

21<sup>st</sup> Century Trust Issues:

Arms Control Verification Regimes and Their State Parties

Jasmine Moheb

University of California - San Diego Political Science Department

Senior Honors Thesis

2020-21

## Abstract

In just under a century, arms control has both been conceived and become the forefront of international security with its high ambitions of regulating arms advancements that threaten the existing balance of power. Yet, human nature and the incentive mechanisms by which state actors make decisions has stayed relatively constant over time – allowing us to analyze these and apply them to a world that is in continuous flux. Making an effort to generalize patterns from a history of state behavior and apply them to the future is a critical function of political science research; one that I seek to contribute to in the field of arms control through this thesis. The relevance of arms control agreements in our modern global security network is an essential one, yet one that I believe is understudied. In a novel attempt to understand the intricacies of arms control agreements, I pursue the research question “Does the number of state parties in an agreement affect the total number of verification methods employed?” I find that the results are contradictory looking at just these two variables, but focusing on the inner workings of bilateral, regional, and international agreements, multilateral agreements tend to adopt more verification methods when they consist of a greater number of state parties. This is especially true in the case of certain verification methods over others. I also discover the role of third parties as well as the power of state players in regard to how they interact with the likelihood of adopting verification methods.

## Acknowledgements

I owe several people a debt of gratitude for their support throughout my Senior Honors Thesis as I could not have gone through a year of remote learning and truly immersed myself in the research process without their help.

Before Professor Erik Gartzke agreed to be my thesis advisor, his class on Weapons of Mass Destruction sparked a passion in me to pursue a career and research interest in arms control. Through attending his weekly classes, faculty lectures about his research on Cross-Domain Deterrence, and a coffee chat where he patiently let me ask 2 hours' worth of questions on his experiences, I am immensely grateful for all you did to help me throughout my thesis and how you inspired me prior to it.

Professors David Lake and Simeon Nichter were truly the best course instructors I could have asked for as every week was a learning process where we were able to learn the mechanics of political science research. This course enabled me to gain skills and an understanding of the science behind political science I had never been exposed to before, and I credit much of that invaluable academic growth to them. Thank you to Professor Lake for always providing insightful feedback and making time to meet with students, all of whom recognized how truly lucky they were to have you as a resource throughout this course.

As a student who entered this program terrified of data and completely unaware of how to read a regression output, I greatly appreciate the help from our Teaching Assistant Alexandra Lange who helped me understand how to use R and patiently explained what I certainly should have remembered from POLI 30. A special thank you to my friend Jane Abrams whose knowledge of R saved me more times than I could count.

Finally, a shout-out to my friends and colleagues in the Senior Honors Thesis program - Troy and Benny - both of you have been with me since our freshman year of college in political science classes, campaigns, student organizations, and this experience would not have been the same without you. Thank you for your steadfast faith in me and I look forward to seeing you two take the world by storm after UCSD.

# Table of Contents

<b>Abstract .....</b>	<b>2</b>
<b>Acknowledgements.....</b>	<b>3</b>
<b>1     Introduction.....</b>	<b>5</b>
<b>2     Literature Review.....</b>	<b>8</b>
2.1     Arms Control and International Agreements	
2.2     Verification Regimes in Arms Control	
2.3     The Role of Third-Party Verifiers	
<b>3     Theory and Hypothesis.....</b>	<b>17</b>
3.1     Hypothesis	
3.2     Analytical Claim and Theory	
3.3     Alternative Explanations	
<b>4     Case Studies .....</b>	<b>22</b>
4.1     Bilateral: SALT I and II	
4.2     Regional: CFE Treaty	
4.3     International: NPT	
<b>5     Quantitative Methodology.....</b>	<b>26</b>
5.1     Research Design	
5.2     Measurements	
5.2.1     Dependent Variable: Total Verification Methods	
5.2.2     Independent Variable: Number of State Parties	
5.2.3     Covariates: Third Party, BRI Level, CINC Score, Arms Type	
<b>6     Results .....</b>	<b>35</b>
6.1     Summary of Results	
6.2     Regressions	
6.3     Discussion	
6.4     Scope, Limitations, and Suggestions for the Future	
<b>7     Conclusion .....</b>	<b>55</b>
<b>8     References.....</b>	<b>57</b>
<b>9     Appendix.....</b>	<b>61</b>

# 1 Introduction

Arms proliferation is an ongoing threat to the safety of nations and civilians alike. A question that political scientists have yet to answer is how best to control their spread and maintain or eradicate the damage they may cause. Yet, the threat of arms proliferation is alive and well. Presently, nearly 16,000 nuclear weapons are held by nine countries: China, India, Israel, France, North Korea, Russia, the United Kingdom, and the United States – with potentially more on the way.<sup>1</sup> Although some international agreements seek to disarm nuclear weapons entirely, arms have continued to proliferate throughout history and actors do not always keep their promises of limiting nuclear stockpiles. Regional tensions growing in countries like India and Pakistan have also led to concerns regarding the development of smaller nuclear arsenals that may proliferate into more widespread conflict.<sup>2</sup> However, not only are nuclear weapons cause for concern, but conventional weapons diminish global security as well. One of the top security threats that causes tens of thousands of deaths and injuries annually are conflicts fought using small arms and light weapons that are misused and illicitly purchased or seized.<sup>3</sup> This is also the weapon of choice most accessible to terrorists, adding to conflict in various already destabilized regions. The issue of arms proliferation is a critical one that policymakers aim to address through various agreements at the unilateral, bilateral, and multilateral level. While the literature surrounding this issue is comprehensive, specific facets of arms control related to the number of state parties in agreements, I argue, are understudied – and I will be focusing on these in my thesis.

Before examining the background of arms control, it is important to define exactly what arms control is. Arms control is “any international control or limitation of the development,

---

<sup>1</sup> Nuclear Threat Initiative 2015.

<sup>2</sup> Geller 2003. Nuclear Threat Initiative 2015.

<sup>3</sup> Nakamitsu 2020; Schroeder and Stohl 2006.

testing, production, deployment, or use of weapons based on the premise that the continued existence of certain national military establishments is inevitable.”<sup>4</sup> The key phrase in this definition is that arms control accepts the continuation of military establishment – differentiating it from disarmament, which is the complete elimination of weapons. In fact, arms control was devised as a policy solution to replace the global goal of disarmament – as the idea of eliminating weapons seemed unrealistic and dangerous to many countries.<sup>5</sup> Instead, a policy outlook of arms control would allow the world to focus on the balance of power and prioritize the world’s overall stability instead of asking countries to disarm when they clearly had no incentive or trust in other countries to do so.<sup>6</sup> Unlike disarmament, the ultimate goal of arms control is to pave the way for deterrence. This relies on the theory of Mutual Assured Destruction - the concept that no state will make a first strike using a nuclear weapon because of the assurance that the state will be struck back.<sup>7</sup> Using this theory, we can rely on the fact that countries may have nuclear weapons, but this in fact will deter war. Arms control agreements seek to place limits on the number of weapons that are produced or kept in stockpiles. Ultimately, arms control is meant to lessen the likelihood of war or limit war’s destructiveness if it occurs.

Although arms control can be traced back to the Geneva Protocol in 1925, the theory and practice of arms control became most prominent during the Cold War. During talks of disarmament at the height of nuclear weapons proliferation, it became clear that the United States and the Soviet Union did not trust one another to disarm. They both understandably feared that being the first to disarm would allow the other side to have an advantage and potentially grow their own military arsenal to overtake the other instead of achieving mutual disarmament.<sup>8</sup> Therefore,

---

<sup>4</sup> Thompson 2016.

<sup>5</sup> JRank Law Library 2021.

<sup>6</sup> JRank 2021.

<sup>7</sup> Jervis 2009.

<sup>8</sup> JRank 2021.

they began to build on the policy objectives of arms control as a more feasible alternative to disarmament. While this paper does not seek to answer the question of whether arms control is effective, I must note the substantial decline in arsenal sizes of the United States and Soviet Union as nearly 51,000 nuclear weapons were disarmed from the peak of weapons development in 1986 during the Cold War to 2010.<sup>9</sup> The effectiveness in cooperation between the two powers gave credibility to the claim that arms control can prove to reduce tensions between escalating powers while allowing for some level of trust to be maintained between adversaries. While arms control became a prominent bilateral policymaking tool during the Cold War, a shift began to occur after the fall of the Soviet Union. There was a distinct rise in arms control agreements being negotiated at a multilateral level.<sup>10</sup> Arms control has always been a policy implemented at various levels – bilateral, regional, and international – but is becoming more popular now as a multilateral solution to conflict in the world. However, few arms control agreements throughout history have actually been successful and the negotiations process often leads to stalemate.<sup>11</sup> This brought me to a point of curiosity wondering what determines the various levels of commitments that state parties make to arms control agreements.

One facet of arms control that I found especially interesting was the verification regimes implemented by state parties. Verification in arms control is used as a tool to enhance confidence among state parties and therefore contribute to international security. In a technical sense, it means “the collection, collation and analysis of information in order to make a judgement as to whether a party is complying with its obligations.”<sup>12</sup> The ultimate goal of verification is to detect state parties who are not compliant with the agreed policy, deter potential non-compliers, and build trust

---

<sup>9</sup> Kristensen and Norris 2016.

<sup>10</sup> Müller 2000.

<sup>11</sup> Carter 1989.

<sup>12</sup> Dahlman 2010.

among the parties to an agreement to have confidence in its legitimacy. At a surface level analysis, one would think that verification methods are a great way to ensure state parties are abiding by their commitments and therefore the more verification, the better. However, the issue is more complicated than that. In fact, a debate about verification I later examine discusses the level of “appropriate” verification as one that is not too intrusive as to reveal one’s secrets and weaknesses in their arsenal, while holding the adversary accountable and preventing surprise attacks. Essentially, the more verification methods implemented by an agreement, the more transparent all arms programs will be, but the more vulnerable each state party will be to the others. Nevertheless, political scientists agree that verification procedures are still our best bet at restoring trust and enhancing commitment, a critical factor in parties’ willingness to cooperate in arms control agreements.<sup>13</sup>

At the intersection of two crucial matters - I find an interesting research question: With the increasing number of state parties in arms control agreements and varying verification methods by agreement, is there a relationship between the number of state parties in an agreement and the verification regime they choose to implement? I look at several verification methods to account for the verification regime and do a comprehensive analysis of arms control agreements from 1925 to 2010, factoring for several controls, to assess this potential relationship. Going forward with the threats posed by arms proliferation, the role of verification regimes will play an important part in the cooperation between states and the success of agreements. It is time to address whether the number of state parties affects the verification regimes agreed upon and consider what this means for arms control agreements now and in the future.

## 2 Literature Review

---

<sup>13</sup> Carter 1989.



A Russian proverb, later quoted by President Ronald Reagan during the negotiations of the IMF Treaty, says “Trust, but verify.” However, a truth universally acknowledged is that in politics that there is no such thing as trust. This especially applies to disarmament and arms control as it is difficult to trust an adversary, or even an ally, with carrying out actions that will diminish their individual security. Facing a collective action dilemma, it is always more beneficial for actors in an arms control agreement to choose to defect individually due to the cost they would incur if they were to cooperate and other states chose not to. While many would consider arms control negotiations to be a form of cooperation, John Maurer correctly points out that “the pursuit of arms control is not always a cooperative exercise.”<sup>14</sup> In fact, arms control is often used as a means for countries to advance their own interests. This effectively illustrates the contradiction we see before us: countries that need to cooperate, should cooperate, but cannot rationally do so. However, one way to alter this payoff structure is to allow for cooperation between states by implementing verification methods that enhance trust and transparency.

## **2.1 Arms Control and International Agreements**

The general literature consensus is that verification is an integral part of the arms control process. If any arms control agreement intends to be effective and carried through, the verifiability of the arms it proposes to regulate must be a central focus of the negotiations.<sup>15</sup> Specific instructions as well as a mechanism to keep parties accountable for their commitments is necessary to avoid any circumvention of the agreement. Verification helps enhance confidence among parties to an agreement and enables parties to make judgements as to whether others are complying with the stipulations decided upon.<sup>16</sup> This brings us to a central question that has rarely been addressed

---

<sup>14</sup> Maurer 2018.

<sup>15</sup> NATO Report 2000.

<sup>16</sup> Dahlman 2010.

by the literature surrounding arms control: What determines whether verification methods are pursued and the extent to which they are implemented in arms control agreements? I will be looking at how the number of state parties in an agreement, in particular, may play a role in the number of verification methods that are set in an arms control agreement.

Before explaining the theory behind my hypothesis, we must first understand how international agreements operate. Kenneth Abbott and Duncan Snidal outline two types of law that are relevant to examine here: soft law and hard law. Soft law is used to describe agreements that are not legally binding whereas hard law refers to binding agreements that can be legally enforced in a court of law.<sup>17</sup> Arms control agreements can take the form of both soft and hard law – with verification methods often making agreements “harder” in nature. Richard Williamson explains the utility of hard law as it creates consequences for actors if they do not abide by agreements, and often requires state parties to implement domestic policies to ensure that the agreement’s stipulations are carried through.<sup>18</sup> However, Abbott and Snidal argue that while hard law has utility, it is less likely to be voluntarily agreed upon in multilateral settings. They state that when there are many actors present in negotiations, actors tend to incorporate more soft law because it reduces costs for the actors involved and allows for more compromise between state parties that may have conflicting interests.<sup>19</sup> While this is a compelling claim, arms control agreements and their use of verification may not fit into this one-size-fits-all agreement claim, as I will argue throughout this thesis.

When examining how this theory plays out in arms control, Ola Dahlman claims that while verification is a “tool to strengthen international security,” the more trust that exists between state

---

<sup>17</sup> Abbott and Snidal 2000; European Center for Constitutional and Human Rights E.V. 2021.

<sup>18</sup> Williamson 2003.

<sup>19</sup> Abbott and Snidal 2000.

actors, the less verification that may be required in an agreement.<sup>20</sup> This is where arms control presents itself to be a unique case as Dahlman argues that there is an inverse relationship between trust among actors and verification methods being pursued. Therefore, these clashing points of view come to an impasse when considering whether actors in a multilateral setting would prefer an agreement with a low cost of entry as Abbott and Snidal claim, or whether this is outweighed by the verification methods they may perceive as a necessary cost they are willing to take on as they do not trust the multitude of state actors present.

Daniel Verdier identifies three types of actors in multilateral agreements, helping clarify what the priorities of these actors may be when settling upon an agreement. The three types of actors are (1) states with low compliance costs that will easily join, (2) states with intermediate compliance costs that need some additional incentives to join, and (3) states with high compliance costs that are left out and punished for nonparticipation.<sup>21</sup> In the context of arms control, in most agreements pertaining to arms and especially WMD such as biological, chemical, or nuclear weapons, a minority of state actors in the world actually have the capability and interest in producing these arms. This is exemplified by the “nuclear club” that is limited to the U.S, Russia, U.K., France, China, India, Pakistan, Israel, and North Korea.<sup>22</sup> These countries with nuclear capabilities are likely to fit into Verdier’s third category as they have high compliance costs if they were to join an agreement to reduce or eliminate nuclear warheads. This counteracts the point introduced by Abbott and Snidal as the costs for most actors in multilateral arms control agreements are in fact low because they fit into either the (1) or (2) category, but only a minority of actors fall into the (3) category. Therefore we see that states choosing between the benefits of a soft

---

<sup>20</sup> Dahlman 2010.

<sup>21</sup> Verdier 2008.

<sup>22</sup> BBC News 2020.

agreement that could come at a low cost to them but have no verification methods and a hard agreement that includes verification methods but may be more costly would likely prioritize the latter because their individual costs in upholding the treaty (unless if they are a country with significant arms) are not very high to join. This is why in a multilateral agreement with many actors that do not trust one another but have low compliance costs for arms control, the prediction made by Abbott and Snidal that only soft agreements with minimum requirements would be agreed upon is not necessarily the case.

Before examining the literature pertaining to verification, there is a primordial argument about arms control that is introduced by Julian Schofield. Schofield argues that arms control counteracts the balance of power that maintains peace between powerful world players.<sup>23</sup> The logic behind this argument is that successful arms control requires environments that assure certainty, for example, through means of transparency and verification. However, the balance of power that is required to deter powerful states from acting aggressively thrives in conditions of uncertainty, for example, when several actors can threaten the existing balance, preventing any one party from becoming hegemon. This leads to an entirely different school of thought that argues arms control is inherently counterproductive to producing the uncertainty required for states to be fearful of other balancing state actors, and that we should instead be solely relying on the balance of power to promote safety in this world. However, the balance of power argument comes with its own flaws as it is possible that a world of uncertainty does not end in balancing, but rather arms proliferation as witnessed during the Cold War.

Another question that comes to mind when presented with the dichotomy pointed out by Schofield is how the balance of power, assuming it plays a role in maintaining peace, will be

---

<sup>23</sup> Schofield 2000.

affected by the number of nuclear weapons kept in stockpiles as decided upon in arms control agreements. Seeing as how deterrence will only be effective in a balance of power setting where all actors have some leverage over the others, a potential point of contention is state parties deciding how many nuclear weapons are needed to maintain this balance – avoiding too many as to reduce unnecessary risk, but ensuring there are still enough in existence to present a substantial and reciprocal threat. This question requires a more speculative answer and may vary by agreement but will be one that researchers and policymakers must consider when making decisions pertaining to the target number of arms they want to maintain. While the literature may disagree on whether arms control or the balance of power will lead to a more stable world, my thesis will not be answering the question of which is better. Instead, I will be focusing solely on arms control and exactly how verification mechanisms are agreed upon. This, in turn, should pave the way for future research to have a more comprehensive foundational understanding of verification regimes and cooperation before answering overarching questions pertaining to their effectiveness in global stability.

## **2.2 Verification Regimes in Arms Control**

When analyzing the literature surrounding verification regimes, first, I will clarify the difference between monitoring and verification. Monitoring usually refers to efforts made towards detecting, identifying, or measuring any developments or activities that may be of interest to a state party.<sup>24</sup> For example, one form of monitoring may be done through National Technical Means. In this thesis, for the sake of simplicity, I include this as one method of verification, but acknowledge that it is different in degree of intrusiveness compared to several other types of verification that I measure a verification regime with. Verification, alternatively, refers to a process in which a

---

<sup>24</sup> Meyer 1984.

judgement is made about compliance to a treaty.<sup>25</sup> Using the many technologies available to countries, the verification process of determining treaty compliance is often perceived as subjective and judgmental. This also means that verification can be susceptible to political manipulation based on the actor conducting the verification process.

Acknowledging the flaws and potentially skewed information that may come with verification, there is also a debate surrounding what is truly “good” verification for an arms control agreement. It is worthwhile to consider this as there are limits to how much information countries want to share about their arms programs due to the possibility of sharing too many details about important security measures, and this debate helps set a standard for how much transparency should be exchanged between countries. Karl Pieragostini argues that verification should try to “satisfice” rather than maximize the information extracted from countries.<sup>26</sup> The explanation behind this is that adequate verification of an arms control treaty will be enough to encourage long-term rational behavior. Another perspective on how much verification should be agreed upon is presented by the United States during the Senate ratification of the Intermediate-Range Nuclear Forces (INF) Treaty in 1988. Ambassador Paul Nitze set the standard for effective verification during hearings by stating “if the other side moves beyond the limits of the treaty in any military significant way, we would be able to detect such violations in time to respond effectively and thereby deny the other side the benefit of the violation.”<sup>27</sup> Similar to Pieragostini’s “satisfice” standard, the United States’ perspective on effective verification reaffirms a point that states are not interested in maximizing all verification methods. Instead, they prefer a verification regime that does what is necessary to prevent any potential harm or setbacks their country may end up with if

---

<sup>25</sup> Meyer 1984.

<sup>26</sup> Pieragostini 1986.

<sup>27</sup> Dahlman 2010.

another country is not in compliance with the agreement. States, therefore, have varying degrees of comfort in accepting verification methods, and this adds a layer of complexity to the variable I am studying in how the interaction between states may affect the final verification regime that is decided upon.

Ultimately, the reasoning behind why countries prefer not to maximize the verification methods they agree to is explained by both James Caporaso and Daniel Verdier. Both highlight the realist school of thought and relate it to state actors, proclaiming that through a realist lens, actors are anarchical and self-interested.<sup>28</sup> While the world may benefit as a whole when the maximum number of verification methods are implemented at all times, a state may be giving an adversary access to new technology or military advancements they have made by increasing intrusive access to their arms programs. In an anarchical world, states will always prefer to protect their private information if the opportunity cost to giving it up would diminish the tactical advantage they have against other states that are also acting in an anarchical fashion.

### **2.3 The Role of Third-Party Verifiers**

In order to understand the role that third parties play in arms control agreements and verification, it is important to clarify what exactly a third party in this case may be. Robert Keohane explains the role of an international regime or intergovernmental organization as one that helps describe patterns of cooperation through setting norms and rules for collective choice, outlining decision making procedures, and setting procedures for implementation.<sup>29</sup> Barbara Koremenos, Charles Lipson, and Duncan Snidal further point out that international institutions have a rational and intentional design made by states and actors that try to leverage political power through them.<sup>30</sup>

---

<sup>28</sup> Caporaso 1992; Verdier 2008.

<sup>29</sup> Keohane 2005.

<sup>30</sup> Koremenos, Lipson, and Snidal 2001.

Essentially, political scientists agree that international organizations or institutions are created by states to directly interact with other states and achieve cooperative solutions through an intermediary to ideally achieve the greatest outcome for the individual state party and the global community as a whole.

David Koplow points out that, when comparing bilateral and multilateral settings, the presence of a third party varies significantly. Examining the payoffs from using a third party to conduct verifications, Koplow states that bilateral arms control agreements tend to engage in reciprocal inspection whereas multilateral agreements tend to create or co-opt international organizations to conduct inspections for the agreement.<sup>31</sup> However, not only is it more likely for international organizations to serve as a third party in multilateral agreements, but Koremenos, Lipson, and Snidal say that as the number of actors increases in an agreement that is facilitated through an international institution, centralized bargaining reduces transaction costs and reduces cooperation problems.<sup>32</sup> This essentially adds an incentive for multilateral agreements, as they grow larger to include more state parties, to use a third-party verifier as it makes cooperation more likely and reduces the costs of cooperation – not only during the crafting of the agreement, but as a permanent oversight mechanism.

The topical literature additionally agrees about the benefits of third-party verifiers. Third-party verification in multilateral agreements allows for monitoring, collecting facts, sometimes enforcing treaties, and interpreting or clarifying stipulations.<sup>33</sup> An example of a third-party verifier in arms control is the NATO alliance's High-Level Task Force (HLTF). This is an outlet that provides experts to partners on issues pertaining to arms, coordinates verification activities,

---

<sup>31</sup> Koplow 2017.

<sup>32</sup> Koremenos, Lipson, and Snidal 2001.

<sup>33</sup> den Dekker 2004.



monitors implementation issues, and an arms control database that state parties have access to.<sup>34</sup>

The role of third parties is one that I later argue plays a part in verification regimes state parties choose to adopt due to their likelihood of cooperation and the reduction in cost – and this uniquely affects the multilateral agreements that benefit the most from their presence.

### 3 Theory and Hypothesis

My hypothesis centers around the number of state parties in arms control agreements and the effects this may have on the verification regime adopted in the agreement. Initially, I began my thesis trying to compare bilateral and multilateral agreements. However, upon realization that my number of observations was limited to 45 and only 15 of those were bilateral agreements, I decided that looking at the number of state parties as a whole would produce more comprehensive and confident results.

#### 3.1 Hypothesis

I hypothesize that arms control agreements that consist of more state parties will adopt more verification methods and potentially more intrusive verification methods compared to arms control agreements that have fewer state parties. I theorize that a country's willingness to adopt verification methods is dependent on these factors: (1) cost and capability of adopting verification methods, (2) trust in other state parties to cooperate, (3) guarantee that verification methods will be fairly adjudicated.

Using the number of state parties in the arms control agreement as my independent variable and the verification regime as my dependent variable, I make the following hypotheses:

**H<sub>1</sub>:** Arms control agreements with more state parties will tend to adopt more verification methods than arms control agreements with fewer state parties.

---

<sup>34</sup> NATO Report 2000.

**H<sub>1</sub>:** Arms control agreements with more state parties will tend to adopt more intrusive verification methods than arms control agreements with fewer state parties.

I also include a null hypothesis to check for the lack of a relationship between my independent and dependent variables:

**H<sub>0</sub>:** There is no relationship between the number of state parties and the verification regime of the arms control agreement.

*Independent Variable:* Number of state parties in the arms control agreement

*Dependent Variable:* Number of verification methods

*Secondary Dependent Variable:* Intrusiveness of the verification regime

### 3.2 Analytical Claim and Theory

The analytical claim I make that explains the reasoning behind my hypotheses connects to the three factors I mention earlier that indicate a country's willingness to adopt verification methods. I argue that multilateral agreements with more state parties will tend to adopt a more vigorous verification regime than agreements with fewer state parties because –

1. *Cost and Capability:* The burden of cost for verification methods can be crippling to even the wealthiest countries, and therefore must be considered as a factor in their hesitation to agreeing to an extensive verification regime. When the U.S. took on the burden of verification during the treaty on Conventional Armed Forces in Europe (CFE), the cost of inspections alone was over \$3 million in 1988.<sup>35</sup> This drove the U.S. to decrease its financial investment in arms verification as their spending did not meet a sufficient cost-benefit analysis. The grievance of cost is exacerbated for countries with lower financial capabilities and less access to technology and specialists in arms control. The cost of

---

<sup>35</sup> Krass 1996.

cooperation for states is lower when using shared databases, surveillance technology, or topical experts. Using multilateral resources lowers the cost barrier that exists for many states that may be interested in employing verification methods but previously were not able to find this feasible given their insufficient access to resources. When costs for verification are reduced because there are more actors sharing the burden of the cost, I argue that states are more likely to adopt verification methods and more states sign onto the agreement. As more intrusive verification methods are often more costly, these may also become more accessible using shared resources and evenly distributed financial commitments.

2. *Trust in Other State Parties to Cooperate:* I argue that states generally do not trust one another, and when more states are trying to find a way to cooperate, the only way they will be able to achieve trust will be by enacting verification methods (and likely more rigorous ones) to serve as a form of enforcement or a check on others' actions. When fewer states are involved, there are fewer relationships to navigate and therefore trust and cooperation may be more easily achieved with less verification needed. George Downs, David Rocke, and Peter Barsoon contribute to this by claiming that deeply cooperative regimes have a limited number of members while regimes with large numbers of members engage in shallow cooperation.<sup>36</sup> While this can be interpreted to mean that agreements with fewer members are more cooperative and therefore will be more cohesive in agreeing to a series of verification methods, I argue that these already-cooperative state parties are less likely to be in need of a more comprehensive verification regime. Instead, I argue that the claim that settings with more members engage in shallow cooperation reinforces my point that

---

<sup>36</sup> Downs, Rocke, and Barsoon 1996.

the need for verification due to an initial lack of trust is more salient among state parties in larger multilateral settings and leads to a stronger verification regime as needed for countries to sign on as state parties. To summarize the argument: as the number of states increase, I theorize that the overall trust in the agreement and its state parties decreases, and therefore actors are more likely to seek more verification methods to keep other state parties in check. This may also extend to increasing more intrusive verification methods as a part of the regime as this will enhance confidence and trust for state parties that are insecure in their relationship with a larger group of actors.

3. *Guarantee that Verification Methods Will Be Fairly Adjudicated:* The issue of trust and having an objective party facilitate verification is more contested than many would initially believe. States often criticize the lacking transparency, political bias, weak evidence and faulty intelligence assessments they believe are rampant in verification efforts.<sup>37</sup> Partly, this has to do with their disagreements in interpretation of the standards set by an agreement, but often it has to do with the party conducting the verification procedure in the first place – especially when a single state actor is in this role. Third-party verifiers, alternatively, are set to serve as objective assessors that states place trust in to make unbiased conclusions about the parties’ compliance to an agreement. The presence of third-party actors such as the IAEA (International Atomic Energy Agency) is uniquely more accessible at the multilateral level because larger agreements tend to be negotiated using international organizations, connecting state parties to the many resources in place within them. Multilateral agreements, more often than bilateral ones, co-opt international organizations or third parties to serve as an intermediary for negotiations and oversight in arms control

---

<sup>37</sup> Moodie and Sands 2001.

verification regimes.<sup>38</sup> Therefore, agreements with more state parties that utilize third-party verifiers are more likely to perceive the verification efforts as objective and fairly adjudicated, persuading them to enact more verification methods and likely more intrusive ones. When state parties do not trust the actor(s) conducting verification to be objective, as is the case often with bilateral agreements, I argue that they are less likely to impose a stricter verification regime.

### 3.3 Alternative Explanations

I must address the potential counterfactual to my theory. This is that countries are more likely to accept an agreement if it consists of fewer verification methods because there is a lower barrier to entry and therefore lower buy-in for them. If this is true, then we are likely to see the opposite trend of what my hypothesis predicts, and instead there will be fewer verification methods as the number of parties increases. This would also indicate that countries value cooperation in weaker agreements over ensuring treaty compliance in a more comprehensive agreement. My hypothesis disputes this by essentially claiming that states will prioritize compliance in a more stringent agreement over cooperation in a subpar or symbolic one.

There may also be the issue of reverse causality between my independent and dependent variables. It is possible that instead of the number of state parties influencing the verification regime, that the verification regime (total number of verification methods or intrusiveness) affects the number of state parties interested in participating. My theory currently answers to this by claiming that state parties are the ones who make agreements, not the other way around, and therefore once state parties are discussing an arms control agreement they believe is necessary, they begin to come up with verification methods they are willing to agree upon. Therefore, the

---

<sup>38</sup> Koplów 2017.

verification regime would be contingent on the state parties that choose to craft and adopt the agreement in the first place.

## **4 Case Studies**

In this section, I will be using case studies of bilateral, regional, and international agreements to show how they vary in the arms control negotiations process and how this is actualized in the verification regimes they agree upon. The case studies I will be examining as a representation of bilateral agreements will be the U.S./Soviet Union negotiations of SALT I and II during the Cold War. The regional case study will be about the CFE Treaty (Conventional Armed Forces in Europe). Finally, I will use the NPT (Non-Proliferation Treaty) as an example of an international agreement to finalize my comparative narrative.

### **4.1 Bilateral: SALT I and II**

The Soviet Union and U.S., in a contentious bilateral arms race, understandably were skeptical of one another throughout the arms control agreement proceedings. The USSR was notorious for wanting to avoid intrusive verifications – and the U.S. had to comply with these standards in order to achieve policy outcomes from negotiations.<sup>39</sup> What stands out in my dataset of bilateral agreements during the Cold War is that they heavily relied on National Technical Means (NTM) for their mode of monitoring. I later explain more about what this means, but the simple explanation is that the U.S. and USSR conducted their own verifications rather than using external sources or making cooperative efforts to verify. These primarily unilateral efforts limited the scope of verification and led to some questions as to whether the arsenal limitations in arms control agreements were truly met by each state party. One example demonstrating this was the

---

<sup>39</sup> Schaper 2000.

possibility that nuclear weapons were moved into the Kaliningrad enclave on the Baltic instead of meeting the agreement's stipulation of complete withdrawal.<sup>40</sup>

SALT I (Strategic Arms Limitation Talks) heavily relied on NTM such as aircraft reconnaissance, satellite imagery, and monitoring systems – all of which were non-intrusive and unilaterally operated. It also created a Standing Consultative Commission (SCC) that did no monitoring but was available for consultation, and therefore boosted the confidence of the state parties that provisions were being abided by. Due to these limitations, the scope of the treaty itself was not comprehensive enough to make significant cutbacks on arms but was able to take some limited steps such as freezing the development of strategic missile launchers.<sup>41</sup> SALT II set additional restrictions on arms while maintaining a reliance on NTM as the primary monitoring system. Verification was expanded to include more information exchanges and notification of tests, however, cooperative verification had yet to be agreed upon. The SCC continued in its role as a consultant and set certain standards for verification but did not conduct any verification itself.

Both the SALT I and II talks demonstrate that bilateral agreements are possible but prove to be highly independent and rely on some information observe and exchange, but primarily use one's own national means. In both cases, actors were hesitant to advance intrusive verification in fear that it would allow their respective adversary to gain an edge on their statecraft secrets. They also were able to produce a consultative commission, but this held little to no power as an agent able to conduct verification and rather served as a consultant to the two parties involved. When examining my dataset, SALT I and II had a total of 1 and 2 verification methods, respectively, implemented in total for their agreed-upon verification regime.

## 4.2 Regional: CFE Treaty

---

<sup>40</sup> Ifft 2001.

<sup>41</sup> Schaper 2000.

The Treaty on Conventional Armed Forces in Europe (CFE) was signed in 1990 by thirty state parties in the European region. The treaty further restricted treaty limited equipment (TLE) and set requirements for a comprehensive verification regime that consisted of NTM, information exchange, and both unrestricted and selective on-site inspections. This treaty was exceptionally successful in achieving over 3,000 inspections and leading to the destruction of 50,000 pieces of equipment.<sup>42</sup>

One distinct feature of the CFE treaty was the third parties it created to oversee implementation of the agreement. The state parties created a High Level Task Force (HLTF) as well as a Verification Coordinating Committee (VCC) that would help better develop and coordinate arms control policy and verification. Through these two agents, the state parties have experts they can rely on for issues on arms control, are better able to coordinate verification activities, can monitor issues that arise with their implementation, can train multinational inspection teams, and have access to a joint arms control database.<sup>43</sup>

The CFE Treaty is an example of a regional agreement that was successful in implementing several verification methods and an intrusive verification regime alongside independent agencies that would be able to carry out many of the actions needed to ensure treaty compliance. Examining my dataset, the CFE Treaty had a total of 4 out of 5 verification methods. While this is a strong example of a regional agreement that adopted many verification methods – I also acknowledge that it cannot be an overall representation of all regional agreements and especially how they compare to bilateral ones. Social science research is often situational and U.S./Soviet tensions during the Cold War cannot be completely controlled for when trying to make comparisons with other countries and regions. I try to address more of this comparative in my quantitative section.

---

<sup>42</sup> NATO Report 2000.

<sup>43</sup> NATO Report 2000.



However, for the purpose of the case study, I argue that the CFE Treaty effectively demonstrates how a regional agreement may have an easier incentive structure to enact a more stringent verification regime.

#### **4.3 International: NPT**

As an international effort to prevent the proliferation of nuclear weapons, the Nonproliferation Treaty (NPT) was signed in 1968 and has 191 state parties to-date. The NPT places safeguards for nuclear activities in non-nuclear states to prevent the addition of new nuclear states to the nuclear club. These safeguards are regulated and overseen by the IAEA as well as other non-nuclear weapons states that are parties in the NPT. The state parties agree to routine monitoring, inspection, and access to certain “strategic points” in facilities. The state parties entrust in the IAEA to have access to nuclear sites for inspections as well as areas that have potential for nuclear material to be present. Additional protocols have also been put in place to regulate undeclared activities and nuclear material. This is due to the previous findings of undeclared nuclear programs being developed in Iraq and the People’s Republic of Korea.<sup>44</sup>

Something worth noting in the NPT is that while in this case I aim to use it as an example of an international agreement, every international agreement has different arms it tries to control for and to different degrees. The NPT is a less controversial one as most countries in the world agree with the stopping of nuclear weapons proliferation. What I seek to show in this case study is that even though there were many state parties that signed onto this agreement, they still were able to agree on verification methods greater than or equivalent to what was seen during the SALT agreements that were between 2 parties. My analysis may not be perfect as I am unable to control for the fact that they seek to regulate different arms among different countries in this case study,

---

<sup>44</sup> Dahlman 2010.

but it is able to show a difference in the number of state parties and the verification regime agreed upon as well as the process in approving that. An added factor that has so far differentiated the regional and international agreements with bilateral ones has been a third party being given power to enforce the verification regimes, as seen in this case with the IAEA.

## **5 Quantitative Methodology**

### **5.1 Research Design**

The research design is meant to measure the relationship between two variables: the number of state parties in an arms control agreement (the independent variable) and the number of total verification methods in that agreement (the dependent variable). I was able to do this by compiling data on every arms control agreement from 1925 until 2010 and calculating a total verification method count using five types of verification that an agreement could have employed. There were several covariates I wanted to account for, including whether a third-party monitoring agency was present, the type of agreement that was signed (bilateral, regional, or international), the power of the countries involved in the agreement, and the type of arms that it regulated. Using data from my independent variable, dependent variable, and covariates, I was able to run several linear and multiple regression analyses that show the relationship that exists among them.

### **5.2 Measurements**

Here I will describe the complexities in exactly what my dependent and independent variable measured and how I was able to collect that data along with several covariates.

#### **5.2.1 Dependent Variable: Total Verification Methods**

The dependent variable I calculated for uses the total number of verification methods in an arms control agreement, and this consisted of five separate measures. Using a preliminary database

from the Stockholm International Peace Research Institute<sup>45</sup>, I was able to lay the initial foundation for the categories of verification methods that are typically enacted in an arms control agreement. These verification methods vary in intrusiveness and have been listed below in order from least to most intrusive.

1. ***National Technical Means (NTM)***: This encompasses the national monitoring techniques, such as satellites, that are used by an individual country. NTM is a unique measure as it requires a country's individual capability to produce and run a monitoring program over another country's arms. Therefore, I was cognizant of the reality that countries with less GDP or national means would be at a disadvantage when trying to employ NTM and therefore may be less likely to include it as a stipulation in an arms control agreement. Additionally, NTM is available to countries using their own means regardless of whether it is made official in an agreement, and therefore can be used in covert ways to achieve the same outcome of verification of another country's program. However, I specifically only looked at agreements that had an explicit commitment to using NTM as a verification method, not individual countries' covert operations as this is not publicly available information.
2. ***Information Observe and Exchange***: This verification method accounts for whether countries decided in their agreement to have a formal process of observing and exchanging data to enhance transparency of state parties abiding by the agreement. This often was actualized in the form of a database where countries would share information about, for example, how many warheads they had in their stockpiles. This particular verification method put the burden on an individual country to self-report the steps it had taken to

---

<sup>45</sup> Krass 1985.

follow the agreement. Unlike NTM, information observation and exchange is a fairly low-cost process as it does not require additional technology that may be out of reach for certain countries to report on their activities.

3. ***Black Boxes***: These are forms of technology that collect information on the arms control measures a country is taking and data is reported back to the other countries in the agreement. Black boxes are more intrusive as they give the home country less agency to alter or modify the information that other countries receive about their progress in reduction of arms. However, there are “information barriers” that prevent outside countries from receiving too much data that could be harmful to the country being reported on.<sup>46</sup> These may include computer algorithms determining the data’s validity in a yes/no format, limiting the amount of information being transmitted to potential competitors. Given that black boxes are a niche verification method, I expanded the definition for these in my data collection to include agreements that decided upon extensive technological verification without putting actual inspectors on the ground. This allowed me to have a more holistic view of the agreements with verification stipulations that included black boxes or achieved the same effect as black boxes.
4. ***Selective On-Site Inspections (OSI)***: Selective inspections occur when an individual country (or potentially a third-party agency) requests to visit and inspect another country’s arms program. The country receiving this request has the agency to accept or deny this request. This is more intrusive than the earlier forms of verification because an outside actor will potentially be on the ground inspecting your capabilities, but the country in question is still able to say yes or no to the request. The requests are also often required to

---

<sup>46</sup> Ifft 2001.

be made several weeks in advance of the actual visit, giving the country more control over how they wish to present their arms program to the verifiers.

5. ***Unrestricted On-Site Inspections (OSI)***: Unrestricted inspections are similar to restricted inspections but vary in that actors are allowed on-site inspections at any time without needing to notify the country that is being inspected. This is the most intrusive form of verification of the methods I account for as it both removes the inspected country's ability to deny the request or accept it at a later date.

Using these five verification categories, I was able to code for each agreement's verification provisions in a binary format, assigning a 0 to each category that an agreement did not include and a 1 to each category that an agreement did include. This allowed me eventually to find the sum of how many verification methods each agreement consisted of. Although I note earlier that these verification methods are organized in terms of increasing intrusiveness, seeing as how the difference in intrusiveness between each of the categories was not a constant number to be accounted for, I only looked at total verification methods assigning each a value of 1 instead of weighted values based on intrusiveness as this may have resulted in inaccurate data.

### **Data Collection of Dependent Variable**

As mentioned earlier, the preliminary data I used to base my categorical breakdown by was the SIPRI database that examined arms control treaties from 1925 until 1979.<sup>47</sup> However, this was limited in timeframe as well as missed several additional arms control agreements in that time period. Therefore, I utilized data from several outside sources including the InterAction Council, Arms Control Association, Nuclear Threat Initiative, International Atomic Energy Agency, United

---

<sup>47</sup> Krass 1985.

Nations Office for Disarmament Affairs, and U.S. State Department Archives.<sup>48</sup> One challenge I encountered was that there was not one designated database for all arms control agreements, therefore I had to compile my own using all the major sources that individually covered some of the arms control agreements over time. However, this allowed me to be more comprehensive in my data collection as I was able to cross-reference each of my 0 and 1 values assigned to each category for each agreement and verify them against 2 or more sources. I believe this enabled me to raise confidence in the accuracy of my dataset as I proceeded with the data analysis.

### **5.2.2 Independent Variable: Number of State Parties**

The number of state parties involved in each agreement was less difficult to compile compared to the dependent variable. The definition I used for the independent variable was “state parties” which would include all the states that signed onto that agreement at any point in time. I was able to use the same set of sources as mentioned earlier in my development for the dataset when calculating the number of countries in each agreement. This is readily available information, but sometimes I would find two sources that were disputing the number of countries in an agreement. Therefore, I cross-referenced the number of state parties I assigned to each agreement across 2 or more sources (3 or more if there was dispute).

### **Data Collection of Independent Variable**

The following sources had information on the number of state parties in each agreement: InterAction Council, Arms Control Association, Nuclear Threat Initiative, International Atomic Energy Agency, United Nations Office for Disarmament Affairs Treaties Database, and U.S. State Department Treaties Database. Please see footnote 48.

---

<sup>48</sup> Axworthy and Dean 2011; Arms Control Association Fact Sheets and Briefs; Nuclear Threat Initiative Fact Sheets and Treaty Text; IAEA Publications on Treaties; UNODA Disarmament Treaties Database; U.S. State Department Treaties Database.

### **5.2.3 Covariates: Third Party, BRI Level, CINC Score, Arms Type**

There were several characteristics of the total observations (agreements) I studied that I felt were relevant to account for as potentially confounding variables in the relationship between the number of parties and the verification methods in an agreement. These included whether a third-party verifier (for example, the U.N. or a relevant agency) was set up as a facilitator for information exchange between parties or conducted inspections, the level of the agreement (bilateral, regional, or international), the combined power of the countries involved, and the type of arms it targeted.

#### **Third Party**

One covariate that I examined was whether the arms control agreement included a stipulation that either created a third-party independent verification mechanism or used an existing international organization, such as a body within the United Nations, to serve as the arbiter of information transfer between state parties. I believe that data from this variable may serve as an explanation for the willingness of parties to engage in an agreement that includes verification. As I argue in my theory, I speculate that an independent actor that holds the parties accountable, collects and verifies information, and enhances transparency may positively influence the parties' commitment to verification mechanisms.

The standards I used to determine what constituted a third party was when an independent adjudicative body was created or specifically employed by the state parties to oversee provisions of the agreement. Some examples of these third parties I studied were commissions created by the countries, the Organization for the Prohibition of Chemical Weapons, African Commission on Nuclear Energy, IAEA, and others. However, the criteria for a third party does not include an international organization that served as the meeting point for arbitration between the parties as they crafted the agreement. The third party must specifically be working with the state parties after

the agreement has been signed and throughout the duration of its existence as a body for independent verification and adjudication.

### **BRI Level**

Another covariate that I created is the variable accounting for whether the agreement is one between two parties (bilateral), parties within a restricted region (regional), or open to all countries without limitation (international). For ease of reference, I use a shorthand to summarize this concept - “BRI Level.” The significance of this covariate is that it allows a deeper understanding of not just the number of parties that sign onto an agreement, but what the level of that agreement was. If it was meant for two parties, bilateral relationships may be inherently different in nature compared to multilateral ones as they may require more trust or may have higher stakes, such as those between the U.S. and USSR during the Cold War. However, regional agreements also add a layer of complexity in a multilateral relationship as they restrict the number of parties involved by region. This may lead to different findings between regional and international agreements as countries within a region may be more prone to working together due to their coexisting nature and may already have the infrastructure in place to enhance cooperation and avoid potential cooperation dilemmas. International agreements, alternatively, may have the greatest access to the greatest number of parties, but then also have the most difficult time catering to everyone’s preferences surrounding verification methods as well as the means they have available to employ various verification methods, such as NTM, that may come at high costs. Therefore, the breakdown by level of agreement is an important covariate for better understanding the nuances in numbers of parties that ultimately affects the verification methods agreed to.

I was able to measure BRI level by referring back to the several sources I used in creating my dataset on verification methods as they often indicated membership restrictions in agreements. If there were only two parties involved, I classified the agreement as bilateral. If there were



multiple parties - depending on the classification - medium-sized agreements (anywhere from 3 to 48) were often regional with membership restrictions, and larger agreements tended to be international with no membership restrictions.

### **CINC Score**

The CINC score covariate was used as a unique solution to account for a potentially confounding variable. When considering the role that the number of parties may have on the verification methods they agree upon, it was apparent that more powerful countries may have a greater sway in the decision of verification methods. Not knowing whether they would be in favor of more or less verification, it was important to understand whether the presence of “powerful” countries in an agreement (or many countries resulting in a larger cumulative power) would have a relationship with the total number of verification methods the state parties agreed to. This would require a measure for the power of each country in each agreement, an intimidating feat – but one that, with advice from my advisor, I was able to get around.

Using the Correlates of War database on National Material Capabilities (v 5.0), I examined CINC (Composite Indicator of National Capability) Scores. This is a widely-used index that accounts for each country’s annual values for “total population, urban population, iron and steel production, energy consumption, military personnel, and military expenditure of all state members.”<sup>49</sup> The index collected data on every country from 1816 to 2012, covering the starting point of the first arms control agreement (the Geneva Protocol in 1925) and allowing me to account for every arms control agreement until 2012 (ending with New START in 2010). Only two observations were not accounted for as they were signed after 2012, and therefore are not present in my dataset. The CINC score operates in a way that it assigns the total power of the world in any

---

<sup>49</sup> Correlates of War 2012.

year as a “1.” Then, it uses the annual values I mentioned above to assign proportions of that total power to each country. For example, in 1979, the power of the USSR was 0.1673254 as it applied in the SALT II bilateral agreement and the power of Afghanistan was 0.0018156 as it applied in the multilateral Environmental Modification Convention.

The CINC scores for each country would allow me to better understand how much “power” was in each agreement by doing the following: First, I added a column to my data of the year that each agreement was signed. Second, I added another column with the names of every country in each agreement (anywhere from 2 to 191). Third, I found the CINC score for each country in the year the arms control agreement was signed. An important note here is that in a few cases, a country signed the agreement after the official year it was signed because the country did not exist at the time of the initial signing. In those cases, I found the exact year the country itself signed the agreement and used the CINC score for that country from that year. In these unique circumstances, since I was collecting a few state parties’ data from different years, it was possible that I could get a higher score than 1 - which is the maximum power in the world for any one year. This occurred with only two agreements: the Geneva Protocol in 1925 and the Biological Weapons Convention in 1972 as many countries did not join until later as they became established after the initial year of signing, and therefore the total CINC score for state parties was over 1 (1.144 and 1.004, respectively). Fourth, I assigned those CINC score values to each country for each agreement and found the sum of the CINC scores for all the countries in each individual agreement. This allowed me to create a new column for CINC scores to be able to compare the total power in each agreement and account for how powerful or capable certain countries would be of either intimidating other countries into preferring their verification methods, or accepting more stringent verification methods as they were better equipped to afford the costs associated with verification. As an added measure, I included a column to depict average CINC scores for each

agreement by using the sum of CINC scores per country and dividing it by the number of parties in that agreement. After evaluating what the original and average scores would represent, I decided to use the original sum of CINC scores instead of the average CINC score by country in my regressions as the average score would misrepresent what I was trying to initially control for in my results (total power in an agreement, not by country). However, I included the average CINC score in my dataset if future research necessitates its use.

## **Arms Type**

The final covariate I examined was the categorical variable of the type of arms (conventional, biological, chemical, nuclear, or space) that was the focus of the agreement. I thought this may provide clearer insight into whether the type of arms determined the verification methods that countries decided upon. For example, would the fact that nuclear and biological warfare are less likely to occur than conventional warfare enhance trust among parties? And would this potentially lead to more or less verification methods being put in place?

This information was publicly available using the same sources I had to create my dataset. I was able to add a column to my data that showed what type of arms the agreement tried to control and used this in my regressions to see if any relationship was present.

# **6 Results**

## **6.1 Summary of Results**

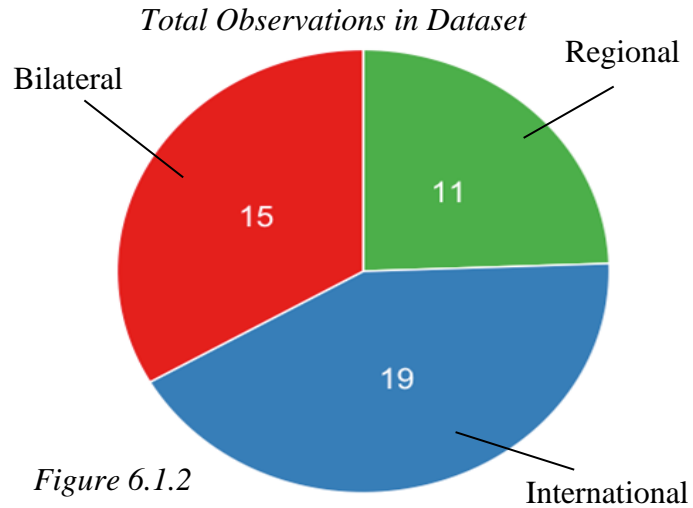
The overall results find that there is an interesting relationship between the number of state parties and total verification methods. While the initial scatterplot (shown in figure 6.1.5) demonstrates a slight negative relationship between the number of state parties and total verification methods, when the data is broken down to bilateral, regional, and international agreements, this outcome shows the opposite. Additionally, significant results I find when comparing each of the verification methods to the number of state parties is that there is a negative

relationship between the number of state parties and National Technical Means. I speculate as to why this may be later in my results. Another significant relationship I explore is the presence of a third-party verifier and how this affects Selective On-Site Inspection. I use multiple variable regressions to demonstrate these relationships and account for other statistically significant results.

Before presenting findings from the overall data, there are some descriptive statistics of the dataset I compiled that will help in providing some context of the variation in data that produced the results seen above. Figure 6.1.1 shows the distribution of bilateral and multilateral agreements along the total 45 observations in a single variable comparison. There are only 15 cases of bilateral agreements, providing some knowledge of the total verification methods they employed. However, this leads me to conclude that the total distribution will be a better assessor of overall patterns in the data. Figure 6.1.2 shows a visualization of this overall distribution, reflecting the fact that international agreements make most of the observations in my dataset, followed by bilateral agreements, and then regional agreements.

*Figure 6.1.1*

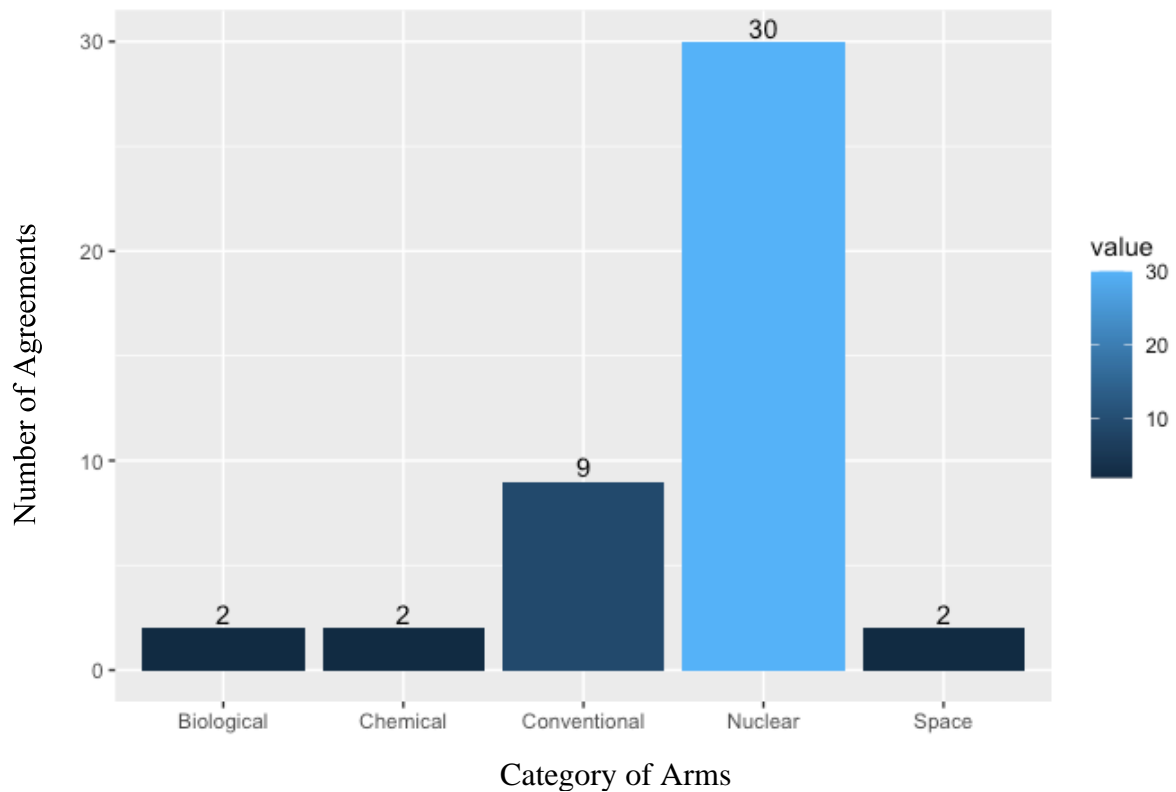
<b>Type of Agreement</b>	<b>Distribution</b>
Bilateral (2 actors)	15 (33.3%)
Multilateral (3+ actors)	30 (66.6%)
➤ Regional	➤ 11
➤ International	➤ 19
<b>Column Total</b>	<b>N = 45</b>



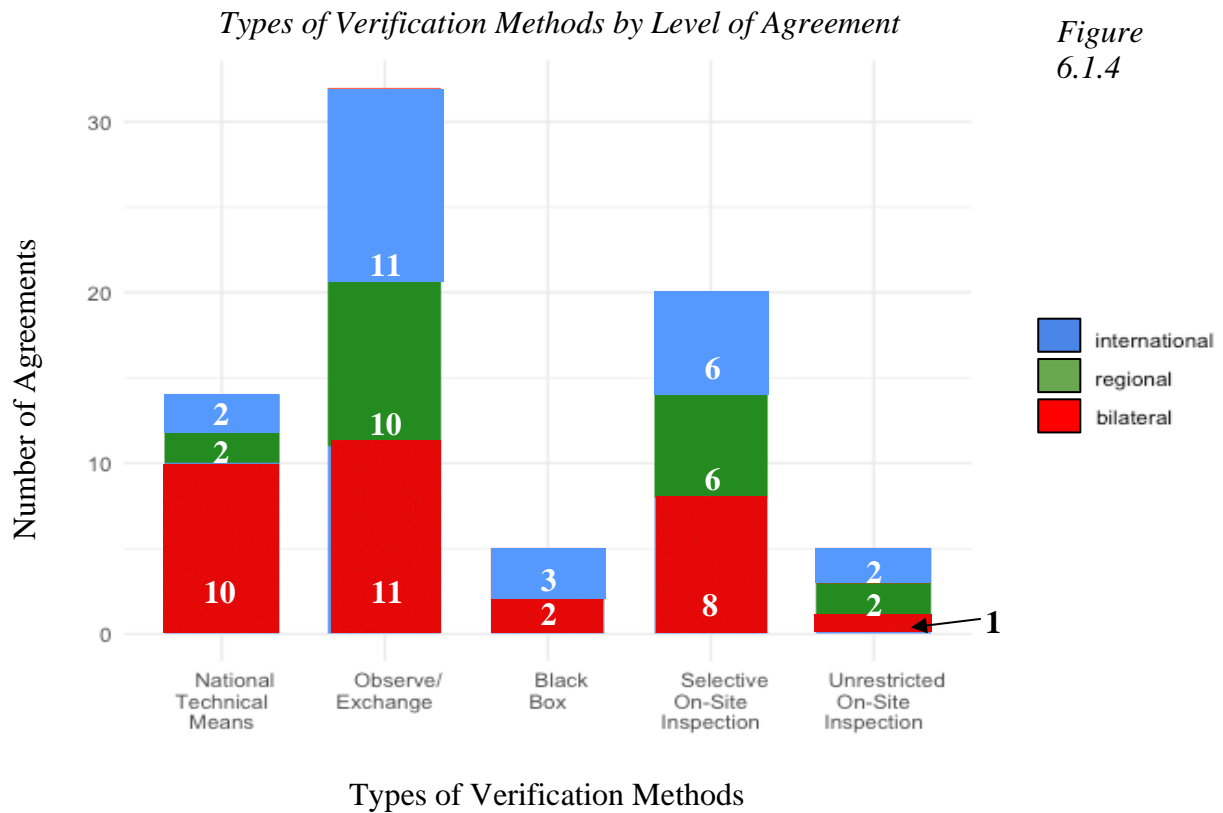
Another variable I used to compare the observations in my dataset was the type of arms each agreement sought to control. Figure 6.1.3 shows a visualization of the number of agreements that fit in each of 5 categories of arms: biological, chemical, conventional, nuclear, and space. What stands out most from this data is that the vast majority of arms agreements were nuclear agreements in nature. It is worth distinguishing between the types of agreements as verification methods may vary based on the type of arms that are being verified. Additionally, there may be minor differences in state parties' levels of commitment to agreements based on the type of arms that are being controlled - I address this in the section pertaining to limits.

Figure 6.1.3

Arms Type by Number of Agreements

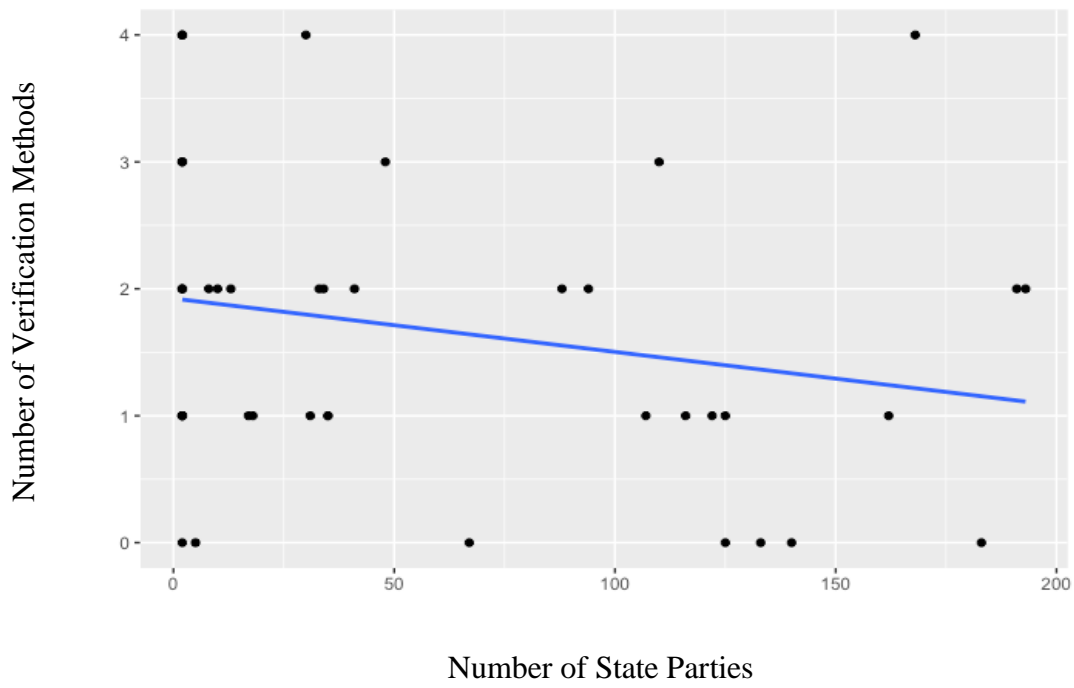


Another visualization that helps understand the dataset comparing my independent and dependent variable is Figure 6.1.4. In this stacked bar chart, I organize the observations or number of agreements by verification method on the X axis. I use five verification methods to produce the “total verification” variable that I use later to make regressions. In each of those five verification methods, the total number of agreements are organized by the BRI level. As represented in the legend, this is stacked by bilateral, regional, and international agreements. What is worthy of mentioning here is that the Observe and Exchange verification method was the most popular among the agreements, followed by Selective On-Site Inspection, and National Technical Means. When I later analyze the regressions, those three verification methods will be producing the most statistically significant results. The Black Box and Unrestricted On-Site Inspection verification methods, alternatively, seem to be the least used among the state parties.



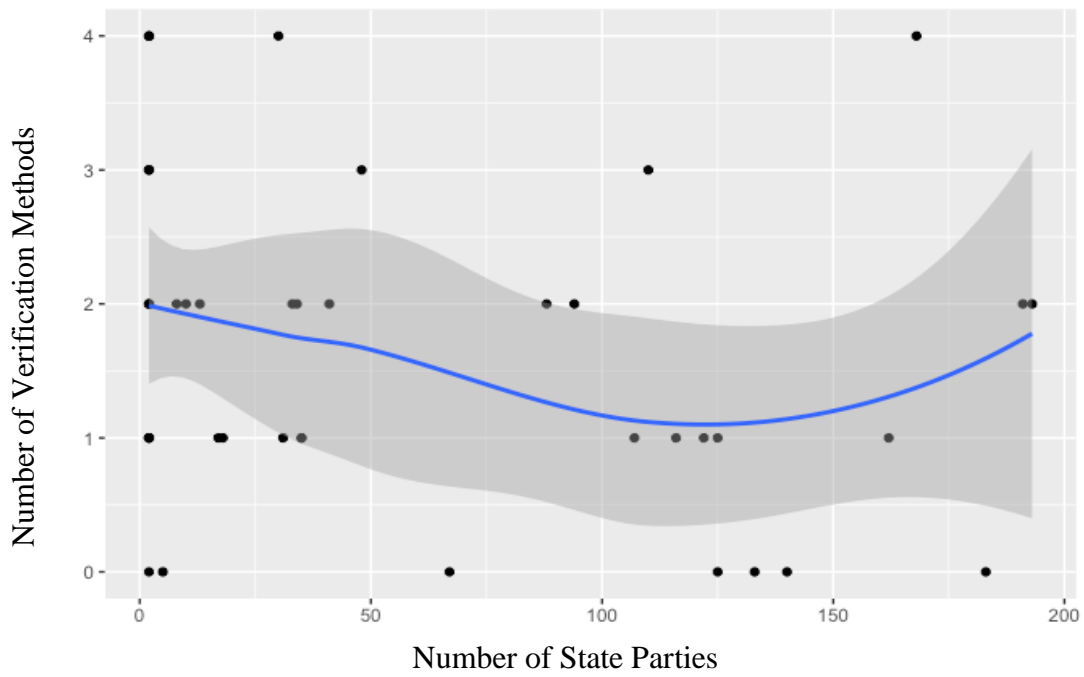
Figures 6.1.5 and 6.1.6 show scatterplots that represent the relationship between the total number of verification methods and the number of state parties in agreements. Figure 6.1.5 shows a regression line that has a slightly downward trend, indicating that as the number of state parties in an agreement increases, the number of verification methods decreases. This shows the opposite relationship between my two variables compared to what I had initially predicted. However, Figure 6.1.6 presents a more nuanced explanation. In this scatterplot, I use the LOESS method (Local Polynomial Regression) to fit a smooth curve between my independent and dependent variables. This allows for a more accurate regression line throughout the dataset. The LOESS method tells a slightly different story than the straight regression line as it shows a decline and then an upwards curve that indicates a positive relationship where the increasing state parties correlated with an increase in the verification methods.

*Total Verification Methods and Number of State Parties*



*Figure 6.1.5*

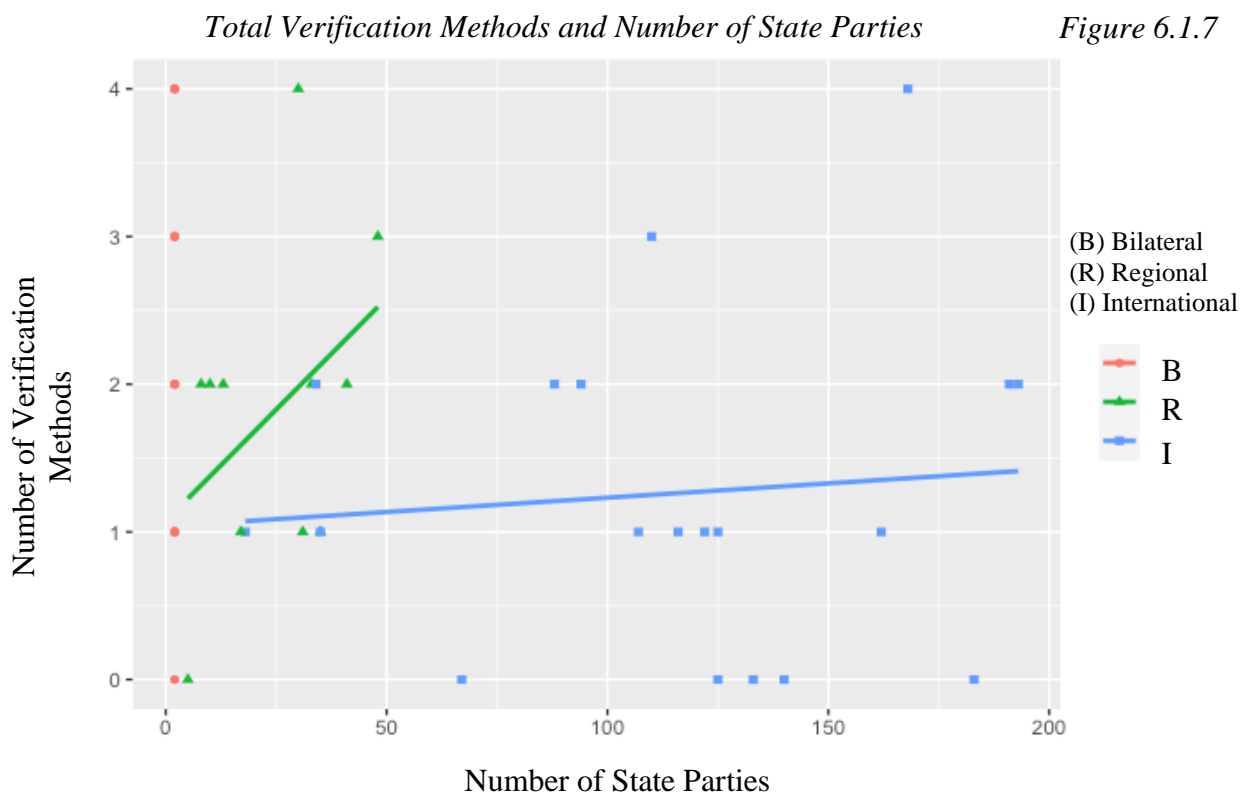
*Total Verification Methods and Number of State Parties  
(LOESS Method)*



*Figure 6.1.6*



To better understand the explanation as to why the number of verification methods trends downwards and then upwards, I created a legend for Bilateral, Regional, and International agreements. This allowed me to make separate regression lines for each of these three categories, showing an entirely different pattern than what was indicated in the first regression. Looking at Figure 6.1.7, the regression lines in regional and international agreements indicate a positive relationship between the number of parties and verification methods. Although Figure 6.1.7 does not show the number of bilateral agreements with each total verification method count, my dataset indicates there is a wide spread of verification methods along bilateral agreements (1 agreement has 0 verification methods, 5 agreements have 1 verification method, 3 agreements have 2 verification methods, and 3 agreements have 4 verification methods). The spread of bilateral agreements may be the reason why the overall regression line, without distinguishing the BRI level of agreement, tends to be downward sloping.



## 6.2 Regressions

The following regressions display the level of significance in the relationships introduced in the summary of results section. In the discussion section, I will give explanations of the significant regressions conducted as well as potential explanations for the following results.

### Linear Regressions

*Figure 6.2.1* Relationship between the number of parties and total verification methods:

```
Call:
lm(formula = total ~ numberparties, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-1.91539 -0.91539  0.08461  0.87999  2.78320

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.923809   0.237564   8.098 3.44e-10 ***
numberparties -0.004208   0.002841  -1.482   0.146
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.187 on 43 degrees of freedom
Multiple R-squared:  0.04857,    Adjusted R-squared:  0.02644
F-statistic: 2.195 on 1 and 43 DF,  p-value: 0.1458
```

This regression shows the relationship between my independent variable (the number of parties in an agreement) and the dependent variable (total verification methods). It indicates that as the number of parties in an agreement increases by 1, the total verification methods decreases by approximately 0.004. This reflects the initial regression line I found in Figure 6.1.5. The P-Value of 0.1458 indicates that this regression output has a 14.58% chance of error. Due to this P-Value, this regression output is not necessarily significant, but given its relevance to the earlier scatterplot results, it was important to include in this regression section for reference.

I also conducted a series of linear regressions indicating the relationship between the independent variable (the number of parties in an agreement) and each individual verification method that was used to compute my dependent variable (total verification methods). These five verification methods were (1) National Technical Means, (2) Information Observe and Exchange, (3) Black Boxes, (4) Selective On-Site Inspection, and (5) Unrestricted On-Site Inspection. Although the linear regressions comparing the number of parties and the last four verification methods proved to lack degrees of significance, the regression output between the number of parties and NTM was notable and I include it in Figure 6.2.2 below.

*Figure. 6.2.2* Relationship between the number of parties and National Technical Means:

```
Call:
lm(formula = ntm ~ numberparties, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.4315 -0.3666 -0.1765  0.5685  0.9399

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.436013   0.090412   4.822 1.81e-05 ***
numberparties -0.002237   0.001081  -2.070  0.0445 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4516 on 43 degrees of freedom
Multiple R-squared:  0.0906,    Adjusted R-squared:  0.06945
F-statistic: 4.284 on 1 and 43 DF,  p-value: 0.04452
```

This regression indicates that as the number of parties in an agreement increases by 1, the likelihood of that agreement including National Technical Means as a verification method decreases by 0.002237. This supports the regression line in Figure 6.1.5. The significant P-value of 0.04452\* indicates that this regression output has a 4.452% chance of error.

Another series of regressions I ran was to determine the relationship between the presence of a third-party verifier and each of the five individual verification methods behind the dependent variable. After running each regression, I found that the presence of a third party had some potentially significant effect on NTM and a significant effect on Selective On-Site Inspections. The purpose of this series of regressions was to find out whether a third-party verifier would increase or decrease the number of verifications agreed upon.

*Figure 6.2.3* Relationship between third party and NTM:

```
Call:
lm(formula = ntm ~ thirdparty, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.45  -0.20  -0.20   0.55   0.80

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.4500     0.1020   4.411 6.78e-05 ***
thirdparty   -0.2500     0.1369  -1.827  0.0747 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4562 on 43 degrees of freedom
Multiple R-squared:  0.072,    Adjusted R-squared:  0.05042
F-statistic: 3.336 on 1 and 43 DF,  p-value: 0.07471
```

Although this regression does not have a significant output, because it was close to being significant, I decided to include it in the results section. This regression indicates that when a third-party verifier is present, the chance that an agreement includes NTM decreases by 0.2500. The P-value of 0.07471 indicates that the regression output has a 7.471% chance of error. This indicates that the presence of a third-party decreases the likelihood of adopting National Technical Means as a verification method.

Figure 6.2.4 Relationship between third party and Selective On-Site Inspection:

```
Call:
lm(formula = selective ~ thirdparty, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.600 -0.250 -0.250  0.400  0.750

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.25000    0.1065   2.348  0.0235 *
thirdparty    0.35000    0.1429   2.450  0.0184 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4762 on 43 degrees of freedom
Multiple R-squared:  0.1225,    Adjusted R-squared:  0.1021
F-statistic: 6.003 on 1 and 43 DF,  p-value: 0.01843
```

This regression indicates that when a third-party verifier is present in an agreement, the likelihood of that agreement including Selective On-Site Inspection as a verification method increases by 0.3500. This supports the regression line in Figure 6.1.6 using the LOESS method as it demonstrates the upward-trending curve in the multilateral data. The significant P-value of 0.01843\* indicates that this regression output has a 1.843% chance of error.

I also ran another series of regressions comparing the CINC Score with the total verification methods as well as each individual verification method and was unable to find any results denoting statistical significance. However, I still include CINC Score as a part of my multiple regressions as they are a valuable factor to account for as a covariate.

### Multiple Regressions

After running multiple regressions on total verification methods and each individual verification method, I found three that produced statistically significant results. These were the

verification methods of NTM, Selective On-Site Inspection, and Unrestricted On-Site Inspection using the variables of total number of state parties, CINC Score, and Third-Party verifier.

*Figure 6.2.5* Relationship between NTM and Number of Parties, CINC Score, Third Party:

```
Call:
lm(formula = ntm ~ numberparties + cincscore + thirdparty, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.6923 -0.3036 -0.1601  0.3579  1.0289

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.463848   0.123275   3.763 0.000527 ***
numberparties -0.004907   0.001692  -2.900 0.005973 **
cincscore      0.585902   0.295571   1.982 0.054178 .
thirdparty    -0.242958   0.127304  -1.908 0.063346 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4242 on 41 degrees of freedom
Multiple R-squared:  0.2348,    Adjusted R-squared:  0.1789
F-statistic: 4.195 on 3 and 41 DF,  p-value: 0.01117
```

This regression shows how NTM is affected by three variables: the number of parties, CINC Score, and whether a third party is present. The results indicated the following:

- When the number of parties increases by 1, NTM decreases by 0.004907 with a 0.5973%\*\* chance of error in the regression output - holding constant the CINC Score and third party.
- When the CINC Score of an agreement increases by 1, NTM increases by 0.585902 with a 5.4178% chance of error in the regression output - holding constant the number of parties and third party.
- When there is a third-party verifier present, NTM decreases by 0.242958 with a 6.3346% chance of error in the regression output - holding constant the number of parties and CINC Score.

*Figure 6.2.6* Relationship between Selective On-Site Inspection and Number of Parties, CINC Score, Third Party Verifier:

```
Call:
lm(formula = selective ~ numberparties + cincscore + thirdparty,
    data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.7201 -0.3551 -0.1218  0.3835  0.9090

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.3720952   0.1385047   2.687   0.0104 *
numberparties  0.0005006   0.0019010   0.263   0.7936
cincscore     -0.3405413   0.3320874  -1.025   0.3112
thirdparty     0.3479058   0.1430319   2.432   0.0195 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4767 on 41 degrees of freedom
Multiple R-squared:  0.1616,    Adjusted R-squared:  0.1003
F-statistic: 2.634 on 3 and 41 DF,  p-value: 0.06257
```

This regression indicates how Selective On-Site Inspection is affected by three variables: the number of parties, CINC Score, and whether a third party is present. The results indicated the following:

- When the number of parties increases by 1, Selective OSI increases by 0.0005006 with a 79.36% chance of error in the regression output - holding constant the CINC Score and third party.
- When the CINC Score of an agreement increases by 1, Selective OSI decreases by 0.3405413 with a 31.12% chance of error in the regression output - holding constant the number of parties and third party.
- When there is a third-party verifier present, Selective OSI increases by 0.3479058 with a 1.95% chance of error in the regression output - holding constant the number of parties and CINC Score.

*Figure 6.2.7* Relationship between Unrestricted On-Site Inspection and Number of Parties, CINC Score, Third Party Verifier:

```
Call:
lm(formula = unrestricted ~ numberparties + cincscore + thirdparty,
    data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.35774 -0.17786 -0.04761  0.03232  0.80222

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.024465   0.085428  -0.286  0.77603
numberparties -0.002993   0.001173  -2.552  0.01452 *
cincscore      0.659477   0.204827   3.220  0.00251 **
thirdparty     0.025775   0.088220   0.292  0.77163
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.294 on 41 degrees of freedom
Multiple R-squared:  0.2026,    Adjusted R-squared:  0.1443
F-statistic: 3.473 on 3 and 41 DF,  p-value: 0.02449
```

This regression indicates how Selective On-Site Inspection is affected by three variables: the number of parties, CINC Score, and whether a third party is present. The results indicated the following:

- When the number of parties increases by 1, Unrestricted OSI decreases by 0.002993 with a 1.452% chance of error in the regression output - holding constant the CINC Score and third-party.
- When the CINC Score of an agreement increases by 1, Unrestricted OSI increases by 0.659477 with a 0.251% chance of error in the regression output - holding constant the number of parties and third-party.
- When there is a third-party verifier present, Unrestricted OSI increases by 0.025775 with a 77.163% chance of error in the regression output - holding constant the number of parties and CINC Score.



### 6.3 Discussion

In this discussion section, I will aim to give my explanation behind the varying scatterplot regression lines in 6.1 as well as each of the regressions produced in section 6.2.

First, I want to reiterate my initial hypothesis – I expected to see the number of total verification methods increase as the number of state parties in agreements increased. I argued that this was because of the costs and capability of actors, trust in other state parties to cooperate, and the guarantee that verification methods will be fairly adjudicated. One of the most surprising outcomes from my thesis was the difference in regression lines between the singular line that encompassed the entirety of my data and the several lines that would allow me to see the regressions at the bilateral, regional, and international level. In Figure 6.1.5, the regression line disputed my hypothesis by showing that as the number of state parties increased, the number of total verification methods decreased. However, Figure 6.1.7 showed an entirely different pattern. In this scatterplot with different regression lines by level of agreement, it was evident that the regional agreements' verification methods increased with a greater number of state parties and so did international agreements, albeit to a lesser degree.

At a theoretical level, there are several potential explanations for why regional and international agreements were so different compared to bilateral agreements. The one that I present has mainly to do with the history of bilateral agreements. As noted in the introduction section of my thesis, bilateral agreements became most prominent during the Cold War and are in most cases between the United States and Soviet Union (now Russia). Authors of arms control literature point out that the U.S. and USSR initially started bilateral agreements with the U.S. proposing more verification methods as a way to be more favorable in the public eye and the Soviet Union would often reject these, not requiring the U.S. to carry through with their

propositions.<sup>50</sup> However, as verification mechanisms became more cost-effective and cooperation became easier, the two countries agreed upon more verification. This especially occurred in the form of NTM being accepted by both sides as they had the necessary funds and technology to utilize this tool. This explanation I believe best answers why bilateral agreements are so spread out along the various total verification methods. Holding constant these bilateral agreements, we must then take a look at the multilateral agreements and the pattern they displayed.

I believe the trend showing both regional and international agreements increasing in verification methods as the number of state parties increased is due to the theoretical mechanisms I laid out in my hypothesis and theory section. However, this does not account for the difference between the regional and international agreements. By the nature of regional agreements, they are restricted to a certain number of countries within a region that can join. This holds true for arms control agreements as well – and this differs from international agreements as these are agreements like the Non-Proliferation Treaty or Geneva Convention that are open to all countries in the world with no geographical restrictions. I believe that, to a certain extent, Abbott and Snidal are correct in their argument and this may explain the difference between the two regression lines. They argue that as the number of verification methods (and therefore stringent restrictions) increases, there are more costs for a greater number of actors to sign onto.<sup>51</sup> They expect to see fewer verification methods as the number of actors in a multilateral agreement increases. While I do not think this theory reflects the whole of my data, I think it can explain the difference between regional and international agreements. This is because it is more likely that regional agreements with a cap on the number of countries (up to 50 parties in my dataset) are able to agree upon a set list of verification methods, whereas too much verification may deter agreements that have 100+ state

---

<sup>50</sup> Oelrich 1990.

<sup>51</sup> Abbott and Snidal 2000.

parties. In the latter case, it may be more difficult to get them to agree to sign onto an agreement with such a high cost of verification (both monetary and strategic costs). I will note, however, that my explanation is a preliminary one and there is not much research on this particular subject, therefore I believe for a more robust explanation, there would need to be more research done on this matter.

The next point of discussion is the explanation behind the statistically significant results and why they were concentrated around NTM, Restricted and Unrestricted OSI. To summarize the relevant results found in the regressions, I found that NTM decreases as the number of state parties increases, NTM decreases when a third party is present, Selective OSI increases with a third-party present, and Unrestricted OSI increases with higher CINC. The multiple regressions reinforced each of these points, holding constant other covariates that I had considered.

First, I will address why NTM decreased with an increase in state parties whereas this was not the case with other verification methods. While a multitude of arguments could explain this, the strongest of these is presented by Koplow, who states that bilateral arms control agreements tend to participate in reciprocal inspections whereas multilateral agreements tend to create or co-opt international organizations in order to conduct inspections for their agreements.<sup>52</sup> Using this logic, we may expect to see bilateral agreements as entities most likely to use reciprocal inspection such as NTM. Additionally, as referenced earlier, the majority of bilateral agreements are between the U.S. and Russia, both countries that are capable of funding their own technical monitoring programs and countries that are highly skeptical of one another and hesitant to share information. It makes sense that NTM would therefore be most used in bilateral agreements but be too costly to implement in multilateral agreements where a country does not have the domestic capability to

---

<sup>52</sup> Koplow 2017.

produce the technologies necessary for doing monitoring alone. This may also play a part in the explanation as to why NTM tends to decrease when a third party is present. Perhaps this is not necessarily because of the third party directly but because a third party is more likely present in a larger multilateral agreement, and these tend to not have NTM as frequently as bilateral agreements. Additionally, when a third party is present – as I will discuss in a moment – it is able to co-opt the costs for a series of verification methods. While a third party can perform a series of tasks in verification, it cannot provide National Technical Means to other countries. This is meant to be a uniquely domestic capability that states may not need when other more inexpensive verification methods are being outsourced to a third party. Again, this topic also does not have a large body of research behind it, but given the observations I have in my dataset, that is the most likely explanation I can derive from the regression.

Next, a trend found was that Selective On-Site Inspection tended to increase when a third party was present. This is closer to what I expected with my hypothesis. This, too, has a potential explanation that is tied to the history of that particular verification method. Oelrich explains that over time, On-Site Inspection became harder because it no longer consisted of just “annoying paperwork” and it started to aggravate actors due to technical concerns that arose.<sup>53</sup> These included, for example, identifying tags attached to weapons, as close On-Site Inspection can reveal enough information about a weapon, allowing for an enemy to better understand how to defeat it. Naturally, it makes a lot of sense, as I predicted in part of my theory section, that when the costs of a verification method are very high, a third party will be a valuable asset in lowering these costs for state parties. Therefore, when the technical execution of Selective On-Site Inspection needs to be thoroughly thought-out as to not reprimand any actors and diminish their defenses, a positive

---

<sup>53</sup> Oelrich 1990.

relationship between third-party presence and state parties' willingness to adopt Selective OSI is a logical connection.

One final statistically significant outcome was that an increase in CINC Score played a part in increasing Unrestricted On-Site Inspection. While there is no literature to speak on this, I can speculate that Unrestricted OSI, as an extremely intrusive verification method, is more likely used in agreements where countries are not going to lose anything by being completely transparent about their particular arms program. This is more likely the case for certain agreements such as space weaponization or biological weapons that countries take seriously and cases where they do not have an incentive to hide covert programs. The highest CINC scores across my dataset were attributed to agreements that were in this category – agreements that typically encompassed most state actors because they were the least controversial (i.e. Geneva Protocol, Outer Space Treaty, NPT, etc.). This is something that future researchers may want to look into as they analyze the incentives behind Unrestricted OSI more in-depth.

#### **6.4 Scope, Limitations, and Suggestions for the Future**

The first limitation in my thesis is the number of cases. Although I coded for every arms control agreement until 2010, there were only a total of 45. This leaves a large margin of error due to external factors that may contribute to the verification methods agreed upon. This came to fruition in some of my explanations in the discussion section that were centered around the history of verification methods and how they have changed over time or how countries' responses to these methods have changed over time. This also therefore limits the scope of my research as it is less capable of being applied in a policy setting. Although some relationships were found, this does not guarantee an outcome where state parties directly affect the number of verification methods consistently. Over time, I would like to see my data be replicated and extended to include many more agreements to allow for more robust conclusions.

Another limitation I had as a student with nearly 20 weeks to complete this thesis was creating a database from square one. Although the SIPRI database I modeled my verification method count from provided a good starting point, there were only a handful of observations in it. Additionally, information regarding the verification methods in each arms control agreement is not kept in a centralized location and therefore I had to use multiple different resources to determine the verification methods in each. Some of this was written in a bullet point format while others were written in a long treaty text of several pages. Although I did my best to remain unbiased, a lot of the analysis I had to do in order to create my database was subjective as agreements often had different terms for actions that would accomplish the same goal as the verification method I was accounting for. Additionally, over time, verification methods have changed, and I would have liked to account for changes in technologies that states were able to develop and refine from 1925 to 2010. If this research is replicated, I believe this is a necessary next step in order to make sure the database is more comprehensive and detailed to include the different verification methods that I may not have accounted for.

One covariate I especially liked was the CINC Score. However, I do not think it was all-encompassing for the many factors that make countries different from one another. For example, the technical capabilities of countries may have played a part – but so may have their commitment to the agreement. It is fair to assume, for example, that India or Pakistan may be more committed to holding an adversary accountable to an arms control agreement whereas Switzerland and Sweden would be less committed to this. In this example, it is possible that the countries also had the same CINC scores (hypothetically). This is another flaw in my data as I would have liked to incorporate how important a country perceived arms control to be for their national security. Ultimately, calculating CINC scores alone took nearly two weeks of calculations, therefore it would

not be feasible to calculate each country's perceived level of threat in a time-sensitive thesis seminar. However, I would encourage future political scientists to explore that potential covariate.

Finally, I believe another limitation in my research was the lack of comprehensive literature surrounding the differences between bilateral, regional, and international agreements. This approach that I took was a novel one in arms control research and therefore there was no existing body of research to use in order to fully analyze and understand the differences between the three types of agreements. I think this will partially be aided by more observations in the future, particularly bilateral agreements not consisting of the U.S. and USSR, to make more well-rounded conclusions from the overall data. More insight into the inner workings of regional agreements and how they differ from international agreements would also be interesting to add to my analysis as the initial regressions I make present some interesting findings to start with.

## **7 Conclusion**

What began as a simple question of “what enhances commitment to arms control agreements” became an interesting topic for a thesis - encompassing verification regimes and how state parties contribute to their adoption. However, it is only fitting that a research question is answered with several more questions, as is the case with my thesis. While overall trends were found in my research, with the increasing number of state parties causing a slight decrease in total verification methods - breaking this down by type of agreement (bilateral, regional, and international) posed many more puzzles that I needed to solve.

Additionally, variation in the level of significance by verification method hinted at the fact that they are not necessarily all the same and they have nuanced reasoning behind each instead of fitting into one category of “verification regime” as political science has thus far classified them. Taking one significant result at a time, I was able to present some theories using relevant literature

and my own regression analysis, allowing me to propose some answers as to why these may vary. Ultimately, while neither of my two hypotheses were perfectly accurate, my regressions formulated some confidence in the first hypothesis that especially in multilateral agreements, as the number of state parties increase, so do verification methods. I also found that with regards to my second hypothesis, while there was no quantitative mechanism to compare the intrusiveness in verification regimes, multilateral agreements tended to adopt more Selective OSI whereas bilateral agreements adopted more NTM, denoting that multilateral agreements with a greater number of state parties were more likely to adopt more intrusive verification methods.

Creating a dataset using raw data from various sources and connecting them into one cohesive story was quite a challenge - but proved to be worth the effort in the end as I was able to come up with some unique conclusions that had not been explored before in the field of political science. I hope that with this research, future scholars will spend more time trying to connect how a verification regime works and what enhances it.

Social science by its nature is vague and difficult to create concrete conclusions with, but as we learn more about topics like arms control, we are better able to prepare policy solutions and understand what to implement in order to make a safer and more cohesive global network. As we move forward into a future with more multilateral agreements and hopefully enhanced state cooperation, it will be interesting to see how arms control agreements play out and whether we expect to see more or less verification as a byproduct of the trust between state parties. Looking back in history, we can recognize that a change needs to be made to the proverb “trust, but verify” and rather emphasize that the lack of trust, in fact, is what leads to the verification regimes that potentially deter future inter-state conflict and maintain peace in the modern era.



## 8 References

Abbott, Kenneth W., and Duncan Snidal. "Hard and Soft Law in International Governance."

*International Organization*, vol. 54, no. 3, 2000, pp. 421-456.

"Arms Control and Disarmament - History." *History - Weapons, Nuclear, War, and Power -*

*JRank Articles*, JRank Law Library, 2021, [law.jrank.org/pages/4438/Arms-Control-](http://law.jrank.org/pages/4438/Arms-Control-Disarmament-History.html)

[Disarmament-History.html](http://law.jrank.org/pages/4438/Arms-Control-Disarmament-History.html).

Axworthy, Thomas, and Ryan Dean. "A Scan of Arms Control Treaties with Lessons Learned."

*InterAction Council*, May 2011.

Caporaso, James A. "International Relations Theory and Multilateralism: The Search for

Foundations." *International Organization*, vol. 46, no. 3 (Summer, 1992), pp. 599-632.

Carter, April. "Success and Failure in Arms Control Negotiations." *SIPRI*, 1989,

[www.sipri.org/publications/1989/success-and-failure-arms-control-negotiations](http://www.sipri.org/publications/1989/success-and-failure-arms-control-negotiations).

Correlates of War Project. *National Material Capabilities*. (Version 5.0). 2012.

<https://correlatesofwar.org/data-sets/national-material-capabilities>

Dahlman, Ola. "Verification: to Detect, to Deter and to Build Confidence." *Disarmament Forum*

*/ UNIDIR*, vol. 3, 2010, pp. 3-13.

den Dekker, Guido. "The Effectiveness of International Supervision in Arms Control

Law." *Journal of Conflict and Security Law*, vol. 9, no. 3, 2004, pp. 315-330.,

[doi:10.1093/jcsl/9.3.315](https://doi.org/10.1093/jcsl/9.3.315).

"Disarmament Treaties Database." *UNODA*, United Nations Office of Disarmament Affairs,

2021.

Downs, George W., and Rocke, David M., and Barsoom, Peter N. "Is the Good News About

Compliance Good News?" *International Organization* 50, 3, Summer 1996, pp. 379-406.

- “Fact Sheets and Briefs.” *Arms Control Association*, 2021.
- “Fact Sheets and Treaty Text.” *Nuclear Threat Initiative*, 2021.
- Geller, Daniel S. “Nuclear Weapons and the Indo-Pakistani Conflict: Global Implications of a Regional Power Cycle.” *International Political Science Review*, vol. 24, no. 1, Jan. 2003, pp. 137-150., doi:10.1177/0192512103024001009.
- “Hard Law/Soft Law.” *ECCHR*, European Center for Constitutional and Human Rights E.V., 2021, [www.ecchr.eu/en/glossary/hard-law-soft-law](http://www.ecchr.eu/en/glossary/hard-law-soft-law).
- Ifft, Edward M. “Verifying Nuclear Arms Control and Disarmament.” *Verification Yearbook*, 2001, pp. 25-42.
- Jervis, Robert. “The Dustbin of History: Mutual Assured Destruction.” *Foreign Policy*, 9 Nov. 2009, [foreignpolicy.com/2009/11/09/the-dustbin-of-history-mutual-assured-destruction/](http://foreignpolicy.com/2009/11/09/the-dustbin-of-history-mutual-assured-destruction/).
- Keohane, Robert O. “After Hegemony: Cooperation and Discord in the World Political Economy, Chapter 4.” *After Hegemony: Cooperation and Discord in the World Political Economy*, Princeton University Press, 2005, pp. 49-64.
- Koplow, David A. “Eve of Destruction: Implementing Arms Control Treaty Obligations to Dismantle Weaponry.” *Harvard National Security Journal*, vol. 8, 2017, pp. 158-235.
- Koremenos, Barbara and Lipson, Charles and Snidal, Duncan. “The Rational Design of International Institutions.” *International Organization*, Autumn, 2001, vol. 55, no. 4, pp. 761-799.
- Krass, Allan S. “The Costs, Risks, and Benefits of Arms Control.” *Center for International Security and Arms Control*, Feb. 1996.
- Krass, Allan S. “Verification: How Much Is Enough?” *Stockholm International Peace Research Institute*, July 1985.

- Maurer, John D. "The Purposes of Arms Control." *TexasScholarWorks*, Texas National Security Review, November 2018, [dx.doi.org/10.26153/tsw/870](https://doi.org/10.26153/tsw/870).
- Meyer, Stephen M. "Verification and Risk in Arms Control." *International Security*, vol. 8, no. 4, 1984, pp. 111-126., [doi:10.2307/2538565](https://doi.org/10.2307/2538565).
- Moodie, Michael, and Amy Sands. "Introduction: New Approaches to Compliance with Arms Control and Nonproliferation Agreements." *The Nonproliferation Review*, vol. 8, no. 1, 2001, pp. 1-9., [doi:10.1080/10736700108436834](https://doi.org/10.1080/10736700108436834).
- Müller, Harald. "Compliance Politics: A Critical Analysis of Multilateral Arms Control Treaty Enforcement." *The Nonproliferation Review*, vol. 7, no. 2, 2000, pp. 77-90., [doi:10.1080/10736700008436811](https://doi.org/10.1080/10736700008436811).
- Nakamitsu, Izumi. "Half of All Violent Deaths Involve Small Arms and Light Weapons." *UN News*, United Nations, 5 Feb. 2020, [news.un.org/en/story/2020/02/1056762](https://news.un.org/en/story/2020/02/1056762).
- Norris, Robert S., and Hans M. Kristensen. "Global Nuclear Weapons Inventories, 1945-2010." *Bulletin of the Atomic Scientists*, vol. 66, no. 4, July 2010, pp. 77-83, [doi:10.2968/066004008](https://doi.org/10.2968/066004008).
- "Nuclear Safety Conventions." *IAEA*, International Atomic Energy Agency, 2021.
- "Nuclear Weapons: Which Countries Have Them and How Many Are There?" *BBC News*, BBC, 14 Jan. 2020, [www.bbc.com/news/newsbeat](https://www.bbc.com/news/newsbeat).
- Oelrich, Ivan. "The Changing Rules of Arms Control Verification: Confidence is Still Possible." *International Security*, Spring, 1990, vol. 14, no. 4 (Spring, 1990), pp. 176-184.
- Pieragostini, Karl. "Arms Control Verification: Cooperating to Reduce Uncertainty." *The Journal of Conflict Resolution*, Sep., 1986, vol. 30, no. 3 (Sep., 1986), pp. 420-444.

- “Report on Options for Confidence and Security Building Measures (CSBMs), Verification, Non-Proliferation, Arms Control and Disarmament.” *PRESS COMMUNIQUE M-NAC 2(2000)121*, North Atlantic Treaty Organization, 14 Dec. 2000.
- Schaper, Annette. “Verifying Nuclear Arms Control and Disarmament.” *Verification Yearbook*, 2000, pp. 57–69.
- Schofield, Julian. “Arms Control Failure and the Balance of Power.” *Canadian Journal of Political Science / Revue Canadienne De Science Politique*, vol. 33, no. 4, Dec. 2000, pp. 747–777.
- Schroeder, Matt, and Rachel Stohl. “Arms Control Today.” *Small Arms, Large Problem: The International Threat of Small Arms Proliferation and Misuse / Arms Control Association*, Arms Control Association, June 2006, [www.armscontrol.org/act/2006-06/features/small-arms-large-problem-international-threat-small-arms-proliferation-misuse](http://www.armscontrol.org/act/2006-06/features/small-arms-large-problem-international-threat-small-arms-proliferation-misuse).
- “The Nuclear Threat.” *Nuclear Threat Initiative - Ten Years of Building a Safer World*, 31 Dec. 2015, [www.nti.org/learn/nuclear/](http://www.nti.org/learn/nuclear/).
- Thompson, Kenneth W. "Arms control". *Encyclopedia Britannica*, 9 Aug. 2016, <https://www.britannica.com/topic/arms-control>.
- “Treaties Database.” *Arms Control Archives*, United States Department of State.
- Verdier, Daniel. “Multilateralism, Bilateralism, and Exclusion in the Nuclear Proliferation Regime.” *International Organization*, vol. 62, no. 3, 2008, pp. 439–476., doi:10.1017/s0020818308080156.
- Williamson, Richard L. Jr. "Hard Law, Soft Law, and Non-Law in Multilateral Arms Control: Some Compliance Hypotheses," *Chicago Journal of International Law*, vol. 4, no. 1, Article 7, 2003.

## 9 Appendix

The Appendix consists of the following data representations:

- My original dataset (in 3 pages due to spacing limits) .....62
  - Page 1: Name of Treaty and Year Signed
  - Page 2: Countries (cut off in some cases because it would not fit on the page)
  - Page 3: All other variables
    - Number of parties
    - CINC score
    - BRI level:  
1 = bilateral, 2 = regional, 3 = international
    - Type of arms:  
1 = conventional, 2 = biological, 3 = chemical, 4 = nuclear, 5 = space
    - Most Intrusive Verifier Count (“highest”) – not used
    - Total verification count
    - Each of the verification methods:  
0 = not in agreement, 1 = in agreement
    - Third party presence:  
0 = no third-party, 1 = third-party present
    - Average CINC Score – also not used

*Please email me for access to the dataset if you are interested.*

- Visualizations of Data: Figures 6.1.1 – 6.1.4 .....65
- Scatterplots: Figures 6.1.5 – 6.1.7 .....67
- Regression Outputs: Figures 6.2.1 – 6.2.7 .....68

	<b>Name.of.Treaty</b>	<b>Year.Signed</b>
1	Geneva Protocol	1925
2	Antarctic Treaty	1959
3	Partial Test Ban Treaty PTBT	1963
4	Outer Space Treaty	1967
5	Treaty of Tlatelolco	1967
6	NPT	1968
7	Sea-Bed Treaty	1971
8	BWC	1972
9	ABM Teaty	1972
10	SALT I	1972
11	Threshold Test Ban Treaty	1974
12	CSCE Final Act, Helsinki	1975
13	PNE Treaty	1976
14	Environmental Modification Convention	1977
15	Moon Treaty	1979
16	SALT II	1979
17	Convention on the Physical Protection of Nuclear Material	1980
18	Convention on Certain Conventional Weapons CCW	1981
19	Treaty of Rarotonga: South Pacific Nuclear Free Zone	1985
20	Missile Technology Control Regime MTCR	1987
21	INF Treaty	1987
22	India-Pakistan Non-Attack Agreement	1988
23	Treaty on Conventional Armed Forces in Europe CFE	1990
24	START I	1991
25	Open Skies Treaty	1992
26	India-Pakistan Agreement on Chemical Weapons	1992
27	Joint Declaration of North and South Korea on the Denuclearization of the Korean Peninsula	1992
28	START II	1993
29	CWC	1993
30	Framework between the US and DPRK	1994
31	Convention on Nuclear Safety	1994
32	Southeast Asian Nuclear-Weapon-Free-Zone (SEANWFZ) Treaty (Bangkok Treaty)	1995
33	CTBT	1996
34	African NW Free Zone Treaty, Pelindaba Treaty	1996
35	Ottawa Treaty/Mine Ban Treaty	1997
36	Inter-American Convention Against the Illicit Manufacturing of and Trafficking in Firearms, Ammunition, Explosives, etc.	1997
37	Lahore Declaration	1999
38	Inter-American Convention on Transparency in Conventional Weapons Acquisitions	1999
39	SORT - Strategic Offensive Reductions Treaty	2002
40	International Code of Conduct Against Ballistic Missile Proliferation ICOC	2002
41	International Convention for the Suppression of Acts of Nuclear Terrorism	2005
42	Central Asia Nuclear Weapons Free Zone CANWFZ	2006
43	Convention on Cluster Munitions CCM	2008
44	Central African Convention - Kinsaha	2010
45	New START	2010

Countries
Afghanistan (0.0035203), Albania (0.0003666), Algeria (0.0040212), Angola (0.001839), Antigua and Barbuda (2.79E-06), Argentina (0.0076569), Australia (0.007786),
Argentina (0.0069331), Australia (0.0070529), Belgium (0.0073392), Brazil (0.0141299), Bulgaria (0.0027232), Chile (0.0022294), China (0.1169668), Czechia (0.0111),
Afghanistan (0.0016639), Antigua and Barbuda (2.69E-06), Argentina (0.0068134), Armenia (0.000754), Australia (0.006986), Austria (0.0035973), Bahamas (0.0000),
Afghanistan (0.0014258), Algeria (0.0018977), Antigua and Barbuda (2.69E-06), Argentina (0.006574), Armenia (0.0006372), Australia (0.0077712), Austria (0.00299),
Antigua and Barbuda (2.88E-06), Argentina (0.006574), Bahamas (0.0003502), Barbados (0.0000174), Belize (0.0000176), Bolivia (0.0005816), Brazil (0.0176207), C
Afghanistan (0.0013737), Albania (0.000628), Algeria (0.0018951), Andorra (2.5E-06), Angola (0.0014602), Antigua and Barbuda (2.82E-06), Argentina (0.0064259),
Afghanistan (0.0014013), Algeria (0.0019072), Antigua and Barbuda (2.69E-06), Argentina (0.0061449), Australia (0.0075723), Austria (0.0030118), Bahamas (0.000),
Afghanistan (0.0014494), Albania (0.0006421), Algeria (0.0019503), Andorra (2.39E-06), Angola (0.0024164), Antigua and Barbuda (5.39E-06), Argentina (0.0060021),
US (0.1647432), USSR (0.1708357)
US (0.1647432), USSR (0.1708357)
US (0.1526986), USSR (0.1678685)
Austria (0.0028151), Belgium (0.006466), Bulgaria (0.0040516), Canada (0.0117008), Cyprus (0.0002015), Czechoslovakia (0.0095648), Denmark (0.0015658), Finl
US (0.1424627), USSR (0.1760274)
Afghanistan (0.0018156), Algeria (0.002199), Antigua and Barbuda (2.69E-06), Argentina (0.0061343), Australia (0.0073455), Austria (0.0026624), Bangladesh (0.00
Armenia (0.0007098), Australia (0.0071011), Belgium (0.0067073), Chile (0.0027443), Kazakhstan (0.0035262), Kuwait (0.0010842), Lebanon (0.0004617), Mexico (0
US (0.1372161), USSR (0.1673254)
Afghanistan (0.0011965), Albania (0.0006006), Algeria (0.0029125), Andorra (2.72E-06), Angola (0.0009789), Antigua and Barbuda (2.46E-06), Argentina (0.006531),
Afghanistan (0.0012083), Albania (0.0005695), Algeria (0.0028747), Antigua and Barbuda (2.46E-06), Argentina (0.0065189), Australia (0.0075), Austria (0.0026226)
Australia (0.007018), Cook Islands, Fiji (0.0000774), Kiribati (2.84E-06), Nauru (4.08E-07), New Zealand (0.0006309), Niue, Papua New Guinea (0.000191), Samoa (0
Argentina (0.0056297), Australia (0.0068841), Austria (0.0026306), Belgium (0.0049726), Brazil (0.0256679), Bulgaria (0.0036501), Canada (0.0112496), Czech Rep
US (0.1340296), USSR (0.173181)
India (0.058537), Pakistan (0.0099851)
Armenia (0.0006818), Azerbaijan (0.0008929), Belarus (0.0021041), Belgium (0.0050293), bulgaria (0.0028196), Canada (0.0109736), Czech Republic (0.0073517),
US (0.1372464), USSR (0.1033478)
Belgium (0.0051323), Bulgaria (0.0023204), Belarus (0.0029166), Canada (0.0120168), Czech and Slovak Federal Republic (0.0046325), Denmark (0.0016568), Fra
India (0.0633383), Pakistan (0.0119537)
North Korea (0.0139724), South Korea (0.0219628)
US (0.1535161), USSR (0.0580243)
Afghanistan (0.0014201), Albania (0.0006413), Algeria (0.0038948), Andorra (2.31E-06), Antigua and Barbuda (2.58E-06), Argentina (0.0055883), Armenia (0.000491),
US (0.1456364), DPRK (0.0120823)
Albania (0.0006439), Angola (0.0015878), Argentina (0.0057977), Armenia (0.000754), Australia (0.0078339), Austria (0.0025891), Bahrain (0.0004811), Bangladesh
Brunei Darussalam (0.0001357), Cambodia (0.0010541), Indonesia (0.0144839), Lao Peple's Democratic Republic (0.000591), Malaysia (0.0038936), Myanmar (0.0
Afghanistan (0.0011902), Albania (0.0005686), Algeria (0.0039661), Andorra (2.55E-06), Angola (0.0014602), Antigua and Barbuda (2.94E-06), Argentina (0.006023),
Algeria (0.0039661), Angola (0.0014602), Benin (0.0003601), Botswana (0.0001795), Burkina Faso (0.000543), Burundi (0.0004309), Cabo Verde (0.0000203), Cam
Afghanistan (0.0012208), Albania (0.0005676), Algeria (0.0039644), Andorra (2.63E-06), Angola (0.0015193), Antigua and Barbuda (2.71E-06), Argentina (0.006090),
Antigua and Barbuda (2.71E-06), Argentina (0.0060907), Bahamas (0.0000347), Belize (0.0000194), Bolivia (0.0009079), Brazil (0.0263278), Chile (0.0032887), Col
India (0.0662212), Pakistan (0.0132955)
Argentina (0.0061934), Barbados (0.0000124), Brazil (0.0257628), Canada (0.0114702), Chile (0.0031783), Costa Rica (0.0002855), Dominican Republic (0.000858),
US (0.1537646), USSR (0.049657)
Afghanistan (0.00158), Albania (0.0003992), Andorra (2.55E-06), Antigua and Barbuda (3.15E-06), Argentina (0.0058882), Armenia (0.0007098), Australia (0.007516),
Afghanistan (0.0012113), Algeria (0.00385), Antigua and Barbuda (3.18E-06), Argentina (0.0058913), Armenia (0.0006268), Australia (0.0075398), Austria (0.002478),
Kazakhstan (0.0031166), Kyrgyzstan (0.0003552), Tajikistan (0.0003442), Turkmenistan (0.0007561), Uzbekistan (0.002279)
Afghanistan (0.0014628), Albania (0.0002812), Andorra (2.63E-06), Antigua and Barbuda(3.24E-06), Australia (0.0074601), Austria (0.0023804), Belgium (0.003768),
Angola (0.002309), Cameroon (0.0011671), CAR (0.0002013), Chad (0.0006173), Congo (0.000423), Equatorial Guinea (0.0000527), Gabon (0.0001889), Sao Tome
US (0.1480978), USSR (0.0394878)

numberparties	cincscore	brilevel	typearms	highest	total	ntm	observeexchange	blackbox	selective	unrestricted	thirdparty	standardcinc
133	1.144	3	2	0	0	0	0	0	0	0	1	0.0086
48	0.902	2	4	5	3	1	1	0	0	1	0	0.01879
125	0.846	3	4	1	1	1	0	0	0	0	0	0.00677
110	0.987	3	5	5	3	0	1	0	1	1	0	0.00897
33	0.056	2	4	4	2	0	1	0	1	0	1	0.0017
191	0.951	3	4	4	2	0	0	1	1	0	1	0.00498
94	0.886	3	4	5	2	0	1	0	0	1	1	0.00943
183	1.004	3	2	0	0	0	0	0	0	0	0	0.00549
2	0.336	1	4	2	2	1	1	0	0	0	1	0.168
2	0.336	1	4	1	1	1	0	0	0	0	0	0.168
2	0.321	1	4	4	3	1	1	0	1	0	0	0.1605
35	0.547	2	4	2	1	0	1	0	0	0	1	0.01563
2	0.318	1	4	4	3	1	1	0	1	0	1	0.159
67	0.711	3	1	0	0	0	0	0	0	0	1	0.01061
18	0.083	3	5	2	1	0	1	0	0	0	1	0.00461
2	0.305	1	4	2	2	1	1	0	0	0	0	0.1525
162	0.969	3	4	2	1	0	1	0	0	0	1	0.00598
125	0.942	3	1	0	0	0	0	0	0	0	0	0.00754
13	0.008	2	4	4	2	0	1	0	1	0	1	0.00062
35	0.683	3	4	2	1	0	1	0	0	0	0	0.01951
2	0.307	1	4	5	4	1	1	0	1	1	1	0.1535
2	0.069	1	4	2	1	0	1	0	0	0	0	0.0345
30	0.483	2	1	5	4	1	1	0	1	1	1	0.0161
2	0.241	1	4	4	4	1	1	1	1	0	0	0.1205
34	0.422	3	1	4	2	0	0	1	1	0	1	0.01241
2	0.075	1	3	0	0	0	0	0	0	0	0	0.0375
2	0.036	1	4	4	1	0	0	0	1	0	1	0.018
2	0.212	1	4	4	3	1	1	0	1	0	1	0.106
193	0.946	3	3	4	2	0	1	0	1	0	1	0.0049
2	0.158	1	4	4	1	0	0	0	1	0	1	0.079
88	0.88	3	4	4	2	0	1	0	1	0	1	0.01
10	0.049	2	4	4	2	0	1	0	1	0	1	0.0049
168	0.566	3	4	4	4	1	1	1	1	0	0	0.00337
41	0.043	2	4	4	2	0	1	0	1	0	1	0.00105
122	0.449	3	1	2	1	0	1	0	0	0	1	0.00368
31	0.072	2	1	2	1	0	1	0	0	0	0	0.00232
2	0.08	1	4	2	1	0	1	0	0	0	0	0.04
17	0.075	2	1	2	1	0	1	0	0	0	0	0.00441
2	0.203	1	4	2	2	1	1	0	0	0	0	0.1015
140	0.654	3	4	0	0	0	0	0	0	0	0	0.00467
116	0.869	3	4	2	1	0	1	0	0	0	1	0.00749
5	0.007	2	4	0	0	0	0	0	0	0	1	0.0014
107	0.249	3	1	2	1	0	1	0	0	0	0	0.00233
8	0.005	2	1	4	2	0	1	0	1	0	1	0.00063
2	0.188	1	4	4	4	1	1	1	1	0	0	0.094



## Visualizations of Data: Figures 6.1.1 – 6.1.4

*Figure 6.1.1*

Type of Agreement	Distribution
Bilateral (2 actors)	15 (33.3%)
Multilateral (3+ actors)	30 (66.6%)
➤ Regional	➤ 11
➤ International	➤ 19
<b>Column Total</b>	<b>N = 45</b>

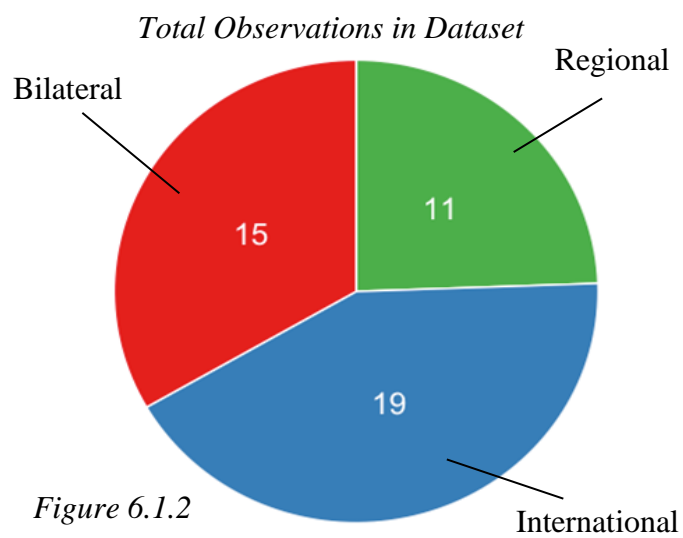
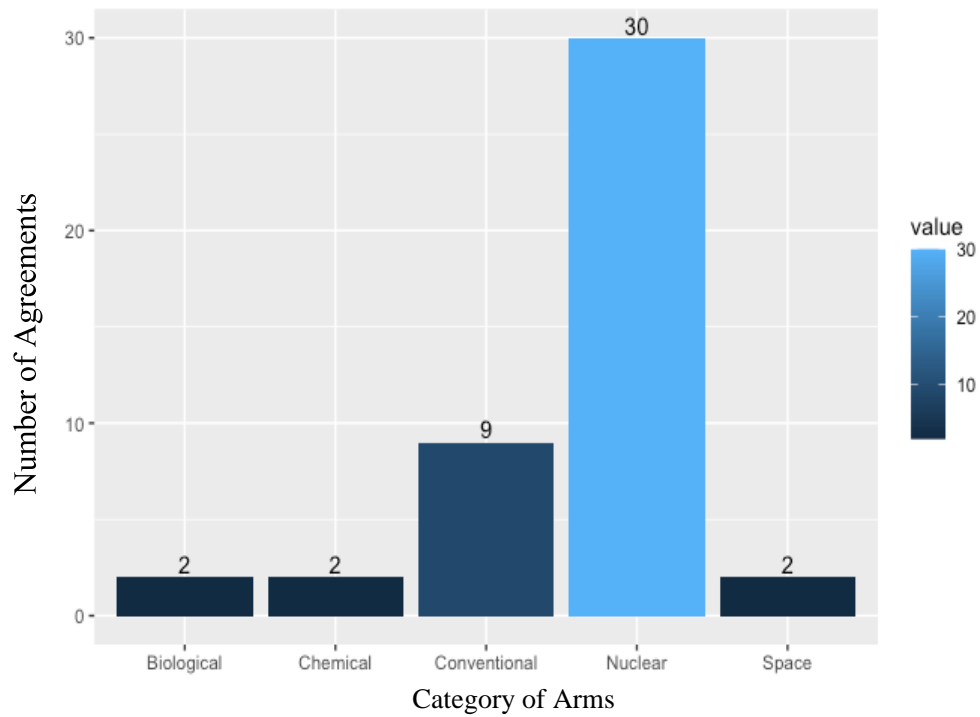
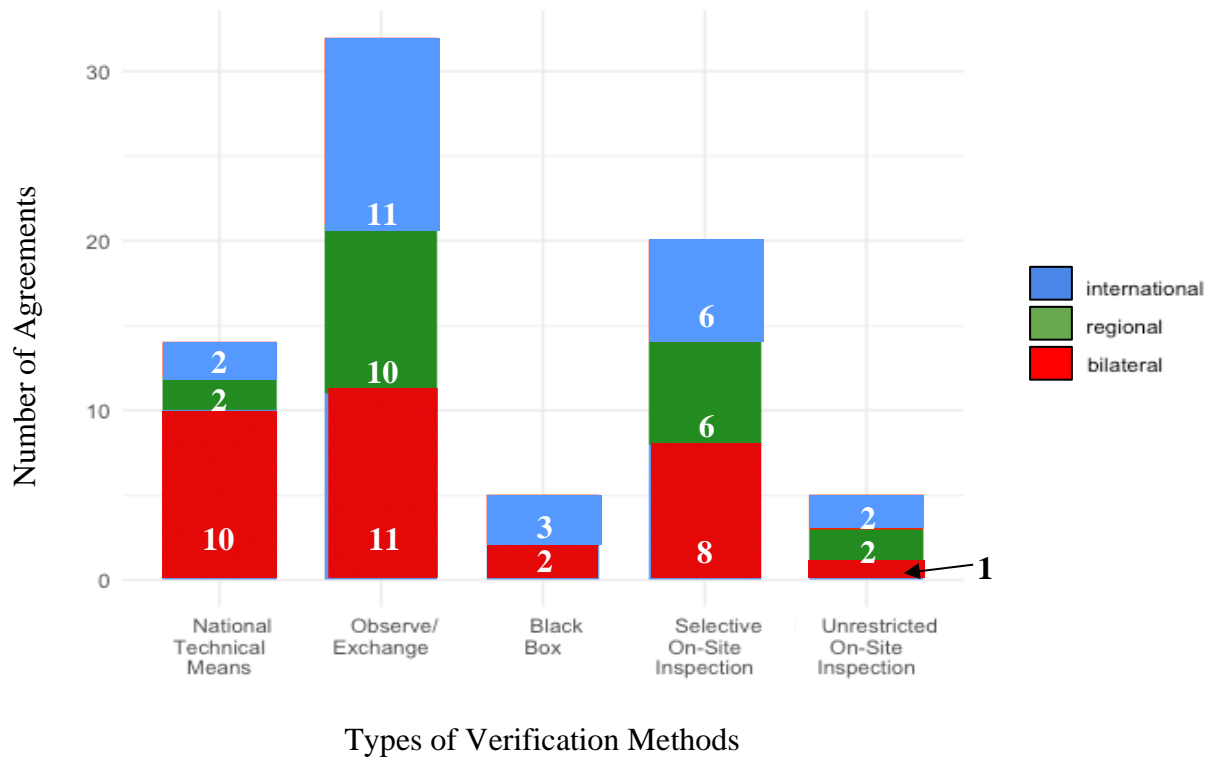


Figure 6.1.3 Arms Type by Number of Agreements



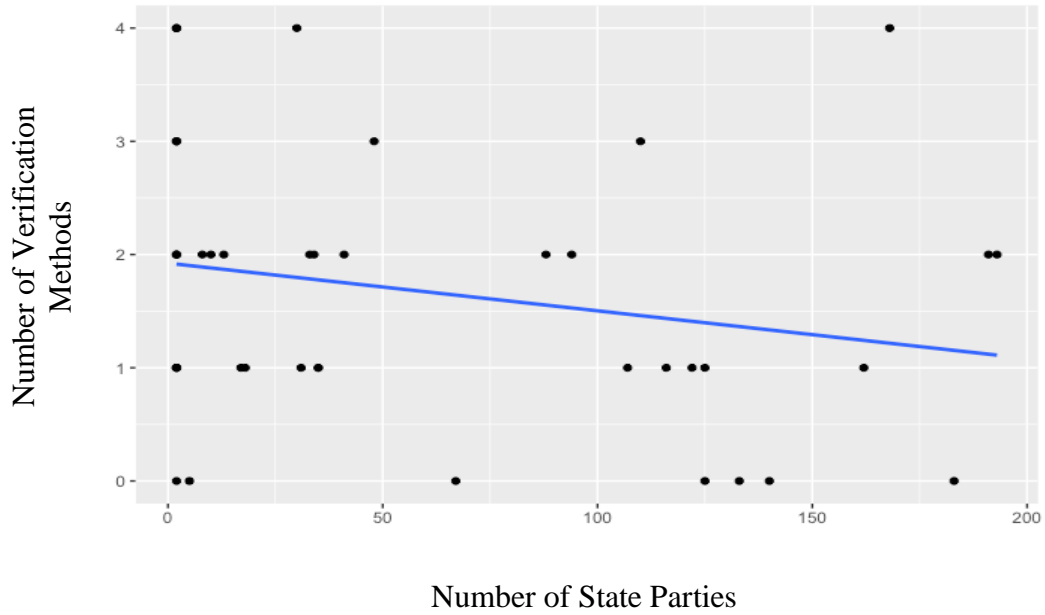
Types of Verification Methods by Level of Agreement

Figure 6.1.4



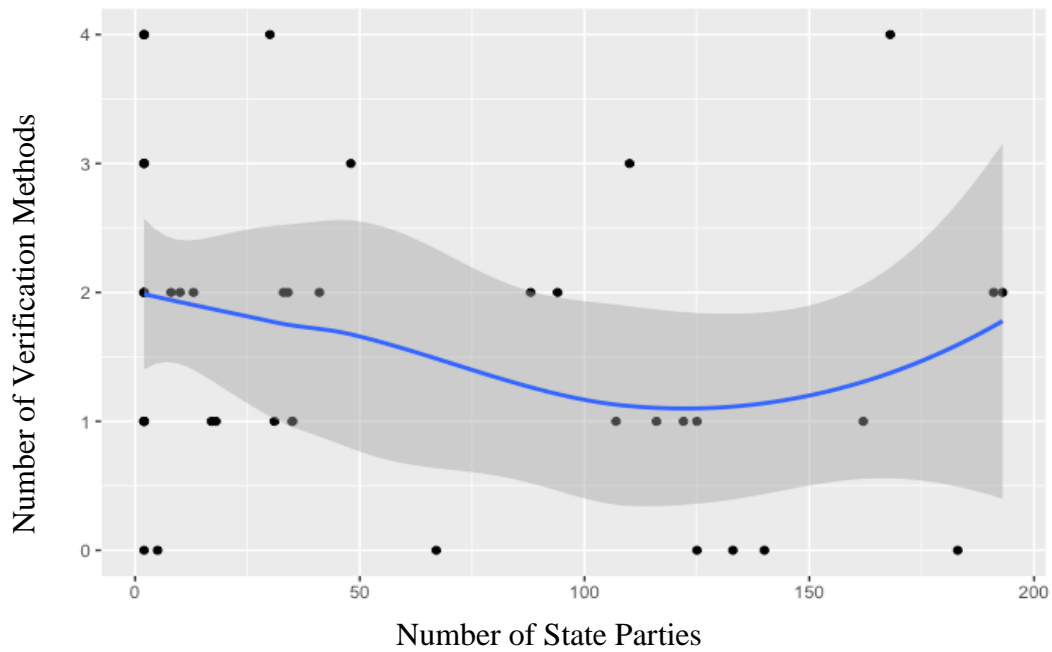
## Scatterplots: Figures 6.1.5 – 6.1.7

*Total Verification Methods and Number of State Parties*

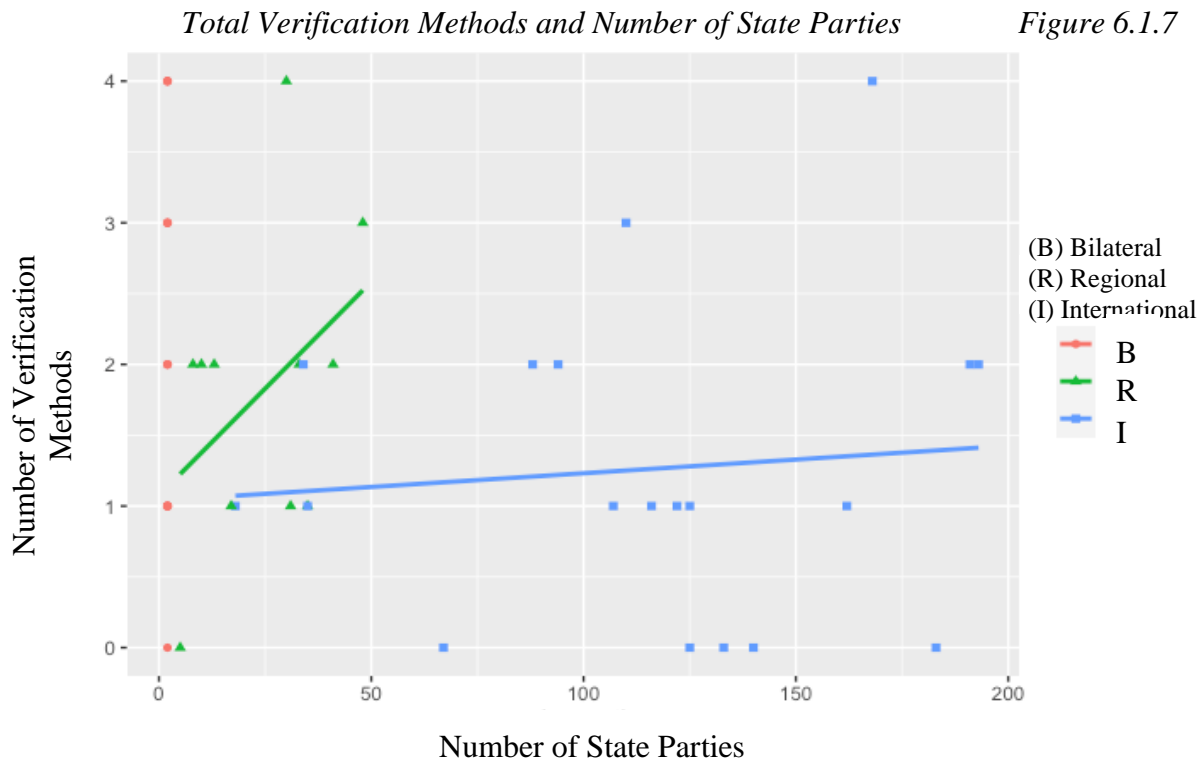


*Figure 6.1.5*

*Total Verification Methods and Number of State Parties  
(LOESS Method)*



*Figure 6.1.6*



## Regression Outputs: Figures 6.2.1 – 6.2.7

*Figure 6.2.1* Relationship between the number of parties and total verification methods:

```
Call:
lm(formula = total ~ numberparties, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-1.91539 -0.91539  0.08461  0.87999  2.78320

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.923809   0.237564   8.098 3.44e-10 ***
numberparties -0.004208   0.002841  -1.482   0.146
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.187 on 43 degrees of freedom
Multiple R-squared:  0.04857,    Adjusted R-squared:  0.02644
F-statistic: 2.195 on 1 and 43 DF,  p-value: 0.1458
```

Figure 6.2.2 Relationship between the number of parties and National Technical Means :

```
Call:
lm(formula = ntm ~ numberparties, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.4315 -0.3666 -0.1765  0.5685  0.9399

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.436013   0.090412   4.822 1.81e-05 ***
numberparties -0.002237   0.001081  -2.070  0.0445 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4516 on 43 degrees of freedom
Multiple R-squared:  0.0906,    Adjusted R-squared:  0.06945
F-statistic: 4.284 on 1 and 43 DF,  p-value: 0.04452
```

Figure 6.2.3 Relationship between third party and NTM:

```
Call:
lm(formula = ntm ~ thirdparty, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.45  -0.20  -0.20   0.55   0.80

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.4500     0.1020   4.411 6.78e-05 ***
thirdparty   -0.2500     0.1369  -1.827  0.0747 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4562 on 43 degrees of freedom
Multiple R-squared:  0.072,    Adjusted R-squared:  0.05042
F-statistic: 3.336 on 1 and 43 DF,  p-value: 0.07471
```

Figure 6.2.4 Relationship between third party and Selective On-Site Inspection:

```
Call:
lm(formula = selective ~ thirdparty, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.600 -0.250 -0.250  0.400  0.750

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.2500     0.1065   2.348  0.0235 *
thirdparty    0.3500     0.1429   2.450  0.0184 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4762 on 43 degrees of freedom
Multiple R-squared:  0.1225,    Adjusted R-squared:  0.1021
F-statistic: 6.003 on 1 and 43 DF,  p-value: 0.01843
```

Figure 6.2.5 Relationship between NTM and Number of Parties, CINC Score, Third Party:

```
Call:
lm(formula = ntm ~ numberparties + cincscore + thirdparty, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.6923 -0.3036 -0.1601  0.3579  1.0289

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.463848   0.123275   3.763 0.000527 ***
numberparties -0.004907   0.001692  -2.900 0.005973 **
cincscore     0.585902   0.295571   1.982 0.054178 .
thirdparty    -0.242958   0.127304  -1.908 0.063346 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4242 on 41 degrees of freedom
Multiple R-squared:  0.2348,    Adjusted R-squared:  0.1789
F-statistic: 4.195 on 3 and 41 DF,  p-value: 0.01117
```

*Figure 6.2.6* Relationship between Selective On-Site Inspection and Number of Parties, CINC Score, Third Party Verifier:

```
Call:
lm(formula = selective ~ numberparties + cincscore + thirdparty,
    data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.7201 -0.3551 -0.1218  0.3835  0.9090

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.3720952   0.1385047   2.687   0.0104 *
numberparties  0.0005006   0.0019010   0.263   0.7936
cincscore     -0.3405413   0.3320874  -1.025   0.3112
thirdparty     0.3479058   0.1430319   2.432   0.0195 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4767 on 41 degrees of freedom
Multiple R-squared:  0.1616,    Adjusted R-squared:  0.1003
F-statistic: 2.634 on 3 and 41 DF,  p-value: 0.06257
```

*Figure 6.2.7* Relationship between Unrestricted On-Site Inspection and Number of Parties, CINC Score, Third Party Verifier:

```
Call:
lm(formula = unrestricted ~ numberparties + cincscore + thirdparty,
    data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-0.35774 -0.17786 -0.04761  0.03232  0.80222

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.024465   0.085428  -0.286   0.77603
numberparties -0.002993   0.001173  -2.552   0.01452 *
cincscore     0.659477   0.204827   3.220   0.00251 **
thirdparty     0.025775   0.088220   0.292   0.77163
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.294 on 41 degrees of freedom
Multiple R-squared:  0.2026,    Adjusted R-squared:  0.1443
F-statistic: 3.473 on 3 and 41 DF,  p-value: 0.02449
```