Palmer (2010) finds the opposite in the case of the US. Regardless of the direction of the previous findings, we have reason to account for changes in the leadership in the countries in the dyad.

(ix) A and B’s trade dependencies: Russett and Oneal (2001) as well as several other studies argue that countries that depend on each other through commerce are less likely to initiate MIDs against each other (Hegre et al 2010, Oneal and Russett 1999). To measure trade dependency, I used Barbieri et al’s (2009) Trade dataset. Country A’s trade dependency is measured as the share of the imports it receives from B over its GDP – and vice versa for B. This is one of the commonly-used measurement approaches in previous studies (see Oneal and Russett 1999). I include both countries’ trade dependency variables in all the models.

(x) A’s and B’s new stronger allies: lastly, I created two variables indicating whether either of the countries sign an alliance treaty involving a stronger power. Morgan and Palmer (2003) show that signing a military alliance treaty with a stronger power increases a country’s resources available to execute policies including MID initiation. In models where A is the primary initiator, I accounted for whether A signed any alliance treaty with a militarily stronger power within the six months prior to the realized dyad (16,858, or 4.17% of the total observations).
In models where A is B’s primary target, I included the binary variable indicating whether or not B signed an alliance treaty with a stronger power within six months prior to the observed time (1,210, or 0.30% of the total observations).\textsuperscript{150}

(ii) MID Escalation to War:

Sample: To test the third hypothesis, I created a sample consisting of MID dyads as units. Like the MID onset analysis, the sample focuses on years between 1949 and 2010. There are in total 1,389 MID dyads (country A vs country B) that I was able to record between these years. If two countries have multiple MIDs in the same quarter year, this is counted as one MID. Chances are high that these MIDs are related to the same dispute/source of conflict. Overall, the US appears in 231, the UK in 85, France in 58, Russia/USSR in 230, Iran in 116, Iraq in 156, Israel in 120, China in 180, North Korea in 70, India in 80 and Pakistan in 63 MIDs against their opponents.

Outcome Variable: Relying on Maoz et al (2019), I created a binary variable which takes the value of 1 if the countries A and B escalate their observed MID to war. Of the 1,389 observations, 61 (4.39%) are coded as Country A and B fighting a war. The limitation of this information is that we cannot tell whether A was the first-attacker whose actions started the war. To address this issue, I included a

\textsuperscript{150} To calculate the alliance variable, I relied on Leeds et al’s (2002) ATOP project which records international military alliance treaties as well as obligations specified in these treaties. I used the treaty-member level version of the dataset and extracted the treaties that the countries of interest are recorded to have signed at any time between 1949 and 2010. Next, for all these treaties, I recorded each member’s material military capabilities relying on National Material Capabilities’ Dataset’s composite CINC score (Greig et al 2016, Singer et al 1972). Next, I recorded if the treaty had members with CINC scores strictly greater than that of the country of interest. Ultimately, I created a binary variable which takes the value of 1 if the country of interest signed any treaty with a stronger power within the last six months before the observed quarter year in the sample.
number of binary variables that explain A’s role in the MID – which I detail in the “control variables” subsection.

**Missile Tests:** Using the data I collected on missile tests, I created six different variables. The first variable captures the total number of missile test launches within the last quarter year before the observed MID. The global maximum number of missile tests for this variable is 107, the median is 1 and the mean is 8.80. The second variable records the total number of a country’s missile tests in the last 6 months prior to the observed MID. For this variable, the global maximum number of missile tests is 185, the median is 1 and the mean is 18.14. These two variables are the core treatment variables that I use to assess Hypothesis 3.

The remainder of the variables specify subsets of the main treatment variable based on whether the tested missiles were already commissioned or were tested for the first time. It is possible that a missile test-fired for the first time will not have as much contribution to the first-strike incentives of a nation than the tests of already-commissioned missiles that are available for service. Hence support for the third hypothesis may be clearer if we specify whether the tested missiles are commissioned or not. In that regard, I created two variables that calculate the total number of (i) new missiles tested for the first time; (ii) already-commissioned missiles within the last quarter-year. I also recorded the total number of these missiles within the six months prior to the observed MID dyad.

**Controls:** Part of the variables I use for the MID onset analyses are also included in the MID escalation analyses. Accordingly, first, I control for the number of previous MIDs between the countries in the observed MID dyad. Theoretically, the
more times two countries met each other previously, the more likely they may be informed about their relative bargaining positions in their observed MID. Thus, I expect a negative association between the total number of previous MIDs between the countries A and B and the probability of escalation to war. The maximum number of previous MIDs in this sample is 83, the median is 10 and the mean is 17.42.

Second, I again control for whether A and B share any border by land. Of the 1,389 MID dyads, 529 (i.e. 38.08%) share a border, and 860 do not. In other words, the majority of the countries in the MID dyads do not directly share a border. There is reason to believe that contiguity is associated with escalation. For example, Braithwaite and Lemke (2011) try different measurements of conflict/crisis escalation. The contiguity is the only variable that robustly increases the risk of escalation regardless of how escalation is measured. On the other hand, theoretically, a possibility is that contiguity is more related to the onset of conflict than conflict escalation (Blainey 1973, Bremer 1992). I nevertheless add the contiguity variable in line with more recent empirical findings.

Third, I account for the relative material capabilities between A and B, using their relative CINC score again (Greig et al 2016). Some studies suggest that parity in relative military capabilities may create greater uncertainty in relative bargaining power among countries in crisis and increase the risk of escalation (Blainey 1973, Bremer 1992). However, other studies hint that if two countries with material capability disparity are already in a crisis-like situation, this may increase the chance of escalation. The reason is that the weaker country must be over-estimating
its own capabilities, and/or the stronger one is under-estimating the weaker one’s capabilities (Morgan 1994). In addition, most recent wars we observe happen to be between countries with seemingly large material capability disparities (Colaresi et al 2008).

Fourth, I include the joint democracy indicator to the MID escalation models. Several studies suggested and found that this variable may have an impact on crisis escalation (Fearon 1994, Mousseau 1997, Schultz 2000). More specifically, if both countries in the dyad are advanced democracies, the expectation is that they are less likely to end up in a war and more likely to solve their crises without diplomacy or actions short of war.

Fifth, I control for the GDPs of the countries A and B in the observed MID dyad. Greater economic wealth has been found in various studies to be positively associated with countries’ decision to use force (Palmer and Morgan 2006). This could be because greater wealth makes it cheaper for countries to use force and even afford fighting a war. Therefore, independent of their relative material military capabilities, countries’ decision to escalate their ongoing MID to war may be affected by their levels of GDP. For the present analysis, I took the natural logarithm of the both countries’ GDP and lagged it by one year.

Lastly, I created three different variables to specify the role of the missile testing country in the overall MID. Recall that the main outcome variable (MID escalation to war) does not specify whether the missile-testing country made the first move that initiated the war. I sought to remedy this issue by accounting for Country A’s role in the MID. Thus, I first created a dummy variable indicating
whether or not A is the primary initiator in the MID. The variable takes the value
of 1 if A is actually the primary initiator, and 0 otherwise. In 693 of the observations
(49.89%), A is the primary initiator. Next, I created another dummy variable that
takes the value of 1 if A was the primary target and 0 otherwise (527 observations,
37.94% of the total). Lastly, I included a binary variable that is coded as 1 if A
joined the MID on the side of the primary target and 0 otherwise (108 observations,
7.78% of the total). Once all these three dummy variables are included in the
relevant statistical models, the remaining reference category (i.e. when all these
three variables are 0) is that A is the joiner on behalf of the primary initiator.

Summary statistics for all the variables for both samples are presented in
Tables 13 and 14 in the Appendix for this paper.

STATISTICAL ANALYSES
(i) MID Onset Results:

In the statistical models to test the first two hypotheses, I use logistic
regressions with time polynomials. I also cluster the standard errors by directed
dyads. The time polynomials measure (i) the quarter-years since the country A last
initiated a MID against B for the models testing the Hypothesis 1, and (ii) quarter
years since A previously was the primary target in a MID initiated by B for the
models testing the Hypothesis 2. Then I calculated the squared and the cubed
versions of these time variables. This approach is one way to account for the
duration dependency in binary time-series cross-section models (Carter and
Additionally, it is possible to simulate values for the time polynomials in the post-estimation phase and see the substantive impacts of the covariates of interest across time – which I present.\textsuperscript{151} TABLE 9 displays the results for the logistic models that investigate the Hypothesis 1.

Overall, results in Table 9 are not supportive of Hypothesis 1. It appears that missile tests of a country do not consistently increase the chances that it initiates a MID against another. Models 1, 2 and 3 are ran on all the dyads. Model 1 shows that the number of public missile tests conducted within a quarter year prior to the observed time has a typically positive effect that is not statistically significant within 95 percent confidence margins ($\beta = 0.006$, t-value: 1.94). Model 2 of Table 9 considers tests of missiles launched for the first time. I still fail to find a statistically significant association between missile tests and MID initiation by A. Model 3 investigates the impact of commissioned missile tests. The results show that tests of commissioned do not have a statistically significant association with the outcome of interest. Models 4, 5, and 6 show the results of the first three models in the relevant dyads sample. The results do not change. There seems to be no consistent association between missile tests of A within the last quarter of the observed time, and the probability that the country A initiates a MID against B.\textsuperscript{152}

\textsuperscript{151} The estimated model for the first hypothesis is: \( \text{Prob}(Y=1) = \frac{e^{X\beta}}{1+e^{X\beta}} \) where \( X\beta = \alpha + \beta_1\text{A’s missile tests} + \beta_2\text{Number of Previous MIDs} + \beta_3\text{CINC A/CINC B} + \beta_4\text{Joint Democracy (dummy)} + \beta_5\text{Contiguity (dummy)} + \beta_6\text{A’s GDP (1 year logged, and lagged)} + \beta_7\text{B’s GDP (1 year logged, and lagged)} + \beta_8\text{Percent change in A’s military spending} + \beta_9\text{A and B shares an alliance (dummy)} + \beta_{10}\text{New Leader in A (last 6 months)} + \beta_{11}\text{A’s Trade Dependency} + \beta_{12}\text{B’s Trade Dependency} + \beta_{13}\text{A’s New Stronger Ally (dummy)} + f(\text{time, } \delta) + \alpha \) such that \( f(\text{time, } \delta) = \delta_1\text{ time} + \delta_2\text{ time}^2 + \delta_3\text{ time}^3 \) and time= quarter years since the last time A initiated a MID against B, and \( \alpha = \text{Intercept} \).

\textsuperscript{152} In the Appendix TABLE 18, I account for missile tests conducted within the last 6 months prior to the quarter year of observation. The results are slightly different. Moreover, A’s missile tests...
Table 9: Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 1)

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: A Initiates a MID against B</th>
<th>All Dyads</th>
<th>Relevant Dyads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>A's Missile Tests (prev quarter)</td>
<td>0.006</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>A's Tests of New Missiles (prev quarter)</td>
<td>0.011</td>
<td>0.054</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
<td>(0.091)</td>
<td></td>
</tr>
<tr>
<td>A's Tests of Commissioned Missiles (prev quarter)</td>
<td>0.009</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>Number of Previous MIDs</td>
<td>0.126*** (0.006)</td>
<td>0.126*** (0.006)</td>
<td>0.119*** (0.006)</td>
</tr>
<tr>
<td>CINC A/CINC B</td>
<td>-0.0005 (-0.0005)</td>
<td>-0.0004 (-0.0004)</td>
<td>-0.0004 (-0.0004)</td>
</tr>
<tr>
<td>Joint Democracy (dummy)</td>
<td>-0.726*** (-0.152)</td>
<td>-0.737*** (-0.152)</td>
<td>-0.726*** (-0.152)</td>
</tr>
<tr>
<td>Contiguity (dummy)</td>
<td>1.540*** (0.132)</td>
<td>1.502*** (0.132)</td>
<td>1.254*** (0.132)</td>
</tr>
<tr>
<td>A's GDP (1 yr lagged, logged)</td>
<td>0.140*** (0.018)</td>
<td>0.162*** (0.013)</td>
<td>0.147*** (0.013)</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>B's GDP (1 yr lagged, logged)</td>
<td>0.201*** (0.014)</td>
<td>0.203*** (0.014)</td>
<td>0.202*** (0.014)</td>
</tr>
<tr>
<td>% Change in Mil. Spending, A</td>
<td>-0.540** (0.195)</td>
<td>-0.542** (0.193)</td>
<td>-0.534** (0.193)</td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
<td>(0.143)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Share alliance (ATOP)</td>
<td>-0.232 (-0.212)</td>
<td>-0.215 (-0.212)</td>
<td>-0.221 (-0.212)</td>
</tr>
<tr>
<td>New Leader in A (prev. 6 months)</td>
<td>0.103</td>
<td>0.110</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>(0.212)</td>
<td>(0.212)</td>
<td>(0.212)</td>
</tr>
<tr>
<td>A's Trade Depend.</td>
<td>15.648 (13.968)</td>
<td>11.452 (14.113)</td>
<td>14.101 (14.072)</td>
</tr>
<tr>
<td>B's Trade Depend.</td>
<td>1.860*** (0.234)</td>
<td>1.881*** (0.234)</td>
<td>1.874*** (0.234)</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.231)</td>
<td>(0.234)</td>
</tr>
<tr>
<td>A's New Stonger Ally (prev. 6 months)</td>
<td>0.456</td>
<td>0.464</td>
<td>0.461</td>
</tr>
<tr>
<td></td>
<td>(0.295)</td>
<td>(0.295)</td>
<td>(0.295)</td>
</tr>
<tr>
<td></td>
<td>(0.791)</td>
<td>(0.788)</td>
<td>(0.802)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Polynomials</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>367,701</td>
<td>367,701</td>
<td>367,701</td>
<td>176,919</td>
<td>176,919</td>
<td>176,919</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-3,025,400</td>
<td>-3,028,025</td>
<td>-3,027,164</td>
<td>-3,330,665</td>
<td>-2,332,258</td>
<td>-2,331,824</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
<td>6,086,080</td>
<td>6,090,050</td>
<td>6,088,329</td>
<td>4,695,331</td>
<td>4,698,516</td>
<td>4,697,648</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001

Some of the control variables in Table 9 across all the estimated models show similar patterns regardless of what type of missile tests are considered as the main treatment variable. First, regardless of any sample or covariate specification I

within the last 6 months has a statistically significant and positive effect in the likelihood that it initiates a MID against B. I find the same finding in the relevant dyad sample. That being said, this positive effect is highly small and very close to zero.
find a statistically significant positive relationship between A’s MID initiation and the number of previous MIDs between A and B – which is in line with findings from the rivalry literature (see Colaresi et al 2008). Second, direct contiguity between A and B statistically significantly increases the chances that country A initiates a MID against B. Third, in all the six models of Table 9, the greater trade dependency B has on A, the more likely A is to attack B. Nevertheless, A’s trade dependency on B does not have a robust impact on its MID initiation. Fourth, A and B’s GDPs have a consistent and positive impact on the risk of A initiating a MID against B. Although positive association between A’s GDP and its MID initiation is intuitive and line with previous literature (Morgan and Palmer 1997, Palmer and Morgan 2006), the same cannot be argued for why greater GDP of B increases the probability that A attacks B. This finding may depend on the sample where there are only 11 specific countries coded as country A. Fifth, interestingly, increase in A’s military spending has a statistically significant and negative impact on the likelihood that it initiates a MID against B. Sixth, as expected, if both A and B are joint democracies, A is significantly less likely to attack B. On the other hand, the remaining control variables either do not have statistically significant effects, or their effects vary by sample or covariate specification.

153 I still include the contiguity variable to the relevant dyad sample since there are major powers such as the US, USSR, Britain and France that have their neighbors as well as distant adversaries. The variable naturally shows variance in these countries’ cases.

154 For example, the impact of sharing an alliance is greater than zero, yet statistically significant in only relevant dyad models. Also, change in A’s leadership, and A signing a treaty with a stronger power have positive effects on the outcome of interest, whereas these effects are not statistically significant.
TABLE 10 displays the estimates from the 6 models that investigate Hypothesis 2. Generally-speaking, there is some support for Hypothesis 2. The expectation here is that the more missiles A test-fires, the more likely it is to be the primary target of B in a MID. Model 1 of TABLE 10 investigates the impact of all missile tests within the previous quarter year on B initiating a MID against A in the full sample. The estimated coefficient for the missile test is positive, and statistically significant within the 95 percent confidence margins (β =0.013, t-value: 3.25). A similar estimate is found if the relevant dyad sample is used in Model 4 of TABLE 10. On the other hand, there is no statistically significant association between new or commissioned missile tests of A in the last quarter year, and the probability that A is the primary target of a MID initiated by B. In fact, the impact of new missile tests is typically negative, but also has values greater than zero within the 95 percent confidence margin in Models 2 and 5.
Although all missile tests seem to have a statistically significant and positive effect on the likelihood of A being targeted by B in a MID, the magnitude of the coefficient hints that this effect is very weak.\textsuperscript{155} To better interpret this
impact, I calculated the predicted probabilities of A being the primary target of a MID initiated by B across different values of the “all missile tests” variable. Figure 25 Displays the results.

**Figure 25: Substantive Effect Analyses for Model 1 of TABLE 10**

To produce the left-plot in Figure 25, I calculated the average-over-the sample probabilities across different simulated values of the missile tests variable using Model 1 of TABLE 10. This method was used elsewhere and presents an opportunity to make predictions strictly based on the values that the variables in the sample take. This way, we can avoid arbitrarily holding the values of the variables at their arbitrary moments (Chiba et al 2015, Dutch and Stevenson 2008). More specifically, I calculated the predicted probability of outcome for each observation the marginal value of missile tests very small. Nevertheless, one also should recall that the overwhelming portion of the sample takes the value of 0 for the outcome variable.
in the sample, using the estimated β.s in TABLE 10, Model 1. I repeated this operation using 11 different simulated values for the “All Missile tests” variable ranging from 0 to 100 (with increments of 10). Ultimately, I obtained 11 vectors of N predicted probabilities where N is the total number of observations. For each vector, I recorded the mean of the predicted probabilities. I also took the 97.5th and 2.5th percentile values of these predicted probabilities to see the 95% margin surrounding the average over the sample predictions.156 As the left plot displays the results, the impact of missile tests on the probability that A is the primary target is very weak, and almost trivial.

The right plot in Figure 25 displays whether the substantive impact of the missile tests vary based on the simulated values of the time polynomials (i.e. quarter years since A was the primary target of a MID initiated by B).157 Accordingly, I repeated three average-over the sample analyses. First, I fixed the value of missile tests to 0, and simulated values of time polynomials between 0 and 251 (with 1 quarter year increments). Then I extracted the average probability over the sample. Second, I repeated this procedure by setting the value of missile tests to 4.158 Lastly, I again repeated the same procedure by setting the value of missile tests to 107 – which is the maximum. The left plot in Figure 25 shows that the impact of missile tests on the probability that A is the primary target is very weak, and almost trivial.

156 The gray shaded lines in the left plot of Figure 1 display the area between these margins. These values should not be confused with 95 percent confidence margins. I simply sought to display an array of predicted probabilities obtained from the densest part of the distribution of these probabilities.
157 Carter and Signorino (2010) suggest that this analysis is beneficial to make better sense of how the variables of interest impact the binary outcome variable across different time polynomials. It is possible that what we find as a strong impact may actually be weaker compared to had we simulated the time polynomials to other values. To see if that is that case, I chose to display how the impact of missile tests vary across time.
158 4.44 is the average number of missile tests, and 0 is the median. Yet, 4.4 cannot be used as a simulated value, thus I instead used 4.
tests sharply declines from time 0 until time 100 (i.e. time since last MID). From
time 100 and on, the impact of missile tests remains more or less constant.
Moreover, the more recently A was attacked by B in a MID, the stronger the
positive impact of A’s missile tests is on the likelihood that it will be attacked by B
again. In addition, there is no substantial difference between predicted probabilities
when missile tests are set to 0 (plain black line) and to 4 (red dashed line). Yet, if
we set missile tests to its maximum value, then the predicted probabilities are higher
compared to 0 and 4 missile tests. That said, these differences between the predicted
values slightly shrink as time since previous MID increases (from 100 and on).

The time polynomial analysis also presents a substantial detail about the
impact of missile tests compared to those of alternative factors. If the impact of
missile tests is highly sensitive to how recent a MID transpired between A and B,
then what we may be capturing is not only the impact of missile tests \textit{per se}, but
also the fact that A and B recently fought against each other in a MID. Yet, in the
models with limited dependent variables, there is an inherent interaction between
the covariates. Then, what we may be observing as missile tests’ dependence on
recent MIDs may also be simply a reflection of the logistic specification.\(^{159}\)

Last but not least, when we turn to the control variables for TABLE 10, we
observe several differences compared to TABLE 9 results. First, the relative CINC
score appears to have a statistically significantly negative impact on the likelihood

\(^{159}\) Berry et al (2009) suggest that CART models are useful to investigate these interactions.
Nevertheless, when I used a CART model, the input did not involve missile tests. In other words,
the role of missile tests was so trivial that the CART model did not consider is as a meaningful
variable to divide the sample. Thus, I do not show the results in this paper.
that A is attacked by B in a MID in all the model specifications. This means that as the balance of military capabilities favor B, it is more likely to initiate a MID against A – which is not a strikingly counter-intuitive finding. Second, I find that joint democracy dummy is negative, but this time statistically insignificant in all the models. Third, only B’s GDP has a consistently positive and statistically significant impact on the outcome variable for TABLE 10 across all the models. Yet, the impact of A’s GDP is not statistically significant, and its direction of impact varies across different models. Fourth, it seems that B’s increased military spending is positively associated with its likelihood of attacking B – which is indirectly in line with Palmer and Morgan (2006). Fifth, it also appears that if B signs a military alliance treaty with a stronger power, it is more likely to attack A – which is consistent across all variable and sample specifications of TABLE 10. Sixth, any military alliance between A and B appear to reduce the risk that B attacks A. Seventh, if there is a leadership change in B within the six months prior to the observed quarter year, it is more likely to initiate a MID against A. Eighth, greater trade dependency of B to A increases the chances of the former attacking the latter – which is not in line with Russett and Oneal (2001). Finally, the impacts of contiguity and the number of previous MIDs are similar to what we find in TABLE 9.

(ii) MID Escalation:

In order to evaluate Hypothesis 3, I ran logistic models with standard errors clustered by directed country dyads. Given that the sample is a cross section of MID
dyads, I did not have to control for any serial dependence among observations. Table 11 presents the estimates from six different logit models that investigate Hypothesis 3.

In general, it is safe to suggest that there is no statistical support for Hypothesis 3 according to the analyses presented in TABLE 11. Models 1 and 2 investigate the impact of missile tests within the 1 and 2 quarter years prior to the realized MID dyad in the sample.160 Contrary to the expectations, missile tests typically have a negative, yet statistically insignificant impact in both models. The impact of new missile tests, investigated in Models 3 and 4 vary on the time interval: new missiles within one quarter year prior to the realized MID has a statistically insignificant positive impact, whereas new missiles within two quarter years has a statistically insignificant negative effect. Lastly, the impact of the commissioned missiles -investigated in Models 5 and 6- is negative and statistically insignificant. Overall, regardless of the type of missiles and the time interval in which we count the number of missiles, it seems that there is no statistically significant association between missile tests and MID Escalation to war.

160 For space reasons, I added the Missile tests within the last 2-quarter year analyses for MID onset models to the Appendix.
I find the following for control variables in models 1 through 6 in TABLE 11: first, number of previous MIDs, contiguity, and the relative CINC score between A and B appear not to have any consistent and statistically significant effect on the likelihood that they escalate their MIDs to war. Second, joint
democracy – quite unexpectedly- has a statistically significant and positive impact on the likelihood that A and B find themselves in a war. This finding is potentially dependent on the sample. There are only 11 different countries as country A, and one of these is the US. The US is the country that involves in MIDs the most, and often fights wars (see Palmer et al 2015). Lastly, I find that both countries’ GDP have a statistically significant and negative effect on the risk of MID escalation to war which is in line with what Mousseau et al (2003) find.

(iii) Selection Problem:

The sample of MIDs is potentially not a random one. Moreover, countries typically do not randomly appear in MIDs, but there are various reasons -as found in Tables 9 and 10- why countries find themselves in a MID. Therefore, if the unit of analysis is a MID dyad, this is not a random but a censored sample. The censored part is the dyads which did not end up in MIDs. Not accounting for this censor problem can result in biased results (Reed 2000). To address this problem, I conducted a series of Heckman Probit Models. TABLE 12 displays the results. I used all the dyads for this analysis. The relevant dyad sample does not present different results (yet not reported in here).
The Heckman Probit model comprises 2 steps. The first is the selection model which estimates the probability that A and B finds themselves in a MID using probit specification. The variables included in this model in addition to Missile Tests are A and B’s GDPs, trade dependency, relative CINC score, joint democracy, contiguity dummy, alliance sharing dummy, previous MIDs, percentage change in B’s military spending, leadership change in B, and whether or not B signs a new alliance treaty with a stronger power. Even though I sought to include almost all the variables from TABLES 9 and 10, I selectively removed

---

161 Note that, there is no direction here. I do not estimate whether A initiates the MID or B.
several statistically insignificant ones from the selection part.162 Otherwise, my analyses showed that the selection model was a completely different model from the outcome variable.

The second stage of the Heckman Probit is the outcome model. In addition to Missile Tests variable, I included A and B’s GDPs trade dependency, relative CINC score, joint democracy and the contiguity dummy. Ultimately, Model 1 of TABLE 12 investigates the impact of all missile tests within a quarter year, Model 2 does so for the missiles tested for the first time, and Model 3 does so for only the commissioned missiles. As the ρ estimates indicate, the outcome and the selection models are statistically significantly correlated in each model in TABLE 12 – which alludes to the inherent selection problem for the MID sample.

Overall, despite the Heckman Probit specification, the findings for the covariates of interest remain more or less the same. As Model 1 of TABLE 12 indicates, given that A and B found themselves in a MID, there is no statistically significant association between this MID escalating to war and A’s missile tests. This non-finding persists if we account for different subsets of missiles (i.e. new missiles in Model 2, and commissioned missiles in Model 3). Additionally, the impact parameter of missile tests is typically smaller than zero within the 95% confidence margins for Models 1 and 3. This effect, however, is greater than zero if we account for the tests of new missiles in Model 2. Yet, this positive effect is not statistically significant. At the same time, we observe that missile tests in all the

162 These are percentage change in A’s military spending, leadership change in A, whether or not A signs a new alliance treaty with a stronger power. Notably, the variables included in the selection model are the ones that drive the risk that B initiates a MID against A.
selection models have a statistically significant and positive impact on the likelihood that A and B find themselves in a MID. Nevertheless, the effect of missile tests is discernibly small for all and commissioned missile tests.

**Robustness Checks for Paper III:**

Regarding Model 1 of TABLE 9, Model 1 of TABLE 10, and Model 1 of TABLE 11, I conducted four types of robustness checks to assess how the models fit as well as to see if the results depend on modeling, covariate or sampling choices. I also employed an alternative measurement approach and re-tested the hypotheses in a different sample using International Crises Behavior (ICB) dataset (Brecher et al 2017). The results are summarized in this section, but detailed and displayed in the Appendix section.

(i) **Rare events Logit:** Since both MID onset and MID escalation to war are rare occurrences, I estimated rare events logit specification with the same covariates from the first models of TABLES 9, 10 and 11. Rare events logit is one way to correct the bias in the sample when the events of interest are rarely realized (King and Zeng 2001). Accordingly, I used Choirat et al’s (2019) “zelig” package in R which present a built-in function to estimate rare events logit models. The results are presented in TABLE 20 of the Appendix. The estimates remain largely the same with this specification compared to logit models of TABLEs 9, 10 and 11. The main difference is that more frequent missile tests of A has a statistically significant and positive -yet weak- impact on the likelihood that it initiates a MID against B.

(ii) **Heatmap Analysis:** I assessed how well Models 1 of TABLE 9, 10 and 11 fit the sample which they are estimated on. Even though the purpose of this paper is
not to make a forecast of countries’ MID involvement or MID escalation, it is still crucial to make inferences using models that make reasonable predictions. To evaluate whether this was the case, I employed the diagnostic presented by Esarey and Pierce (2012). This diagnostic is run in R with the package “heatmapFit” (*ibid*). This package calculates the empirical probabilities of the binary outcomes of interest in the sample. Then, it generates a plot which compares these empirical probabilities with the predicted probabilities from the estimated models.\(^{163}\) The plots are presented in the Appendix. The results in general show that the Models made reasonable predictions at lower probabilities of MID onset (i.e. Models 1 of TABLE 9 and 10). Yet, given the lack of adequate observations, higher probabilities of MID are either overpredicted or underpredicted. On the other hand, Model 1 of TABLE 11 appears to have accurately predicted most of the MID escalation cases.

(iii) **Country sensitivity:** I conducted a simple sensitivity analysis to see if the estimates from Model 1 in Table 9, Table 10 and Table 11 are driven by observations belonging to a particular country. The analysis is detailed in the Appendix. For Model 1 of TABLE 9, I find that absent Russia/USSR observations, the impact of A conducting missile tests on the likelihood that it initiates a MID against B becomes positive and statistically significant. For Model 1 of TABLE 10, I find that the statistical significance is dependent on country specifications. Moreover, the impact of conducting missile tests on being targeted by another in a MID becomes statistically insignificant once we remove the US-based dyads.

\(^{163}\) See Esarey and Pierce (2012) for a more detailed explanation of how the plot is generated.
Lastly, the non-finding in Model 1 of TABLE 11 is not sensitive to any country specification. There is still no statistically significant association between missile tests and MID escalation.

**(iv) International Crises:** In this case, I used a different outcome variable – the country’s involvement in international crises. Some of the MIDs a country becomes involved in can be trivial military encounters such as fishery disputes or border violations/clashes. These low intensity incidents may not be actively executed or managed by the high-level decision-makers who also plan missile tests. As an alternative to MID-based outcome variables, relying on Brecher et al’s (2017) International Crisis Behavior (ICB) Project Dataset, I created a country-quarter year sample.\(^\text{164}\) The outcomes of interest are crisis onset and crisis escalation to war.\(^\text{165}\) For both crisis onset and escalation models, I ran probit regressions. For the escalation model, I added the Inverse Mills Ratio obtained from the crisis onset models to account for the selection problem -as countries do not randomly appear in crises.\(^\text{166}\) The results do not substantially differ from the main findings I obtained using the dyadic MID dataset (Maoz et al 2019). Moreover, there is a weak but positive association between a country’s likelihood of getting involved in a crisis, and its missile tests. Given that it involves in a crisis, there is a negative association between missile tests and crisis escalation. The only difference is that the negative

---

\(^{164}\) I did not use a dyadic sample because the only possible dyadic version of ICB is presented through EUGene Software (Bennet and Stam 2000). This version’s year coverage is until the 2001 which reduces the variance for the outcomes of interest drastically. Thus, I chose the unit of interest as a country. The results of this ICB-based analyses are not a direct test of the hypotheses. Nevertheless, the findings are informative about the possible association between missile tests and crises.

\(^{165}\) See Appendix for how the variables were measured and their frequencies in the sample.

\(^{166}\) See Appendix for technical details.
association between crisis escalation and missile tests is statistically significant. – which still suggests that there is no support for the Hypothesis 3.

**Discussion**

Should policy-makers and security analysts worry when countries test-fire missiles in public? Will these countries be more militaristic when managing their international problems? Will they eventually cause war? This paper presented good news for those who are curious about these questions. These results are obtained through a large data collection project on 11 countries’ missile tests and a series of multivariate statistical analyses. Hence, to my knowledge, this is the first time we have empirically-driven implications about the consequences of missile tests.

Overall, I fail to find any robust support for the three Hypotheses that I empirically investigated using multivariate statistical models. First, test firing missiles has directionally no consistent impact on the likelihood that the missile-testing country initiates a MID. The positive association that I expected in Hypothesis 1 can only be observed if we remove Russian-based observations in the sample. Nevertheless, this association is weaker than many other factors that significantly increase the likelihood of the outcome of interest.

Second, there appears to be some support for Hypothesis 2 such that the more missiles a country test-fires, the more likely it is to be the primary target of a MID. However, these results are highly sensitive to the US observations— as found in the robustness checks analyses. If the US observations are removed, I fail to find support for the second hypothesis.
Third, I find absolutely no support for missile tests and MID escalation. The statistical analyses I presented indicate that recent missile tests of a country typically have a negative effect on escalation. Yet, this effect can also be positive – which means that the results are statistically insignificant. Interestingly however, when I considered alternative sample and measurement approach using the ICB dataset (Brecher et al 2017), I found a statistically significant and negative association between missile tests and international crisis escalation. Thus, not only is there no support for the Hypothesis 3, but also it is possible that a story contrary to the “first-strike advantage” is going on.

There are three main questions that ought to be answered given the findings in the present paper. First of all, given weak or no support for the hypotheses, are the pundits and politicians wrong when they worry about missile tests? Is the US overreacting to North Korea’s missile tests? If the American experts expect that North Korea is going to wage wars exclusively because of a few missile tests, this is unwarranted given the present findings. There many reasons why North Korea would wage war against South Korea, and when we control for these factors, it appears that the missile tests of North Korea have basically no discerning impact on its MID and war involvement.

The advantage of the multivariate statistical analyses is to compare missile tests against competing explanations for why countries get involved in MIDs and escalate their MIDs to wars. Accordingly, alternative factors such as the democracy, territorial contiguity, economic wealth, leadership changes or previous
MID encounters are more reliable predictors of MID onset or MID escalation than missile tests.

Second, what does the weak relationship between missile tests and MIDs imply given the argument I developed? Above all, my findings do not cast doubt on Jervis’ (1976) argument. There can actually be some level of insecurity that public missile tests create in the international relations. After all, there are more than a handful incidents in which one or more countries protest the public missile tests of another. This shows that public missile tests are typically unwelcomed acts. However, it may be that the missile tests are not threatening enough to create a “shoot first or die” situation which leads countries to get involved in MIDs.

Third, how do we interpret the statistically-significant results in TABLE 10? It would be hard to claim that there is no association whatsoever between missile tests and MID onset. I find limited support for the second hypothesis in TABLE 10 (Models 1 and 4). A possibility is that the weak association we observe is because of the reaction that countries get when they test missiles. However, if we take a closer look at MID incident narratives (Palmer et al 2015), we observe that this is not the case. In other words, countries do not typically initiate MIDs because one of their adversaries test-fired a missile. The only exception that the MID dataset records is the Chinese missile tests in the Taiwanese Strait Crisis (Dispute3 number: 4088).167

167 “Dispute3” is the name of the variable which records the unique number that the MID data collectors assign to each recorded dispute.
The narratives of the MID dataset however hint to an alternative story: there are 4 cases in which a country (typically South Korea or Japan) initiated a MID against another (typically North Korea), under the suspicion that the latter would test-fire a missile (Dispute3 numbers: 4125, 4226, 4322, and 4456). However, in none of these cases that the suspected country actually test-fired any missile (Palmer et al 2015). Hence, it appears that the pundits’ concerns about missile tests are not empirically well-founded. I was able to show this in a large-N statistical setting. In addition, a closer look into MIDs indicate that missile tests do not typically trigger MIDs, but there are only a handful in which the suspicion of missile tests did.

Third, where do we move from here? One way to interpret the findings of this paper would be that missile tests are a trivial detail in international politics. This is already a new statement that we would not be able to make absent the data collection and the empirical analyses I conducted presently. Nevertheless, I cannot claim that we should abandon studying missile tests. First and foremost, the robustness checks using the ICB dataset showed that there may potentially be a negative association between crisis escalation and missile tests. Yet, these findings are obtained given the fact that the missile tests had a positive influence on the incidence of the crisis. Recall that findings using the MID sample weakly alluded to this pattern if the MID is initiated against the missile-testing country. The question is whether this pattern -although not robust enough yet- is telling us an alternative story that needs to be unraveled.
One possibility is that countries sometimes test-fire missiles to intensify their conflict with their adversaries. In this intensification, the public tests may serve as a coercive mechanism for the missile testers to secure a bargaining leverage. Several previous studies coin (and empirically support) the idea that countries sometimes self-select themselves to MIDs or crises in general to make certain demands against their opponents in the shadow of war. However, their intension is not to start a war, but to use the risk of war to secure a preferable solution to the conflict (Fearon 1994, 2002, Reed 2000). In line with this argument, a possibility could be that countries sometimes test-fire missiles by risking wars to get the resolutions they desire.

For example, after its country’s intense missile testing period until late 2017, Kim Jong-Un finally agreed to negotiating with South Korea and the US on various occasions. As the negotiations failed, North Korea resumed its missile tests. Admittedly, it could be that Kim Jong-Un halted his country’s tests simply as a gesture of cooperation. Alternatively, it could be that the missile tests partially communicated North Korea’s bargaining stance to South Korea and the US. Yet, once the negotiations failed, North Korea decided to communicate through missile tests again. It could be that we can expect more missile tests from North Korea until new negotiations begin.

There is always a risk that the case of North Korea is the only one that supports the missile tests as a bargaining tool argument. Nevertheless, it is worth further consideration: should this idea be generalizable, it would be an important contribution to our understanding of the role of the missile tests in the international
politics. Right now, we can argue that the role of missile tests is not to stir militarized conflicts and wars. That said, the investigation of the bargaining story requires a more thorough data collection project. We need data on occasions in which missile-testing countries decided to meet their adversaries to conduct negotiations.

References for Paper III


King and Zeng (2001) “Logistic Regression in Rare Events Data” Political Analysis 9(2): 137-163


Appendix for Paper III

(1) Summary Statistics

Table 13: Descriptive Statistics (MID Involvement Models)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>A initiates MID against B (H1-DV)</td>
<td>404,547</td>
<td>0.002</td>
<td>0</td>
<td>0.042</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Quarter years since A last initiated MID against B</td>
<td>404,547</td>
<td>98.156</td>
<td>88</td>
<td>67.698</td>
<td>0</td>
<td>252</td>
</tr>
<tr>
<td>B initiates MID against A</td>
<td>404,547</td>
<td>0.001</td>
<td>0</td>
<td>0.037</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Quarter years since B last initiated MID against A</td>
<td>404,547</td>
<td>99.823</td>
<td>90</td>
<td>67.941</td>
<td>0</td>
<td>252</td>
</tr>
<tr>
<td>A's missile tests (prev. quarter)</td>
<td>404,547</td>
<td>4.227</td>
<td>0</td>
<td>11.327</td>
<td>0</td>
<td>107</td>
</tr>
<tr>
<td>A's missile tests (prev. 6 months)</td>
<td>404,547</td>
<td>8.447</td>
<td>1</td>
<td>22.117</td>
<td>0</td>
<td>194</td>
</tr>
<tr>
<td>A's tests of new missiles (prev. quarter)</td>
<td>404,547</td>
<td>0.130</td>
<td>0</td>
<td>0.409</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>A's tests of new missiles (prev. 6 months)</td>
<td>404,547</td>
<td>0.256</td>
<td>0</td>
<td>0.625</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>A's tests of commissioned missiles (prev. quarter)</td>
<td>404,547</td>
<td>1.936</td>
<td>0</td>
<td>5.648</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>A's tests of commissioned missiles (prev. 6 months)</td>
<td>404,547</td>
<td>3.875</td>
<td>0</td>
<td>11.075</td>
<td>0</td>
<td>96</td>
</tr>
<tr>
<td>Number of Previous MIDs A-B</td>
<td>404,547</td>
<td>0.061</td>
<td>0</td>
<td>1.484</td>
<td>0</td>
<td>83</td>
</tr>
<tr>
<td>A's GDP (1 yr lagged, logged)</td>
<td>394,911</td>
<td>19.923</td>
<td>20.199</td>
<td>1.723</td>
<td>15.220</td>
<td>23.300</td>
</tr>
<tr>
<td>B's GDP (1 yr lagged, logged)</td>
<td>386,650</td>
<td>16.921</td>
<td>16.815</td>
<td>2.105</td>
<td>11.705</td>
<td>23.300</td>
</tr>
<tr>
<td>CINC A.CINC B</td>
<td>399,289</td>
<td>1.626.441</td>
<td>22.598</td>
<td>15.678.960</td>
<td>0.003</td>
<td>$26,815.800</td>
</tr>
<tr>
<td>Joint Democracy (dummy)</td>
<td>404,547</td>
<td>0.237</td>
<td>0</td>
<td>0.425</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Contiguity (dummy)</td>
<td>404,547</td>
<td>0.039</td>
<td>0</td>
<td>0.195</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A's Trade Depend.</td>
<td>404,547</td>
<td>0.0003</td>
<td>0</td>
<td>0.002</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B's Trade Depend.</td>
<td>404,547</td>
<td>0.005</td>
<td>0</td>
<td>0.028</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>% Change in Mil. Spending, A</td>
<td>382,649</td>
<td>0.093</td>
<td>0.067</td>
<td>0.358</td>
<td>-0.806</td>
<td>8.478</td>
</tr>
<tr>
<td>% Change in Mil. Spending, B</td>
<td>338,504</td>
<td>0.437</td>
<td>0.076</td>
<td>22.594</td>
<td>-1.000</td>
<td>1,965.292</td>
</tr>
<tr>
<td>Share alliance (ATOP)</td>
<td>404,547</td>
<td>0.145</td>
<td>0</td>
<td>0.352</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>New Leader in A (prev. 6 months)</td>
<td>404,547</td>
<td>0.042</td>
<td>0</td>
<td>0.200</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>New Leader in B (prev. 6 months)</td>
<td>404,547</td>
<td>0.003</td>
<td>0</td>
<td>0.055</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A's New Stonger Ally (prev. 6 months)</td>
<td>404,547</td>
<td>0.019</td>
<td>0</td>
<td>0.137</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>B's New Stonger Ally (prev. 6 months)</td>
<td>404,547</td>
<td>0.001</td>
<td>0</td>
<td>0.038</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 14: Descriptive Statistics (MID Escalation Models)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>MID Escalation to War</td>
<td>1,389</td>
<td>0.044</td>
<td>0</td>
<td>0.205</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A’s missile tests (prev. quarter)</td>
<td>1,389</td>
<td>8.799</td>
<td>1</td>
<td>17.565</td>
<td>0</td>
<td>107</td>
</tr>
<tr>
<td>A’s missile tests (prev. 6 months)</td>
<td>1,389</td>
<td>18.136</td>
<td>1</td>
<td>34.817</td>
<td>0</td>
<td>185</td>
</tr>
<tr>
<td>A’s tests of new missiles (prev. quarter)</td>
<td>1,389</td>
<td>0.185</td>
<td>0</td>
<td>0.471</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>A’s tests of new missiles (prev. 6 months)</td>
<td>1,389</td>
<td>0.404</td>
<td>0</td>
<td>0.813</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>A’s tests of commissioned missiles (prev. quarter)</td>
<td>1,389</td>
<td>3.826</td>
<td>0</td>
<td>8.233</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>A’s tests of commissioned missiles (prev. 6 months)</td>
<td>1,389</td>
<td>7.797</td>
<td>0</td>
<td>16.376</td>
<td>0</td>
<td>85</td>
</tr>
<tr>
<td>Number of Previous MIDs A-B</td>
<td>1,389</td>
<td>17.420</td>
<td>10</td>
<td>18.348</td>
<td>0</td>
<td>83</td>
</tr>
<tr>
<td>A’s GDP (1 yr lagged, logged)</td>
<td>1,389</td>
<td>20.074</td>
<td>20.374</td>
<td>1.816</td>
<td>15.220</td>
<td>23.300</td>
</tr>
<tr>
<td>B’s GDP (1 yr lagged, logged)</td>
<td>1,389</td>
<td>18.926</td>
<td>18.820</td>
<td>2.075</td>
<td>12.248</td>
<td>23.300</td>
</tr>
<tr>
<td>CINC A/CINC B</td>
<td>1,389</td>
<td>44.597</td>
<td>2.958</td>
<td>405.336</td>
<td>0.007</td>
<td>13,300.330</td>
</tr>
<tr>
<td>Joint Democracy (dummy)</td>
<td>1,389</td>
<td>0.102</td>
<td>0</td>
<td>0.303</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Contiguity (dummy)</td>
<td>1,389</td>
<td>0.381</td>
<td>0</td>
<td>0.486</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A was MID initiator</td>
<td>1,389</td>
<td>0.499</td>
<td>0</td>
<td>0.500</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>B was MID initiator</td>
<td>1,389</td>
<td>0.379</td>
<td>0</td>
<td>0.485</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A joined target</td>
<td>1,389</td>
<td>0.078</td>
<td>0</td>
<td>0.268</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

(2) **Bivariate regression results:** The MID involvement models also have time polynomials included as variables. Other than the time polynomials, the only variables of interest are missile tests variables. The results are displayed in Tables 15, 16, and 17.
Table 15: Bivariate Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 1)

<table>
<thead>
<tr>
<th></th>
<th>A Initiates a MID against B</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Dyads</td>
<td>Relevant Dyads</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A’s Missile Tests (prev quarter)</td>
<td>0.010***</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A’s Tests of New Missiles (prev quarter)</td>
<td>0.220***</td>
<td>0.070</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A’s Tests of Commissioned Missiles (prev quarter)</td>
<td>0.021***</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.080)</td>
<td>(0.081)</td>
<td>(0.087)</td>
<td>(0.086)</td>
<td>(0.086)</td>
</tr>
<tr>
<td><strong>Time Polynomials</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>367,701</td>
<td>367,701</td>
<td>367,701</td>
<td>176,919</td>
<td>176,919</td>
<td>176,919</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-4,302.679</td>
<td>-4,309.110</td>
<td>-4,302.304</td>
<td>-3,266.969</td>
<td>-3,266.766</td>
<td>-3,266.870</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
<td>8,615.358</td>
<td>8,628.220</td>
<td>8,614.609</td>
<td>6,543.939</td>
<td>6,543.533</td>
<td>6,543.741</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001
Robust standard errors in parentheses, clustered by Dyads

Table 16: Bivariate Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 2)

<table>
<thead>
<tr>
<th></th>
<th>B Initiates a MID against A</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Dyads</td>
<td>Relevant Dyads</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A’s Missile Tests (prev quarter)</td>
<td>0.006*</td>
<td>-0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A’s Tests of New Missiles (prev quarter)</td>
<td>-0.043</td>
<td>-0.245*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.119)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A’s Tests of Commissioned Missiles (prev quarter)</td>
<td>0.009</td>
<td></td>
<td>-0.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.035***</td>
<td>-2.977***</td>
<td>3.019***</td>
<td>2.826***</td>
<td>2.819***</td>
<td>2.819***</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.099)</td>
<td>(0.102)</td>
<td>(0.104)</td>
<td>(0.101)</td>
<td>(0.103)</td>
</tr>
<tr>
<td><strong>Time Polynomials</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>330,358</td>
<td>330,358</td>
<td>330,358</td>
<td>157,570</td>
<td>157,570</td>
<td>157,570</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
<td>6,320.548</td>
<td>6,325.482</td>
<td>6,322.813</td>
<td>4,930.437</td>
<td>4,927.035</td>
<td>4,929.122</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001
Robust standard errors in parentheses, clustered by Dyads
(3) Missile tests within 6 months prior to the observation: For the models estimated in TABLEs 9 and 10, I changed the treatment (i.e. missile test-related variables) such that this time the missile tests are counted not within the last quarter, but within the last six months prior to the realized dyad quarter year.
Table 18: Logit Estimates, MID Initiation vs Missile Tests (Hypothesis 1)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>A Initiates a MID against B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Dyads</td>
<td>Relevant Dyads</td>
</tr>
<tr>
<td>A's Missile Tests (prev 6 months)</td>
<td>0.004**</td>
<td>0.004*</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>A's Tests of New Missiles (prev 6 months)</td>
<td>0.079</td>
<td>0.114</td>
</tr>
<tr>
<td>(0.063)</td>
<td>(0.050)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>A's Tests of Commissioned Missiles (prev 6 months)</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Number of Previous MIDs</td>
<td>0.125***</td>
<td>0.126***</td>
</tr>
<tr>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>CINC A CINC B</td>
<td>-0.0005</td>
<td>-0.0006</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.005)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Joint Democracy (dummy)</td>
<td>-0.722***</td>
<td>-0.731***</td>
</tr>
<tr>
<td>(0.152)</td>
<td>(0.152)</td>
<td>(0.152)</td>
</tr>
<tr>
<td>Contiguity (dummy)</td>
<td>1.537***</td>
<td>1.507***</td>
</tr>
<tr>
<td>(0.132)</td>
<td>(0.130)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>A's GDP (1 yr lagged, logged)</td>
<td>0.132***</td>
<td>0.152***</td>
</tr>
<tr>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>B's GDP (1 yr lagged, logged)</td>
<td>0.201***</td>
<td>0.202***</td>
</tr>
<tr>
<td>(0.035)</td>
<td>(0.034)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>% Change in Mil. Spending, A</td>
<td>-0.540***</td>
<td>-0.538***</td>
</tr>
<tr>
<td>(0.196)</td>
<td>(0.193)</td>
<td>(0.194)</td>
</tr>
<tr>
<td>Share Alliance (ATOP)</td>
<td>-0.230</td>
<td>-0.217</td>
</tr>
<tr>
<td>(0.144)</td>
<td>(0.143)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>New Leader in A (prev. 6 months)</td>
<td>0.099</td>
<td>0.104</td>
</tr>
<tr>
<td>(0.212)</td>
<td>(0.212)</td>
<td>(0.212)</td>
</tr>
<tr>
<td>A's Trade Depend.</td>
<td>17.395</td>
<td>12.810</td>
</tr>
<tr>
<td>(13.874)</td>
<td>(14.112)</td>
<td>(14.021)</td>
</tr>
<tr>
<td>B's Trade Depend.</td>
<td>1.865***</td>
<td>1.876***</td>
</tr>
<tr>
<td>(0.236)</td>
<td>(0.228)</td>
<td>(0.232)</td>
</tr>
<tr>
<td>A's New Stronger Ally (prev. 6 months)</td>
<td>0.465</td>
<td>0.486</td>
</tr>
<tr>
<td>(0.295)</td>
<td>(0.298)</td>
<td>(0.295)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-10.985***</td>
<td>-11.535***</td>
</tr>
<tr>
<td>(0.975)</td>
<td>(0.778)</td>
<td>(0.806)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Polynomials</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>367,701</td>
<td>367,701</td>
<td>367,701</td>
<td>176,919</td>
<td>176,919</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-3,024,440</td>
<td>-3,027,239</td>
<td>-3,026,394</td>
<td>-3,239,081</td>
<td>-2,330,729</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
<td>6,082,880</td>
<td>6,088,479</td>
<td>6,086,789</td>
<td>4,692,162</td>
<td>4,695,457</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001
Robust standard errors in parentheses, clustered by Dyad
Table 19: Logit Estimates, Being targeted in a MID vs Missile Tests (Hypothesis 2)

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: B initiates a MID primarily against A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Dyads</td>
</tr>
<tr>
<td>A's Missile Tests (prev 6 months)</td>
<td>0.008*** (0.002)</td>
</tr>
<tr>
<td>A's Tests of New Missiles (prev 6 months)</td>
<td>-0.101 (0.114)</td>
</tr>
<tr>
<td>A's Tests of Commissioned Missiles (prev 6 months)</td>
<td>0.010* (0.005)</td>
</tr>
<tr>
<td>Number of Previous MIDs</td>
<td>0.123*** (0.005)</td>
</tr>
<tr>
<td>CINC A/CINC B</td>
<td>-0.002** (0.001)</td>
</tr>
<tr>
<td>Joint Democracy (dummy)</td>
<td>-0.241 (0.159)</td>
</tr>
<tr>
<td>Contiguity (dummy)</td>
<td>1.572*** (0.156)</td>
</tr>
<tr>
<td>A's GDP (1 yr lagged, logged)</td>
<td>-0.041 (0.045)</td>
</tr>
<tr>
<td>B's GDP (1 yr lagged, logged)</td>
<td>0.224*** (0.037)</td>
</tr>
<tr>
<td>% Change in Mil. Spending, B</td>
<td>0.002*** (0.001)</td>
</tr>
<tr>
<td>Share Alliance (ATOP)</td>
<td>-0.522** (0.184)</td>
</tr>
<tr>
<td>New Leader in B (prev. 6 months)</td>
<td>0.851** (0.279)</td>
</tr>
<tr>
<td>A's Trade Depend.</td>
<td>-72.828 (55.383)</td>
</tr>
<tr>
<td>B's Trade Depend.</td>
<td>2.344*** (0.242)</td>
</tr>
<tr>
<td>B's New Strongest Ally (prev. 6 months)</td>
<td>1.398** (0.522)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-8.315*** (0.872)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Polynomials</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>330,358</td>
<td>330,358</td>
<td>330,358</td>
<td>157,570</td>
<td>157,570</td>
<td>157,570</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-2,156,591</td>
<td>-2,163,360</td>
<td>-2,161,510</td>
<td>-1,673,746</td>
<td>-1,680,542</td>
<td>-1,678,506</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001
Robust standard errors in parentheses, clustered by Dyads

(4) Robustness Checks:

(i) Rare Events Logit: Here, I estimate Rare Events Logit models using the covariates from Model 1s of TABLEs II, III and IV. I use Choirat et al’l (2019)
`zelig` package developed for R. However, this package does not support clustering of standard errors. Therefore, the standard errors are not clustered in the models the results of which are summarized in the following table.

**Table 20: Rare Events Logit Estimates (Hypotheses 1, 2, and 3)**

<table>
<thead>
<tr>
<th></th>
<th>A Initiates MID</th>
<th>A Prim. Target</th>
<th>War A-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missile tests (prev. quarter)</td>
<td>0.006*</td>
<td>0.014***</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Number of Previous MIDas</td>
<td>0.125***</td>
<td>0.123***</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>CINC A/CINC B</td>
<td>-0.0004*</td>
<td>-0.002*</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.001)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Joint Democracy (dummy)</td>
<td>-0.724***</td>
<td>-0.241</td>
<td>0.865*</td>
</tr>
<tr>
<td></td>
<td>(0.148)</td>
<td>(0.157)</td>
<td>(0.380)</td>
</tr>
<tr>
<td>GDP of A (1 yr lagged, logged)</td>
<td>0.138***</td>
<td>-0.036</td>
<td>-0.271**</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.013)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>GDP of B (1 yr lagged, logged)</td>
<td>0.204***</td>
<td>0.227***</td>
<td>-0.185*</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.032)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>Contiguity (dummy)</td>
<td>1.540***</td>
<td>1.554***</td>
<td>-0.376</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.129)</td>
<td>(0.348)</td>
</tr>
<tr>
<td>% Change in Mil. Spending, A</td>
<td>-0.534**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.180)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Change in Mil. Spending, B</td>
<td></td>
<td>0.002***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Share Alliance (ATOP)</td>
<td>-0.225</td>
<td>-0.509***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.143)</td>
<td></td>
</tr>
<tr>
<td>New Leader in A (prev. 6 months)</td>
<td>0.120</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Leader in B (prev. 6 months)</td>
<td></td>
<td>0.873</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.317)</td>
<td></td>
</tr>
<tr>
<td>A's Trade Depend. on B</td>
<td>15.920</td>
<td>-71.320*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(18.280)</td>
<td>(33.590)</td>
<td></td>
</tr>
<tr>
<td>B's Trade Depend. on A</td>
<td>2.448***</td>
<td>3.018***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.742)</td>
<td>(0.756)</td>
<td></td>
</tr>
<tr>
<td>A's New Stronger Ally (last 6 months)</td>
<td>0.485</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.262)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B's New Stronger Ally (prev. 6 months)</td>
<td></td>
<td>1.465***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.431)</td>
<td></td>
</tr>
<tr>
<td>A is the MID initiator (dummy)</td>
<td></td>
<td>-1.168**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.443)</td>
<td></td>
</tr>
<tr>
<td>A is the MID target (dummy)</td>
<td></td>
<td>-1.317**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.474)</td>
<td></td>
</tr>
<tr>
<td>A joins the prim. Target (dummy)</td>
<td></td>
<td>-0.160</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.545)</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-11.230***</td>
<td>-8.437***</td>
<td>7.079**</td>
</tr>
<tr>
<td></td>
<td>(0.746)</td>
<td>(0.829)</td>
<td>(2.152)</td>
</tr>
<tr>
<td>Time Polynomials</td>
<td>Yes</td>
<td>Yes</td>
<td>N/A</td>
</tr>
<tr>
<td>AIC</td>
<td>6086.1</td>
<td>4350.3</td>
<td>474.14</td>
</tr>
<tr>
<td>Observations</td>
<td>367701</td>
<td>330258</td>
<td>1389</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001, standard errors in parentheses
(ii) **Heatmap Fit Analysis:** This analysis involves the comparison of empirical probabilities of the outcome of interest in the sample, and the predicted probabilities from the model of interest. Esarey and Pierce (2012) presents an R package called “heatmapFit” which basically allows us to generate a plot to compare the model fit with the sample.\(^{168}\) In the following plots, the y-axis represents the empirical probabilities of outcome (A initiates MID, A is the primary target of a MID, and A-B fight a war). The x-axis represents the predicted probabilities obtained from the relevant models. The diagonal dashed line is the perfect-fit line (i.e. the location where the empirical and predicted probabilities perfectly align). The thick line (i.e. heat map line) represents the trend of the association between empirical and predicted probabilities. The darker the shade of this line is, the less robust the presented association is in the relevant coordinate. If the heat map line is above the perfect fit line, this means that the model underpredicts the outcome of interest. Lastly, each plot involves a histogram of the distribution of the predicted probabilities on the x-axis.

\(^{168}\) See Esarey and Pierce (2012) for a more detailed explanation of how the plot is generated.
Figure 26: Heat Map Plot of Model 1 of TABLE 9

Figure 27: Heat Map Plot of Model 1 of TABLE 10
(iii) **Country Sensitivity**: For the sensitivity analysis, I re-ran the models 1 of TABLE 9, 10 and 11, 11 times. In each iteration, I dropped the observations belonging to one country to see whether the β estimate for missile tests and its 95% confidence intervals change. The following coefficient plot figure shows the β estimates for “all missile tests” variables within the last quarter and their 95% confidence intervals in each of these iterations.
Figure 29: Country Sensitivity Analysis for Model 1 of TABLE 9

Figure 30: Country Sensitivity Analysis for Model 1 of TABLE 10
The sensitivity analyses highlighted that once the US observations are dropped from the relevant samples, the impact of missile tests on MID onset (whether country A or B is the initiator) can almost be equally likely positive or negative. In the MID escalation, without the US observations, the missile tests have almost a pacifying effect. Most of the estimated $\beta$s in the 95% confidence margins are smaller than or equal to zero (i.e. top coefficient plot in Figure 31).

(iv) *ICB Dataset:* Here, I created a new sample to account for international crises. In this new sample, the unit of analysis is country-quarter year. The countries are the 11 ones for which I collected data on missile tests. The total number of observations is 2,351. The outcome variable is coded as 1 if the country of interest was involved in an international crisis (192 observations), and 0 otherwise. The information is acquired from Brecher et al (2017) – ICB dataset, hereby. The
definition of a crisis is based on Brecher (1996). Similar to MIDs, the crises typically transpire in the shadow of wars.

In terms of war, I relied on the “majres” variable which records the major policy response by the country of interest in the observed crisis. Unfortunately, the ICB dataset does not directly record wars in a fashion similar to MID dataset. Nevertheless, I still recorded a binary variable that is equal to 1 if the country of interest’s response in a crisis was coded as “multiple violent military acts”. This is the most violent category in the major response to a crisis variable that the ICB dataset presents. Of the 192 crises, only 15 reached to this “multiple violent military acts” level (which I treat as escalation to war).

As control variables in the crisis onset models, I include the country of interest’s GDP (logged and 1-year lagged), its Polity II score, percentage change in its military spending from year $t-2$ to $t-1$ where $t$ is the year of observation, a dummy that shows whether the country signed a military alliance with a stronger power in the last 6 months prior to the observed time, and another dummy indicating whether there was a leadership change in the country within the six months prior to the observed time.

For the models that estimated crisis escalation to war, I include the missile test, GDP, and Polity variables. In these models, the unit of analysis is a crisis, and the outcome variable is war. Nevertheless, since countries do not randomly appear in crises, I added the Inverse Mills Ratio to the crisis escalation models based on
the predicted probabilities of the relevant onset model. Models 1 and 2 in TABLE 21 investigate the impact of all missile tests, and Models 3 and 4 investigate the impact of commissioned missiles. Since there were no new missiles tested within a quarter year prior to a crisis, the “new missiles” variable was not investigated.

### Table 21: International Crisis Onset and Escalation Analyses, Probit Estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crisis</td>
<td>War</td>
<td>Crisis</td>
<td>War</td>
</tr>
<tr>
<td>Missile Tests (prev quarter)</td>
<td>0.013*** -0.054**</td>
<td>(0.002)</td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>Tests of Commissioned Missiles (prev quarter)</td>
<td>0.025*** -0.103**</td>
<td>(0.004)</td>
<td>(0.037)</td>
<td></td>
</tr>
<tr>
<td>GDP (logged, lagged)</td>
<td>-0.019</td>
<td>0.127*</td>
<td>-0.016</td>
<td>0.116*</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.050)</td>
<td>(0.040)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Polity II Score</td>
<td>0.014** 0.035</td>
<td>0.014** 0.035</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.025)</td>
<td>(0.005)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>% Chg in Military Spend. (t-2 to t-1)</td>
<td>-0.189</td>
<td>-0.173</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td>(0.142)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Stronger ally (prev. six months)</td>
<td>-0.361</td>
<td>-0.349</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.288)</td>
<td>(0.280)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New leader (prev. six months)</td>
<td>-0.160</td>
<td>-0.175</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.160)</td>
<td>(0.162)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inverse Mills Ratio</td>
<td>-3.970**</td>
<td>-3.773*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.513)</td>
<td>(1.682)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.977</td>
<td>2.225</td>
<td>-1.032</td>
<td>2.115</td>
</tr>
<tr>
<td></td>
<td>(1.021)</td>
<td>(2.200)</td>
<td>(1.029)</td>
<td>(2.459)</td>
</tr>
<tr>
<td>N</td>
<td>2351</td>
<td>192</td>
<td>2351</td>
<td>192</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
Time polynomials are included, but not displayed in models 1 and 3
* p < 0.05, ** p < 0.01, *** p < 0.001

---

169 More specifically, Inverse mills ratio is calculated as \( m(x) = F(x)/f(x) \) where \( x \) is the vector of predicted \( X\beta \) from the crisis onset estimates, \( F(x) \) is the cumulative distribution function, and \( f(x) \) is probability density function of \( x \)
Conclusion

In 2007, when the United States agreed with several Eastern European countries on installing missile defense systems in the latter’s territories, the Russian Premier Putin was less than happy. The expansion of NATO and the US’ military influence and equipment as far as to the former Warsaw Pact nations presented new military challenges for Russia. Putin’s response did not appear to be a mere protest. Russia test-launched a series of new generation Topol’-M missiles. These were new generation missiles which showed that Russia no more completely depended on its old Cold War ballistic technologies. Newspapers interpreted Putin’s missile tests as a sign of warning against the US and its expansion in the Eastern Europe (Isachenkov 2007). Furthermore, these increased tensions caused many to ask whether a direct military confrontation between Russia and the US would take place.

News of missile tests such as the one above are frequently heard in international relations. However, to this date, no scientific work investigated why countries sometimes publicly test more missiles than other times, and what the consequences of these tests are. One reason could be that the scholars of international relations perceive public missile tests as symbolic and inconsequential acts in international politics. Hence, according to them, missile tests may not be worth studying further. However, many news sources and analysts interpret public missile tests as provocative and dangerous acts that exacerbate tensions among nations. The apparent major problem is to figure out which side holds a view that is closer to reality. Should we worry about missile tests in international politics
because they increase tensions and ultimately lead to militarized conflicts? Alternatively, should we let countries get away with publicly test-firing their missiles as many times as they desire?

This dissertation laid out the first ever empirical analyses on publicly-conducted missile tests. My analyses revealed some intuitive and some counter-intuitive findings. Overall, the statistical association between militarized conflicts and publicly-conducted missile tests is at best weak. More specifically, first, I fail to find a generalizable pattern which would suggest that countries test-fire more missiles publicly as a response to military threats. Second, missile tests do not lead to more militarized conflicts, and/or wars. Therefore, the media’s alarming approach on missile tests may potentially be farfetched. Missile tests do not harbor militarized conflicts and wars. Thus, for example, Russia’s Topol’-M tests probably did not increase the risk of a Russo-American war. In fact, my findings highlight possibility that these tests were scheduled to take place as a part of developing new generation Topol’ missiles for the Russian navy, and not as a retaliation against America’s deployment of Patriots to Eastern Europe.

In what follows, I first summarize the analyses and findings from the three papers of this dissertation. Second, I discuss the implications of my findings. Lastly, I lay out the limitations of the present research and how future research can address these limitations.

Summary of the Findings

In the first paper, I presented my data collection project on missile tests. This collection includes 11 countries and their publicly conducted missile tests
between 1949 and 2015. My dataset is also the first cross-country record of missile tests presented in a format fit for statistical operations. Moreover, now, we have enough data to conduct statistical analyses, and to improve our understanding of the role of missile tests in international politics. In addition, the dataset includes variables that delineate various attributes of the missiles test-fired publicly as well as the locations where missile tests are conducted. This detailed coding scheme facilitates nuanced statistical analyses which ultimately can enlighten us more about what role(s) missile tests play in international politics.

Following a thorough depiction of the data collection procedure, the first paper summarized missile testing histories of the 11 countries of interest. These overviews were based on the yearly number of missiles tested by each country that I generated using the data I collected. A crude look at each of the countries’ missile test trends led to the following observation: increases in countries’ missile tests often coincide with their periods of major crisis or militarized conflict involvements. I concluded the Paper 1 with this observation and went on with more rigorous analyses to investigate the relationship between missile tests and international conflict in the next two papers.

Paper II investigated the question of what explains variance in countries public missile test trends. The importance of this paper is that its findings can help directly or indirectly answer highly policy-relevant questions such as why North Korea sometimes test-fire missiles more than usual – such as in 2017? Did the Trump-Kim Jong-Un summits in 2019 guarantee that the latter’s country will no more test-fire missiles in the near future? Why did India conduct an antisatellite
missile test in 2019? Was it related to the country’s recent fatal border skirmish with Pakistan? Through large-N statistical analyses, I sought to uncover a set of generalizable patterns behind increases/decreases in public missile tests, which in return could be helpful in speculating about individual cases such as North Korea, India, Iran, Russia, Pakistan, et cetera…

Based on policy papers and previous research, I presented three main factors that potentially affect variation in the number of missiles a country test-fires in a given year. The first factor was the development of new missile programs – which has been one of the most intuitive and conventional reason why a country would test more missiles than usual. The second factor was the presence of a military threat. These military threats could be in the form of a country’s previous militarized dispute involvements, its adversaries’ missile tests or joint military exercises. The third factor was the presence of an imposed economic sanction regime. After all, sanctions have been used as a policy tool by the UN and the US to curb the number of missiles tested by countries such as North Korea or Iran – yet, no study has really shown whether this tool is actually effective.

The statistical analyses I conducted revealed that of the three main factors, only the recent development of new missile programs consistently accounted for increases in missile tests. I also found that a country’s militarized dispute involvements have a weak effect on the missiles it publicly test-launches subsequently. However, neither adversaries’ missile tests, nor joint military exercises consistently affect a country’s decision to subsequently test missiles. Thus, majority of the factors that are associated with potential military threats are
at best trivial predictors. Also, economic sanctions do not seem to have any impact on missile tests. Besides all the three main potential explanatory variables, I lastly found that a strong economy is a very reliable predictor of increases in missile tests.

Overall, the second paper showed that public missile tests do not seem to be consistently driven by a wish to address foreign military threats. However, two factors appear to consistently affect changes in missile tests. First, a strong economy enables testing more missiles publicly. This makes sense since the US, and USSR/Russia are the top missile-testing countries and dwarf the rest of the world’s yearly number of missile tests. This finding casts doubt on the potential claim that missile tests are used by weak countries such as North Korea or Saddam era Iraq to intimidate strong powers such as the US. It actually appears that both North Korea and Iraq would potentially have test-fired missiles in much greater numbers had they enjoyed economic prosperity like the US. Second, if a country introduces a new missile program and test-fires a missile from this program for the first time, we can expect that more such tests will take place soon. For example, if we observe Iran introducing new cruise missiles during a naval exercise in the Strait of Hormuz, this could be a sign that more tests of these missiles will take place soon.

In Paper III, I investigated whether increases in a country’s public missile tests also affect the risk of its militarized interstate disputes (MID) and war involvement. Relying on Jervis’ (1976) seminal work on threat perception in international relations, I suggested that missile tests of a country should create major security concerns in the world. Furthermore, I anticipated that these security
concerns will create incentives for countries to attack their adversaries before the latter can attack first. Thus, I anticipated that increases in a country’s missile tests should significantly increase the risk of its MID initiation or being targeted in a MID as well as the escalation of these MIDs to war.

The statistical findings in Paper III suggested that there is not much support for the argument that missile tests create security concerns that result in MIDs and wars among nations. More specifically, although there appears to be a statistically significant association between missile tests and involving in MIDs, this association is substantively very weak, and highly sensitive to sample specifications. Also, it is the case that missile tests typically have a negative (yet statistically insignificant) impact on MID escalation to war. Additional statistical measures that address the selection problem in the sample do not change the results. Overall, the results of the third paper suggest that missile tests do not typically stir up trouble in the international politics: there seems to be no convincing statistical evidence that MIDs or wars are driven by increases in a country’s missile tests.

**Implications**

Overall, considering the three papers developed in this dissertation, we ought to remember three main points about publicly conducted missile tests. First, countries appear to conduct public missile tests to ready new missiles for combat as well as to ensure that the already commissioned ones are still combat-ready. Accordingly, for example, if we observe a country testing a brand-new missile for the first time, my analyses show that more public tests of this missile will take place.
Hence, my analyses basically support the very intuitive notion that missile tests are conducted to ensure that a missile is operationally capable.

The second notion we learn is that public missile tests are not a foreign policy tool of poor nations to intimidate stronger nations. Above all, I fail to find satisfactory evidence to suggest that missile tests are typically conducted to fend off foreign military threats. In addition, my analyses consistently show that countries with stronger economies test-fire more missiles. This finding may seem intuitive since having a stronger economy typically enables countries to have a larger selection of foreign/domestic policy tools to address their foreign/domestic policy goals (Clark et al 2008, Palmer and Morgan 2006). Nevertheless, we also know that economically-troubled nations such as North Korea, Saddam-era Iraq, Pakistan and Islamic Republic of Iran often test-fire (or used to test-fire) missiles. Media tends to focus on these countries since they antagonize the United States and its allies across the globe. At the same time, countries with large economies such as the USSR during the Cold War or the US occasionally test-fired missiles in 1,000s per year. This is almost a hundred times more than what nations such as North Korea, Pakistan, or Iran can typically afford to test-fire in a year.

The third notion is that there does not appear to be a consistent association between missile tests and militarized interstate disputes. There are two aspects to note here. First, I fail to find any evidence that suggests that missile tests are a typical response against military threats. Instead, it seems that countries typically test-fire missiles because tests are integral to the development and maintenance of missiles. This, however, does not mean that there are no instances when a missile
is test-launched to intimidate a certain adversary. My analyses suggest that these
instances are just not frequent enough to constitute a generalizable pattern. Hence,
what pundits typically interpret as an act of provocation or defiance is probably
rather a test that was scheduled well ahead of the test day, with the purpose of
assessing the capacity of a missile in question.

Second, increases in missile tests do not subsequently cause greater risk of
militarized interstate disputes and wars. A possibility is that when countries test-
launch missiles publicly, this does not lead to major security concerns, and then
“shoot first or die” incentives in the rest of the world. If we buy into Jervis’ (1976),
this may imply that countries typically see public missile tests as rather ordinary
acts. Then it is even possible that all the demarches and protests issued against
missile testing countries are almost standardized reactions, and not sincere
manifestations of panic and surprise. Then, the good news is that North Korea’s,
Russia’s or Iran’s public missile tests will probably not be a reason behind wars or
militarized conflicts that (will) involve these nations.

All these findings lead to several questions that need to be answered to
better understand the role of publicly-conducted missile tests in international
politics. First, why do we found peaks in countries’ missile test trends during their
major times of conflict in Paper I, but the results in both Paper II and Paper III
failed to corroborate it? The answer is straightforward: the latter two papers
involved multivariate statistical models. In Paper II, I investigated the impact of not
only the involvement in MIDs, but also of other possible factors that could
theoretically alter the number of missiles a country can publicly test-fire in a year.
It seems that when we account for alternative factors such as a large economy or the introduction of new missile programs, the seeming impact of militarized conflict involvement is limited or spurious.

In Paper III, I accounted for many alternative reasons why a country can find itself in a militarized conflict against another. In this multivariate setting, I found that missile tests were not a strong predictor of a country’s militarized conflict involvement. In other words, there are many reasons why tensions among countries rise and ultimately lead to MIDs, and missile tests cannot often be counted as one of these reasons. For example, suppose that the US deploys naval ships to Black Sea, and engages in jet buzz flights over Russian coastal towns following Russia’s Topol’ M tests. According to my analyses, the US’ hostility is more driven by factors such as geographic distance and balance of military power between the two countries than how many Topol’ M.s Russians test. The reason why the Topol’ M tests occurred shortly before the American ship deployment could even have been coincidental: the Russians might have already scheduled these tests before America’s Black Sea deployment. The results in Paper II suggest that this “coincidence” story is highly possible, as missile tests tend to often be scheduled events as parts of missile development programs.

Second, if missile tests appear to not be a direct cause of militarized conflict, should we turn a blind eye on North Korea, Iran, China, India or Pakistan’s missile tests? If our concern was that one of these countries would initiate a war due to their missile tests, this concern would be misplaced. Missile tests do not beget wars. However, there is always a possibility that these countries were militarily
threatened at some point and are trying to bulk up their defenses. After all, missile tests are driven by the introduction of new missile programs, and these programs are often costly and not introduced for no reason. Hence, missile tests at least signal that there is something problematic in the case of the missile-testing country in question, and that is why this country probably decided to improve its missile technology (or sought to ensure that its old missile systems are perfectly in order). So, a missile-testing country is not a country that enjoys completely peaceful relations with others. It has adversaries that may have posed threats to its security at some point.

Third, are missile tests a bargaining tool which countries use to get policy concessions? Statistical results observed in Paper III suggest that this alternative story would not be farfetched. Moreover, it seems that there is a weak, but positive association between missile tests and MID involvement. Yet, missile tests do not cause MIDs to escalate to wars. One possibility is that these findings simply allude to a simple empirical selection problem. Another possibility is that we find a story that is in line with Fearon’s (1994, 2002) argument: countries occasionally select themselves into MIDs (or crises) by publicly test-firing missiles. However, their aim is not to get what they want by fighting a war, but rather to intimidate and pressure their adversaries in a tense crisis environment to secure a favorable resolution in a conflict. Nevertheless, since the impact of missile tests on subsequent MID involvement is particularly weak, we cannot confidently argue

---

170 Countries that test-fire more missiles are also at the same time more likely to get involved in MIDs. The reason these MIDs do not escalate to war due to missile tests could be because of this selection problem.
that countries typically test-launch missiles to employ coercive diplomacy in bargaining.

**Future Work**

As highlighted early on, this dissertation presented the first empirical investigation of the role(s) of publicly-conducted missile tests in international politics. My analyses not only revealed a set of implications, but also led to a number of questions that future scholarly work can address. In this part, I will discuss a number of venues for future research based on the findings from the three papers I developed.

One question that the present work did not fully answer is why some countries initiate their indigenous missile development programs, while others do not. My analyses in Paper II show that once countries develop and test-fire brand-new missiles of their own, there is subsequently an increase in the number of missiles that they test-launch. However, it is also the case that some countries rely more on importing missiles than developing and consequently test-firing their own missiles. In fact, I found a weak but negative association between importing missiles and subsequently test-firing missiles. Many rich western European countries such as Germany, Belgium, Italy or Spain have much stronger economic means than Pakistan, Iran or North Korea to be able to develop and test-fire their own military missiles. Yet, they heavily rely on NATO or US imports, as well as weapons jointly developed by the European Union. One possibility may be that some countries develop their own indigenous missile programs because they lack
of a strong and reliable ally (or an IGO) that can relieve their defense burden – or in this case missile requirements.

The cases of North Korea or Pakistan however casts doubt on the “strong ally providing missiles” argument. Both countries started with importing missiles from other countries such as the Soviets or China. However, both countries later on reverse-engineered their imports and initiated their indigenous missile development programs. So, their self-reliance is not because they did not have access to missiles from much stronger countries. In fact, both countries probably needed imports first, so that they could learn how to develop missiles of their own. More recently, Turkey, although a NATO member, started focusing much more on developing its own missile arsenal – even though it has access to B61 nuclear bombs thanks to its NATO partnership. Therefore, one possible research project in the field of missile tests could be to investigate why countries sometimes decide to nationalize their missile arsenal, and why some countries do not? This is because, as my current results show, missile tests consistently increase as countries develop and introduce their own indigenous missiles

Another venue for future research is the investigation of potential consequences of conducting missile tests other than militarized conflicts and wars. In this study, the findings suggest that tensions created as a consequence of missile tests only rarely lead to MIDs and almost never lead to wars. That said, there still remains the fact that countries do not hide but display their missiles’ performances

---

171 One could argue that Turkey has reasons not to rely on its NATO partnership given how the latter delayed delivery of Patriot systems to the former, and the former ended up agreeing with Russia to purchase alternative missile interceptor systems.
to the rest of the world. There may be consequences behind missile tests that the present three papers have not uncovered yet.

As discussed in the “implications” section, public missile tests may potentially be a bargaining strategy. Although empirical analyses weakly suggest that this may be sometimes the case, future studies can investigate this issue in greater depth. Then, if missile tests are conducted publicly for bargaining purposes, this can have implications on how missile-testing countries approach the management of their conflict. For example, is North Korea more likely to sit on the bargaining table after test-firing enough missiles? This may not be far from the truth since after the large peak in North Korea’s missile tests in 2017, Kim Jong-Un agreed to at least engage in bargaining with the US President Trump. Yet, additional data collection is required to investigate any connection between public missile tests and resort to peaceful bargaining.

* * *

Overall, this dissertation presented more non-findings for many hypotheses it tested than statistically significant findings. Missile tests do not appear to be driven by most types of military threats. A country’s missile tests only have a weak association with its subsequent militarized interstate dispute involvements, and no association with its involvement in wars. It rather appears to be the case that missile tests are driven by introduction of new missile development programs and national wealth. Although these findings seem intuitive, we would not be able to know it, had this dissertation not involve a major data collection project as well as relevant empirical analyses on missile tests. The least one could say is that we now have an
opportunity to expand our understanding on missile tests since we have a proper large-N dataset for future statistical inquiries, and several questions that the findings of this research generated.

References for Conclusion


