The Disclosure Dilemma: Intelligence and International Organizations

Allison Carnegie and Austin M. Carson*

November 11, 2015

*Allison Carnegie is Assistant Professor, Department of Political Science, Columbia University, New York, NY (Email: allison.carnegie@columbia.edu). Austin M. Carson is Assistant Professor, Department of Political Science, University of Chicago, Chicago, IL (Email: austinmcarson@gmail.com). We thank Page Fortna, Ashley Jester, Jeff Kaplow, Chris Kojm, Maggie Peters, Paul Poast, David Steinberg, Jane Vaynman, Felicity Vabulas, Tristan Volpe and the participants of the 2015 APSA conference for helpful comments. We also thank Bryan Schonfeld and Laura Sipe for excellent research assistance. We gratefully acknowledge funding from the Columbia Provost's Grant. All remaining errors are our own.

The Disclosure Dilemma: Intelligence and International Organizations

Abstract

International organizations (IOs) have long been hailed as solutions to problems that emerge when states are unable to credibly commit to specific policies. We identify a largely over-looked issue that impedes cooperation within IOs: they often require sensitive intelligence that they are not able to obtain themselves and member states are unwilling to share. Employing game-theoretic tools, public opinion surveys, statistical analyses, and detailed case study examinations, we build and test a theory that identifies barriers to intelligence sharing, uncovers the circumstances under which IOs facilitate intelligence cooperation, and derives the implications of these dynamics for interstate relations more generally. We argue that states face a disclosure dilemma, or an inability to share intelligence to address areas of common concern due to concerns about revealing sources and methods associated with obtaining that information. This dilemma often impedes intelligence sharing altogether or leads states to share intelligence that sacrifices their long-term security interests in favor of short-term political gains. However, we argue that IOs can ameliorate this disclosure dilemma if they can guarantee that they will not divulge their members' sources and methods. Consistent with our theory, we find evidence in a variety of empirical settings that IOs can harness intelligence to better achieve cooperative goals if they are designed to protect sensitive information and remain politically impartial.

NOTE: What follows are excerpts from a book draft currently in progress. It includes introductory material situating the project, a formal model, and the key findings from an empirical chapter. The overall plan for the book is described at end. We appreciate any comments!

The breakup of Yugoslavia led to a spasm of ethnic violence from 1991 to 1999. The International Criminal Tribunal for the Former Yugoslavia (ICTY) was established in 1993 to investigate alleged war crimes. Early in its mission, the ICTY faced an important information-gathering challenge. The institution's staff found itself reliant on data shared by third party states to document illegal behavior of leaders like Serbia's Slobodon Milosevic. Yet the most conclusive evidence was often derived from national intelligence. Imagery from satellites and drones, signals intercepts, and similarly sensitive intelligence sources gave leading states in Europe and elsewhere information about war-time behavior that was extremely precise. Potential contributors often balked at sharing such sensitive information. If intelligence could be shared, the benefits were substantial: thousands of pages of highly classified American evidence, for example, featured photos from satellites, reconnaissance planes and videotapes showing the locations of secondary graves.[1] Yet the potential costs to a state from providing intelligence were clear as well. During one episode, German imagery intelligence disclosed to the ICTY "highlighted to the Serbs that we knew where the Muslim bodies were buried, thus tipping them off to return to the killing fields and destroy the mass graves in order to remove and scatter the evidence. Keeping some of the information out of the public eye proved essential at that point" (Scheffer, 2012, 274).

The Yugoslavia war crimes tribunal is an example of a general dilemma that states with intelligence often confront. National intelligence can provide states with information that, if shared broadly, could influence policy in ways that further both national self-interest and cooperative goals. Yet sharing intelligence tends to endanger future collection. In general, using intelligence to influence political events requires intelligence claims to be sufficiently precise and detailed so that they are credible. This, in turn, endangers technical, human, and other sources. Exposing sources and methods allows current and future intelligence targets to adapt. As a result, states with intelligence relevant to the mission of international organizations (IOs) tend to have a very high threshold for sharing their sensitive private information. Yet the successes in using intelligence in

[1]See Branigin, William. "U.S. Evidence Enhances Case Against Milosevic." *Washington Post.* May 28, 1999; Manning (2000, 1, 12, 16).

the ICTY suggest that this trade-off in sharing intelligence can be overcome if global governance institutions are equipped to handle and protect sensitive intelligence disclosures. Indeed, the U.S. and other countries insisted during the design phase of the institution that the tribunal "protect confidential information obtained by the Prosecutor." The chief prosecutor noted that "trust and confidence had to be built between the institution and the government concerned" for governments to hand over sensitive information (Moranchek, 2006, 484).

International organizations have long been theorized as solutions to problems that emerge when states are unable to share information and credibly commit to specific policies. We identify an over-looked but common informational problem in world politics: the sensitivity of intelligence sharing and what we call states' "disclosure dilemmas." The response by most states most of the time – withholding their intelligence – leads to missed cooperative opportunities and hobbles the effectiveness of global governance. However, IOs with certain design features can alleviate this disclosure dilemma and facilitate welfare-enhancing intelligence-sharing. Indeed, we demonstrate that in several issue areas member-states and IOs themselves in the last three decades have begun experimenting with building IO capacity to protect sensitive information. When done effectively, IOs can enable states to overcome the disclosure dilemma in three ways: by enabling comparisons among different states' intelligence findings; by investigating intelligence conclusions to authenticate or reject their validity; and, by viewing but safeguarding the sources and methods used to generate intelligence conclusions. Employing game-theoretic tools, public opinion survey experiments, statistical analysis, and detailed case study examinations, we build and test a theory that describes the disclosure dilemma, explains how IOs can address it, and identifies the circumstances under which IOs can successfully perform this function as well as the positive effects that result. While we posit IOs as a possible solution we are clear that success to date has been inconsistent and that information protection in IOs produces its own political and practical challenges.

The Argument in Brief

We analyze the role of intelligence in global governance. International relations theories of international institutions have long argued that information provision is an important reason that states invest in creating IOs. We are interested in a particular kind of private information: explicitly concealed information gathered through clandestine means about another state's behavior and intentions, which we refer to simply as "intelligence." Almost all states gather at least some intelligence on issues of national interest through highly systematized bureaucracies (Andrew, 1979). While adversaries' military activity is the traditional emphasis of intelligence collection, many states also gather intelligence on their allies' activities, economic performance, civil strife, environmental trends, and infectious diseases. This information is derived from many sources such as human and technical assets, where the latter includes imagery intelligence (i.e. aerial or satellite photography), signals intelligence (i.e. interception and decryption of radio, cellular, and other communications), and more exotic tactics like nuclear radiation signature analysis and cyber-espionage (Johnson, 2007; Richelson, 2007).

We argue that states operate in the shadow of a largely unrecognized informational dilemma. Modern states devote precious resources to building national intelligence capabilities. While states use intelligence for national self-interest, the process of gathering information often uncovers evidence of prohibited or dangerous activity by other governments and private actors that exceeds what IOs and non-governmental actors know. If shared, this private information can further international public goods and help IOs perform their missions. However, states face a difficult dilemma. Sharing such intelligence helps to achieve political goals but risks compromising the sources and methods used to obtain it. Sometimes disclosures do occur but the threshold for states to do so is very high. A well-known example demonstrates this point: during the Cuban Missile Crisis, American leaders divulged aerial photography from sensitive U-2 overflights of Cuba in order to ensure that their claims were credible. While sources and methods revelation did not prevent disclosure in this case, U.S. leaders went public only in the face of an unusually acute security threat and the American intelligence community strongly objected (Andrew and Andrew, 1995,

292-298). This basic tradeoff between the political utility of widely shared intelligence and the sources/methods damage from such revelations is the essence of the disclosure dilemma.

We further contend that withheld intelligence contributes to missed opportunities for international cooperation. In domains ranging from nuclear proliferation to war crimes to environmental change, details only available through intelligence collection could, if shared, help to identify problems in need of multilateral attention and energize the enforcement of international norms and laws. These missed opportunities are often difficult to observe when the intelligence remains private. A contribution of this book is that it identifies a host of recent successes in which states have shared intelligence to energize multilateral action. These cases underscore the large potential impact that more robust intelligence-sharing could make in a range of empirical domains.

We then argue that international organizations can play a role in ameliorating this dilemma. Solving the disclosure dilemma requires giving states a way to share the content of national intelligence without jeopardizing the sources/methods used to obtain it. International organizations can do so by receiving and comparing disclosures from multiple states' intelligence services; by investigating intelligence claims that lack sources and methods to authenticate or reject them; and, by receiving and protecting intelligence claims with sources and methods. Each of these, in turn, requires two important institutional features. First, IOs must be designed to receive sensitive intelligence and protect it. This requires confidentiality rules and an organizational capacity to secure information and avoid leaks. Second, IOs must be able to assess the veracity of disclosed intelligence. IOs must be capable of credible and relatively unbiased evaluation of the validity of intelligence-based claims given that the original raw data is not more widely distributed. This requires a secretariat with technical expertise and without a politicized reputation. When IOs perform these functions, governments are free to regularly share relevant intelligence; this can aid in multilateral rule enforcement and in identifying new problems. To be clear, we argue that IOs *can* serve such a function but readily acknowledge they have not regularly done so in the past. To shed light on their potential as solutions, we devote special attention to examples in which states have experimented with using IOs for this purpose. In the area of weapons of mass destruction, for

example, UNSCOM's use and protection of shared intelligence regarding Iraqi compliance played a key role in improving non-proliferation performance compared to efforts in the past.

The Disclosure Dilemma

Our argument rests on the observation that many states gather intelligence about other states' activities, which often provides far greater information about these states' actions than is available publicly or through international organizations. When shared, such unilateral intelligence has the capacity to strengthen international cooperation, the provision of international public goods, and international order. Information that is disseminated among states can foster multilateral punishments for wrong-doing, allow states to share the costs of keeping international order, and allow states to gain access to resources that other states possess. Intelligence sharing facilitates specialization, such that states can collect intelligence in the manner they are especially skilled in, and all states can then have access to better intelligence than they could collect on their own. Since intelligence often concerns other states' responsibilities for crimes and non-compliance with international rules, this intelligence sharing can counter these states' campaigns of deception to foster transparency and accountability throughout the international system.

As neoliberal institutionalism has long argued, lack of information about compliance is a key barrier to an effectively rule-governed international system. Thus, if states disclosed this kind of intelligence, the information environment would be richer facilitating international cooperation. However, states often balk at sharing such intelligence because doing so can jeopardize sources and methods. Intelligence collection often requires states to develop proprietary means of eliciting sensitive information, which then may be exposed along with their intelligence. For example, providing satellite photographs can reveal the locations and capabilities of specific satellites (imagery intelligence), or certain information may expose the presence of informants (human intelligence) or the ability to intercept communications (communications intelligence). Thus, while informed states may wish to share their classified information, they may be unable to do so without incur-

ring unacceptably high costs to future intelligence collection and political costs from exposing surveillance.

One alternative is for states with such intelligence to only release conclusions or specific claims after scrubbing sources and methods. However, this creates a problem of credibility, as such a state could have simply manufactured the content of the report to achieve its own political ends. Scrubbing intelligence thus generally fails: making claims without the raw intelligence materials lead many observers to discount these claims. Scholars have recognized the issues caused by failing to reveal sources and methods. For example, Chesterman (2006, 21) states, "the lack of information about the source makes it difficult to assess reliability...lack of context can undermine preparation of an appropriate response," and "when isolated pieces of information are shared in this way a reasonable response of the recipient is to suspect that they are being manipulated."

As a result, non-disclosure is the most common reaction even when sharing national intelligence could make multilateral action, including institutional performance, more effective. Sometimes disclosure does occur but, as demonstrated by the Cuban Missile Crisis example, the threshold for states to reveal intelligence that sacrifices sources and methods is very high. States tend to only do so in situations where the loss of sources/methods is tolerable because the benefits of providing the information are so great, either due to a high security threat that the intelligence can mitigate or to domestic political rewards associated with revealing the information.

An informed state thus faces a trade-off. It can: 1) Reveal its information and compromise its intelligence methods; 2) Reveal its information but not compromise its methods, in which case the recipient of the information may not trust the information or be able to act on it; or 3) Not reveal its information. None of these represent ideal options; thus the informed state releases its information when the immediate benefits of doing so are high and concerns over sources and methods remain are relatively muted, such as when the informed state expects little future interaction with the target. If these conditions do not hold, intelligence is not shared, often frustrating international goals.

IOs can Solve the Dilemma

One potential solution to the disclosure dilemma is to use IOs to authenticate intelligence and combine it with other information while avoiding broader disclosure of sources and methods. Having a reputable, credible outside entity that can receive and review national intelligence can guard against frivolous and politically motivated claims; observers are then more likely to trust a claim without access to the discloser's sources and methods. Such information can therefore be shared and believed without the discloser worrying that its sources and methods will be jeopardized.[2]

Consider, for example, an analogous situation: journalists' reliance on anonymous sources. Journalists often encounter sources who will only agree to speak under conditions of anonymity. Such sources are often high-placed government or industry insiders whose views are critical for completing a story. However, when using anonymously sourced information, journalists cannot merely make assertions or their claims may not be trusted by their audience. Without identifying how they reached a conclusion, "unnamed sources can really only be heard and not judged...The public may recognize this [at]...the moment when facts become challenged, motives appear suspect, and arguments for trust fall under fire" (Pribanic-Smith, 2012, 138). However, because journalists must protect these sources, they "have only limited recourse in countering the claims of others without revealing too much information about their unnamed sources" (Pribanic-Smith, 2012, 147). Indeed, a host of press scandals which revealed that journalists made up information have made this problem increasingly salient (Pribanic-Smith, 2012), though the need for information vetting has long been recognized (Wulfemeyer, 1982).

One solution to this problem is for editors and fact-checkers to vet a journalist's sources but also protect them. The public then has the knowledge that a reputable individual or organization has ensured the veracity of the information, and can trust the article more fully. As a result, journalists are often required to provide editors with source names (Wulfemeyer, 1982). Indeed, "no guideline

[2]Other solutions exist, such as ad hoc intelligence sharing among states in a hierarchical relationship. However, we note that it is often difficult to get states to agree to give up sovereignty to enter such agreements and does not facilitate collaboration in the same manner, thus not solving the problem completely (Walsh, 2010).

for the use of anonymous sources has more support than the involvement of the editor, both in approving the grant of anonymity and in sharing knowledge of the source's identity...Having an editor know the identity of the source minimizes the risk of both inaccuracy and harm...Involving a second person is a critical check within the system to ensure that the story is accurate and fair in the face of the potential for error or abuse." (Boeyink, 1990, 237).[3]

We argue that the same basic dilemma exists in world politics. The potential for international organizations to play a role in alleviating this dilemma, similar to the role played by the editor in the journalism analogy, has not been explored. The book shows that IOs can serve in just this role. To do so, a properly empowered IO receives intelligence submissions from member-states regarding potential non-compliance or similar issues. The IO secretariat – not the wider body of member-states – gets access to, and authenticates, the submitted intelligence. The IO can authenticate specific intelligence claims in three non-exclusive ways: by comparing multiple intelligence submissions from different states; by consulting relevant data collected in the past by the IO and supplementing with additional investigation; and, when sources and methods themselves are included in the intelligence submission, by scrutinizing the basis of the intelligence claim. Next, the IO shares its conclusions and any supplemental information with the wider regime membership, endorsing the intelligence-based claim while protecting the sensitive information therein (including any sources and methods). The IO therefore ensures that the state's claims further the purpose of the regime and do not simply satisfy the narrow political goals of the disclosing state. Yet often a given disclosure serves both functions. For example, when the U.S. shared information with the International Atomic Energy Agency (IAEA) regarding North Korea's nuclear program, it both satisfied the U.S.'s narrow political interests and bolstered the larger nonproliferation regime.

To perform this function, member-states and the IO itself must create and maintain an organizational capacity to protect sensitive information and prevent leaks. When IOs are able to play this role, member states know that they must share information with IOs or their information simply

[3]Similar dynamics also play out in other areas such as interactions between law enforcement and confidential informants.

will not be believed. If all states with credible information share it with IOs, then only a state with poor or manufactured information would refuse to do so. Importantly, such an expectation can benefit both powerful and weak states. Weak states can benefit because, even while details like sources and methods are not visible to them, they obtain more high quality information from the authentication process. These states thus can act on information that they would not have had access to otherwise. Additionally, powerful states can benefit because an avenue for credibly sharing intelligence to influence policy is opened up while minimizing the damage to intelligence collection. Achieving goals such as building an international coalition to bolster multilateral support for a particular course of action – or simply swaying international opinion – is therefore easier to achieve.[4]

Contributions

The book makes three primary contributions. Most broadly, it provides careful analysis of national intelligence's impact on international organizations in a variety of settings. For decades, scholarly analysis of cooperation-enhancing IOs and states' national intelligence has advanced almost entirely independent of one another. We bring into conversation diverse literatures and research traditions including intelligence studies, the rational design of international institutions, game theoretic models of strategic communication, and international law. The result is a set of findings on the disclosure dilemma and national intelligence that is of broad interest for scholars within and outside of international relations. Moreover, our empirical chapters on different issue areas feature findings about intelligence that will be of interest to scholars working in these specific domains. Our findings on the role of national intelligence in the nuclear non-proliferation regime, for example, sheds new light on an area of enduring interest to international security scholars. Overall, we demonstrate the existence of a recurring theme in everything from peacekeeping to environmental governance to weapons of mass destruction: national intelligence can play an important role in fa-

[4]Powerful states do not always benefit. If they seek to fabricate information, they may find themselves unable to convince other states when they otherwise would have.

cilitating global cooperative goals but only if international organizations successfully address the disclosure dilemma.

We also contribute to long-standing debates about how international organizations contribute to the challenge of interstate cooperation. Beginning with Keohane (1984), scholarship on international institutions has claimed that international institutions facilitate state cooperation by reducing barriers to information sharing. The primary role of IOs is to receive and/or gather information and then disseminate it; this facilitates reputational and other mechanisms that support cooperation over time. In one sense the our findings echo this view: if designed appropriately, we show that IOs can better advance cooperative goals by receiving and acting upon national intelligence that would otherwise remain private information. The aggregate effect is a familiar story of more information shared and better cooperative outcomes. Yet underlying this outcome is the opposite informational role: the erection rather than removal of information barriers. Our central finding is that these institutions can only address inefficient information-sharing if they can *securely protect intelligence received from other states*. In diverse issue domains, we find that states balk at sharing information unless they are confident that an IO will protect its sensitive sources and methods. As a result, we describe a variety of substantive domains in which member-states and IOs themselves work to build organizational capacities for secrecy. Our description of disclosure dilemmas and the international response provides a new and theoretically promising contrast to the prevailing view of how IOs facilitate cooperation.

The third main contribution is to illuminate an important normative tension in global governance. Better equipping international organizations to handle intelligence runs directly against the dominant trend towards accountability and transparency in global governance. Early criticism of a "democratic deficit" in international institutions has swelled into a chorus of calls for greater transparency and openness regarding IO activities in international finance, global trade, European integration, and environmental regulation (Blanton, 2007; Ehring, 2008; Gupta, 2008; Koenig-Archibugi, 2004). Doing so can help to enhance the accountability and therefore legitimacy of international institutions (Grant and Keohane, 2005). However, we identify a surprising counter-

vailing trend. In a number of substantive domains, the attempt to seize opportunities for better IO performance through integrating national intelligence is producing more, rather than less, barriers to openness. In the Yugoslavian war crimes tribunal, for example, efforts to assure intelligence-sharing states that their sensitive information would be protected through new international legal secrecy measures prompted critiques that such measures create "problems for preserving the openness of international criminal trials" (Moranchek, 2006, 479). By shining a light on the overlooked issue of intelligence role in IOs, we describe and theorize an important and increasingly common global governance phenomenon that creates new problems for making international institutions accountable and transparent.

A Model of Intelligence Disclosures

We use a cheap talk model adapted from Crawford and Sobel (1982) to show that when a single informed state gives advice to the international community, only noisy information can be credibly transmitted. The more biased the state, the noisier the information. While there are informative equilibria, these entail a significant loss of information. Revealing sources and methods would allow the message to be verified, but comes at a high cost and thus only occurs when states face urgent foreign policy concerns. However, an IO can improve information transmission by gathering its own information and transmitting it to the international community as an unbiased party, soliciting information from additional states, or obtaining but protecting the sources and methods of the single state. Part of the discussion below closely follows Krishna and Morgan (2008).

The model consists of interactions between two states, A and B, and an observer O, which are denoted by superscripts. The observer could be the international community, specific states, or domestic actors. Prior to the start of the game, B has violated international law with probability θ, which is distributed uniformly on the unity interval. State A possesses sophisticated intelligence collecting capabilities, so that if a violation has occurred, A detects it. A thus knows the value of θ. However, O lacks the ability to detect the violation and is thus unaware of the value of θ. A may

send a costless message to O regarding the value of θ, after which O must make decision y, which is how severely to punish B. Alternatively, A may reveal its sources and methods, in which case O can then observe the precise value of θ. Revealing sources and methods costs c.

O supports the fair application of international law, while A seeks a punishment that reflects the true scope of B's violation plus A's bias toward B.[5] O does not have access to the government's intelligence capabilities and thus has a difficult time ascertaining whether the government seeks a specific policy to further its own partisan policy preferences, or whether it seeks to uphold international legal commitments of other countries. Examples are numerous, as we detail more extensively subsequently. For instance, in the domestic context, U.S. president Reagan sought to support the Contras against the government of Nicaragua, but had difficulty obtaining public support without providing evidence that the Nicaraguan government was in violation of international law. Similarly, the U.S. government wanted to punish the Soviets for chemical weapons use, but needed to convince the public that chemical weapons had actually been used; otherwise, punishment was seen as being politically-motivated and was not supported by the domestic audience. In the international context, the U.S. desired support for the 2003 Iraq War, but other states worried that the U.S. was not providing accurate information regarding Iraq's alleged violations. Analogous examples abound.

O's payoff is thus $U(y,\theta) = -(y-\theta)^2$, and A's payoff is $V(y,\theta,b) = -(y-(\theta+b))^2$, where $b \geq 0$ is a bias parameter that indicates how closely O and A's preferences align.

The timing of the game is as follows:

1. A learns the value of θ.

2. A then chooses whether to reveal its sources and methods.

3. A sends message m to O, leading O to form beliefs determined by the distribution function G.

4. Finally, O chooses whether to punish B.

[5]Other considerations are possible for O; the key aspect of O's utility is that learning of a violation makes O more willing to punish B, and learning that a violation did not occur makes it less willing to punish it.

With all of the game's pieces in place, we now solve the model.

Solving the Model

We solve for the Bayesian-perfect equilibria of the model by backward induction. Suppose for the moment that c is so high that sources and methods are never revealed. We find that an equilibrium exists in which A provides no information to O, even when A's and O's preferences match perfectly, such that $b = 0$. In this equilibrium, O rightly believes that A's message is uninformative and so decides whether to punish B solely based on its previous information. A then has no reason to provide information and thus "babbles" instead, giving uninformative messages. Further, whenever A and O's preferences are not aligned, no information is conveyed by A.

Proposition 1. *When A is even partially biased, information loss is present in all equilibria.*

We obtain this result because if A's message always conveyed the actual value of θ, and O accepted this message as fact, A would then be motivated to distort the message. Specifically, rather than reporting θ, A would report $\theta + b$. O therefore anticipates this and knows that it is not receiving full, unbiased information.

Can O ever obtain credible information from A, even if it is only partly informative? After A provides message m, O chooses y to maximize its utility given G; thus, O selects y such that

$$y(m) = E[\theta|m]. \tag{1}$$

Suppose that A can decide whether to send message m that leads O to administer a minimal punishment to B (y), or can send message m' that leads O to administer a stronger punishment ($y' > y$). Suppose also that O wants to administer justice to B, such that the more B violated the law, the stronger the punishment that O desires. Thus in a state where B undertook a grave violation of the law, $\theta' > \theta$, O would rather implement y' than y, and prefers y to y' in state θ. These preferences satisfy the single-crossing condition; thus, O prefers y' to y for all violations greater than θ'. A

unique state a therefore exists which satisfies $\theta < a < \theta'$ and makes O indifferent between y and y'. This means that the distance between y and O's preferred punishment in state a equals the distance between y' and O's preferred punishment in state a, so that

$$a + b - y = y' - (a + b) \tag{2}$$

. We thus find that A sends message m whenever $\theta < a$ and sends m' whenever $\theta > a$. Returning to equation (1), we see that $y = \frac{a}{2}$ and $y' = \frac{1+a}{2}$. Plugging these values into equation (2), we find

$$a = \frac{1}{2} - 2b. \tag{3}$$

Equation (3) shows that when A is more biased, such that $b \geq \frac{1}{4}$, no value of a will satisfy the equation, so A does not provide any information. Further, when $b < \frac{1}{4}$, O receives less information the greater the value of b. This provides the intuition for the following proposition, which Crawford and Sobel (1982) proves formally:

Proposition 2. *The state space is partitioned into a finite number of intervals in all equilibria. In the most informative equilibrium, the information that O obtains is decreasing in b.*

If A could credibly convey the true value of θ to O, both A and O would benefit because O would select $y = \theta$ and receive a utility of 0, and A would receive $-b^2$, while the pay-offs are below these values in any equilibrium above. Thus, as in standard cheap talk models, information is lost due to A's inability to credibly reveal θ, and more information is lost as the bias grows.

Consider now the conditions under which A chooses to reveal its sources and methods, since doing so allows it to convey the true value of θ, but only at cost c. When A can disclose sources and methods, it receives $-b^2 - c$. If it does not do so, it receives $-(y - (\theta + b))^2$. Thus, A reveals its sources and methods when $-b^2 - c \geq -(y - (\theta + b))^2$, which occurs when the cost of disclosing sources and methods is lower and the distance between A's preferred punishment for B and the punishment that O will actually administer is greater. Note that the effect of the bias is ambiguous.

Proposition 3. *A is more likely to reveal sources and methods the lower the cost associated with doing so, and the further O's punishment decision will be from A's preferred decision otherwise.*

Adding an International Organization

Now suppose states play the same game, but also have the option of providing their intelligence to an IO. If the IO can protect sensitive information contained in intelligence disclosures, then it may serve three functions. Each can serve individually or in combination as an alternative to making credible intelligence-based claims by *publicly* sharing intelligence sources and methods.

First, the IO can collect information from multiple informed states, which can help to corroborate information even if sources and methods are not revealed. While intelligence conclusions from a particular state often are not credible unless they are backed up with sources and methods, receiving the same conclusion from many states makes the intelligence more likely to be believed.

Second, an IO can analyze intelligence-based conclusion that a state reaches without learning of the state's sources and methods, investigating the claim independently and rendering a judgment of validity. An IO with verification or fact-finding powers can do this in a variety of ways including by collecting environmental samples, interviewing relevant government and industry personnel, obtaining privately available information as with commercial satellite imagery, or seeking out relevant expertise. The IO can then disseminate its conclusions to the member states.

Third, the IO can receive conclusions that states reach based on their intelligence collection along with the sources and methods used to obtain that intelligence. While states may not be willing to share their sources and methods with a diverse set of member-states, they may share them with the IO if it is unlikely that the information will be leaked. The IO can then attest to the validity of the information. The wider group of member-states are then much more apt to believe the conclusions that an IO reaches than one that a biased state comes to, and are therefore often more likely to believe the claims when the IO stands behind them.

Representing these ideas formally, suppose there is an additional informed state, D, so that the timing of the game is the same as above, but both A and D learn the value of θ. Then, in the second

step, A and D chooses whether to reveal their sources and methods or to provide their conclusions simultaneously to an IO. In this case, the IO can make both A and D better off. Consider first the outcome when A and D have identical biases. Suppose O selects the most conservative punishment of the two suggestions from A and D, so that if $m_1 < m_2$, O opts for m_1, while if $m_2 < m_1$, O opts for m_2. If D sends the message $m_2 = \theta$, then a message from A of $m_1 > \theta$ does not effect O's decision. Yet if A sends the message $m_1 < \theta$, O's action becomes $y = m_1$; however, now A is worse off. Thus, A reports the true value of θ. The outcome is the same when A and D have opposing biases, but because it is more complex and has been derived in other contexts elsewhere, we do not reproduce the proof here.[6] Because the IO can solve the information problem, states always reveal their conclusions to the IO and all are better off.

Proposition 4. *When IOs can collect intelligence conclusions from multiple informed states, states reveal their conclusions to the IO and all are better off, even if sources and methods are not disclosed.*

In many cases, multiple states do not possess intelligence about violations of international law. Frequently, only one state such as the United States has the highly sophisticated intelligence capabilities required to discern a violation. In such a case, it is still possible for an IO to solve the information problem. In particular, suppose now that A can tell the IO whether it believes a violation occurred and the IO can then independently authenticate it by comparing the submitted intelligence to existing data the IO possesses and conducting additional investigations. Of course, the IO itself could maintain a particular bias, depending the the composition of its staff or membership, or the specific states it relies on for funding. Tasking the IO with acquiring its own information thus represents a trade-off between the international community relying on its own uninformed decision about whether to punish B, or following the biased recommendation of the IO. Suppose the IO possesses bias b. As shown by Dessein (2002),

Proposition 5. *When the IO's bias is not too big and the IO can collect its own information, the*

[6] See Krishna and Morgan (2001) for the proof. Note that the equilibrium is fragile. If an equilibrium refinement introducing the idea that A and D can make mistakes is added, full revelation is no longer possible (Battaglini, 2002).

international community is better off delegating the punishment decision to the IO than in any other equilibria above, even if sources and methods are not disclosed.

However, authenticating intelligence is costly for the IO. An alternative role for the IO is to collect the intelligence from the informed state, which could reveal its sources and methods to the IO. The IO could then easily authenticate that the intelligence was correct. But this too can be costly, this time due to the chance of leaks. The IO secretariat, like any organization, varies in its capacity to protect the sensitive information disclosed to it. Scholars of organizations emphasize that organizational secrecy is effortful (Geser, 1992, 438). Within an organization, information and activity are by default transparent to other parts of the organization, so that protecting sensitive information requires an IO to develop organizational processes to prevent its wider distribution. This can require restricting physical access to documents for certain offices or staff-members, developing punishments for unauthorized disclosure by outside entities like journalists or governments, and cultivation of norms of information protection among staff.

IOs thus vary in their ability to prevent leaks and states react strategically to this risk. If the IO leaks the information to O, A's sources and methods are compromised and A is worse off. Thus, A now only reveals if the probability of a leak, p, is not too high. Further, we may relax the assumption that states may either reveal sources and methods or not, and instead stipulate that states can choose the degree to which they will release sources and methods, or $s \in \{0,1\}$. For example, states may share their raw intelligence, fully releasing sources and methods, or they may distort the intelligence to protect certain aspects of their sources and methods. Intelligence may be scrubbed to varying degrees, or sources and methods may not be revealed at all. The costliness of a potential leak, $c(s)$, is increasing in the degree to which sources and methods were revealed, as the more sources and methods are revealed, the more a leak will compromise future intelligence collection abilities.

A now chooses s to maximize its utility: $-(y(s)-(\theta+b))^2 - pc(s)$, or it chooses s to satisfy $-2(y(s)-(\theta+b))y'(s) - pc'(s) = 0$. Notice that if no sources and methods are revealed, A receives $-(y-(\theta+b))^2$, while if sources and methods are fully revealed, A receives $-b^2 - pc$, since $y = \theta$.

We thus find that the more leaky the IO, the more *A* scrubs its sources and methods.

Proposition 6. *The lower the chance of a leak, the more A reveals its information along with its sources and methods, and the IO then authenticates and shares its conclusions with O.*

In sum, the model demonstrates that absent an IO, states cannot credibly transmit their information collected with intelligence assets without revealing sources and methods. However, since doing so is costly, states often choose not to reveal them, resulting in inefficient information transmission. An IO can solve this dilemma in three ways: collecting intelligence conclusions from multiple states, independently authenticating the intelligence of a single state by gathering additional information, or authenticating intelligence conclusions by viewing a state's sources and methods. While we have unpacked these mechanisms in isolation for clarity, an IO may in practice combine these activities in practice if it possesses the ability to handle sensitive intelligence.

Nuclear Intelligence and Preventing Proliferation

Our theory is highly applicable to the nuclear realm since nuclear weapons represent one of the most common areas in which sources and methods concerns arise. The extreme secrecy in which nuclear weapons are pursed and the critical role of intelligence in detecting these activities makes this a prime setting for disclosure dilemmas to occur. Further, since the introduction of nuclear weapons during World War II, the pull of these weapons has proven irresistible for many states, making these concerns especially common. Nuclear weapons' devastating capabilities and destabilizing effects make the resolution of disclosure dilemmas particularly important for the safety and security of the international community.

The model is well-suited to explain a common problem in the enforcement of the nuclear regime and the important role of international institutions in solving that issue. While disclosure dilemmas occur frequently in this arena, IOs such as the IAEA, UNSCOM, and the CTBTO can often mitigate them if they can protect states' sources and methods. As noted by Fuhrmann (2012, 220), the IAEA's successful functioning "often depends, however, on member countries sharing

information with the IAEA that is obtained by intelligence agencies. States may understandably be reluctant to share sensitive intelligence with an international organization because doing so could compromise their sources and methods."

As these IOs have been evolving overtime to cope with new nuclear concerns, they have been increasingly able to do. For instance, in 2012 the *FP* reported that "beginning two decades ago, the IAEA started relying less on information it gathers during its own field inspections alone and more on information that others provide."[7] Yet new nuclear challenges such as terrorism and new technologies continue to arise as well, making it particularly critical that these institutions can protect sensitive information contained in intelligence disclosures to function more effectively and further non-proliferation goals.

We now explain why disclosure dilemmas are so prevalent in the nuclear domain, and show how IOs can mitigate them. We explain how our model can be adapted to the nuclear context, and demonstrate evidence in favor of the model's propositions. We highlight three IOs: the IAEA, UNSCOM, and the CTBT, comparing and contrasting their abilities to solve disclosure dilemmas. We conclude by discussing the scholarly and policy implications of our argument.

The Challenge of Nuclear Proliferation

When the model is applied to the nuclear weapons context, the relevant actors become a state that may have detected a violation of nuclear norms, the potential violator, and the international community. The potential violator is a state that may have breached its obligations under the Nuclear Nonproliferation Treaty (NPT), which lies at the heart of the nuclear non-proliferation regime. This treaty legally binds signatories to three elements: 1) states that did not possess nuclear weapons prior to 1967 may not obtain them; 2) the five states that tested nuclear weapons prior to 1967 may not help other states obtain them and will work toward disarmament; and 3) non-nuclear weapons states are guaranteed help with energy development and civilian nuclear technology. In practice, violations constitute breaches of the first obligation.

[7]Hibbs, Mark. "Has the IAEA's Information Become Politicized?" *FP*. December 10, 2012.

Why might states attempt to renege on their commitments to the treaty? The extant literature highlights a host of other factors that drive this decision. States may do so when they are presented with the opportunity to proliferate (Monteiro and Debs, 2014*a*), when they face national security threats (Betts, 1977; Thayer, 1995), when the domestic polity supports it (Solingen, 2009), or due to their national identities (Hymans, 2006) and institutions (Walsh, 2001).[8]

Regardless of states' motivations, these activities represent a clear threat to many observers. The set of states most worried about a given state's nuclear development tend to be the state's adversaries and neighbors, since a nuclear weapon in the region could spark instability, or could be used against them in the future. Nuclear development can upset the regional balance of power and can cause surrounding states to seek a bomb as well or to bolster their conventional weapons capabilities. States outside of the region or that are not adversaries of the proliferator may still worry about the implications of a violation on a nuclear arms race more broadly, or their potential involvement in a conflict between the proliferator and one of their allies.

Intelligence and Nuclear Proliferation

How do informed states detect potential violations? In general, states with sophisticated intelligence capabilities lead the international community in its efforts to stop nuclear proliferation. Intelligence therefore may detect intentions to develop nuclear weapons, or bomb-building activities at a variety of phases of its development since obtaining a nuclear weapon is a lengthy process. States must monitor many types of activities: developing a bomb through an indigenous process requires multiple steps to reach a nuclear capacity, as states explore, pursue and then eventually acquire such a weapon. For example, intelligence could determine that a state has purchased heavy water or reprocessing plant parts or was stockpiling uranium, or it could detect a site where the bomb would be built which might include construction, workers, piles of dirt, cement pouring, and machinery.

These activities can be discovered using a variety of methods. As described previously, intel-

[8] See Sagan (2011, 2012) for overviews.

ligence collection can be divided into technical and human sources, with the former drawing on imagery, signals, and measurement techniques. Nuclear-related intelligence that has entered the public sphere through intentional disclosure or leaks over the years suggests that all intelligence collection methods can play a role in monitoring nuclear developments abroad.[9] Two of the most sensitive methods of the early period of the nuclear regime, in the 1950s and 1960s, were U-2 spy planes and communications intelligence. For example, the U.S. uncovered information about Israel's potential nuclear sites through communications intelligence and obtained imagery of the sites via its U-2s. In the last few decades, breakthroughs have been achieved by Western governments targeting programs in North Korea, Syria, and Iran using advanced satellite imagery, analyzing intercepted documents in laptop computers obtained through human sources, and performing careful monitoring of nuclear-related emissions (Richelson, 2007).

The Benefits of Intelligence Sharing

Sharing intelligence findings about aspiring nuclear states has two primary benefits for preventing proliferation. First, sharing intelligence allows a state to corroborate and supplement what it knows. Many states invest resources in building sophisticated intelligence capabilities yet even the most well-informed states often operate with significant uncertainty regarding nuclear targets. The United States, for example, has consistently maintained strong intelligence advantages over most other states in the international system regarding nuclear activities abroad. Yet despite its sophisticated capabilities, frustration with uncertainty about the pace and extent of nuclear development has been common in the history of U.S. intelligence assessments. Gaps in U.S. nuclear intelligence can lead it to miss key nuclear developments in other states which other states with different sources and methods can detect. For example, a Soviet satellite found a potential nuclear test site in South Africa in 1977, which the U.S. did not know about. In general, the Soviets monitored Eastern European states more heavily than the U.S. did, and thus had the ability to detect their

[9] States could hypothetically purchase a nuclear weapon triggering fewer tripwires of intelligence detection. Other means of detection – through sources monitoring international criminal networks, for example – may then play a role.

activities first (Coe and Vaynman, 2015). In several regions, moreover, the United States has historically lacked powerful human intelligence capabilities, which represents an area in which many other states are stronger. This often allows regional neighbors and American allies with advantages in human sources to detect nuclear activities and important technical details with greater skill.

The U.S. and other intelligence collectors can therefore benefit from triangulating privately gathered intelligence about a suspected proliferating state with what each other knows. Ad hoc bilateral intelligence sharing between states is one mechanism; yet this requires faith that the intelligence recipient will protect sensitive information. The result tends to be highly selective intelligence sharing limited to trusted allies and intelligence liaison partners. In the aggregate, states lacking a mechanism to pool intelligence findings while protecting sources and methods face an inefficient informational outcome in which most states most of the time keep their nuclear intelligence private. This leaves states and non-state actors poorly equipped to prevent and reverse progress in nuclear infrastructure and weapons development.

The second benefit of sharing intelligence is that it tends to improve political options for taking action based on what is known about an aspiring nuclear state. Sharing intelligence can be essential to building a multilateral coalition to support more intrusive nuclear inspections, punitive economic sanctions, or even military action. We represent this in the model when the observer decides whether to punish the potential violator. The observer is a state that is uninformed about whether a violation took place. If it knew that a violation occurred, it would want to punish the state, and if it knew conclusively that one did not, it would not desire to inflict a punishment. While these states have biases of their own, we simply assume that learning of a violation increases their desire to punish the violating state. The observer cares about punishment because the more states that acquire nuclear weapons, the less safe the observer's world becomes. States with nuclear weapons could use them or threaten to use them against the observer or the observer's allies. Or, if such a state became nuclear-capable, it could destabilize the region and make other states more apt to acquire these weapons, making it less safe for the observer.

Of course, the observer is not strictly necessary for punishment to occur; put differently, mul-

tilateral action is not the only way to address nuclear proliferation. A state that detects suspicious activity can take measures to prevent proliferation unilaterally. For example, such a states could inflict damage through the loss of security guarantees (Monteiro and Debs, 2014*b*), cuts to foreign and military aid, trade restrictions, asset freezes, investment restrictions, or even preventative attack (Fuhrmann and Kreps, 2010). The U.S. administered many of these penalties, or threats of these penalties, in response to Taiwan's and South Korea's nuclear programs (Miller, 2014; Solingen, 2012). However, there are substantial limits to such an approach and these efforts alone are often unlikely to dissuade a violator. A single state rarely possess the leverage to persuade another state to relinquish its nuclear program. To administer effective sanctions, multilateral efforts are typically necessary, such as those that were effective in rolling back Algeria's and Libya's programs. Intelligence sharing facilitates this process. If done credibly, sharing intelligence can persuade or coerce a greater number of states to support verification and enforcement actions. As a result, the suspected violator will face unified and costly penalties, slowing progress and potentially resulting in reversal of its program.

Sources and Methods and Credible Revelation

While multilateral punishments are most effective, states may not be able to assemble multilateral coalitions if they cannot assure potential coalition partners that they are acting on the basis of sound intelligence. But how can a state convince other states of its claims? After all, informed states possess biases, typically favoring the punishment of their adversaries or states with opposing interests, but not of their allies or states with complementary interests. Thus, informed states have incentives to lie about how reliable their intelligence is, overstating their claims when they have private incentives to administer harsh penalties.

As featured in the model, one way the informed state can enhance its credibility is to release the sources and methods it used to obtain its information. For example, rather than two states swapping their suspicions about the nuclear activity of a third state, one might share satellite photographs of a concealed nuclear facility that substantiates the idea. However, while revealing sources and

methods can enhance the informed state's credibility, doing so is costly. By demonstrating the origins of the intelligence, the informed state also reveals a good deal of information about its intelligence-collecting capabilities–information that it often prefers to keep secret. Most obviously, if the violating state learned of the informed state's sources and methods, it would alter its nuclear activities to avoid detection. For example, if the violator learned that a nuclear site was discovered by satellite, it could move the site underground to avoid future detection. Or, if the violator knew that its activities were detected through signals intelligence, it might alter the way its officials communicate.

More generally, the informed state typically seeks to avoid any other state learning of its intelligence capabilities since today's friend could become tomorrow's target, or could leak the information to another state. Even if today's friend were to remain a friend, states spy on their allies routinely, and prefer that these states are not aware of their activities. Thus, revealing sources and methods behind the detection of nuclear activities can disrupt intelligence-collection more generally–even beyond the nuclear realm.

While states occasionally share sources and methods if they trust each other not to expose them–as the United States and Britain sometimes do–such sharing is frequently impossible if states do not trust each other to do so. For example, when the United States wanted to share intelligence about Iran's nuclear program, the *New York Times* reported, "The United States rarely shares raw intelligence outside a small circle of close allies. But it decided to disseminate a shortened version of the secret warhead briefing. Mr. Joseph and his colleagues presented it to the president of Ghana and to officials from Argentina, Sri Lanka, Tunisia and Nigeria, among other nations. But the administration felt uncomfortable sharing any classified intelligence with another ring of countries. For them, it developed the equivalent of the white paper on Iraq that Britain and the United States published before the Iraq war. The 43-page unclassified briefing includes no reference to the warhead documents, but uses commercial satellite photos and economic analysis to argue that Iran has no need for nuclear power and has long hidden its true ambitions."[10]

[10]Broad, William J., and David E. Sanger. "Relying on Computer, U.S. Seeks to Prove Iran's Nuclear Aims." *The*

As shown in the model, when an informed state is biased, it cannot credibly convey its information, which in this case is intelligence about nuclear weapons proliferation. The more biased the informed state is, the less the international community and domestic observers can discern the truth. The state could reveal its sources and methods to prove its claims, but doing so is costly because it compromises the state's ability to continue its intelligence collection.

Thus, the disclosure dilemma manifests itself in the nuclear domain. Informed states face a difficult choice: either share raw intelligence and jeopardize intelligence collection or only report intelligence-based conclusions without reporting the source or method of reaching that judgment. Unless states reveal their sources and methods, the conclusions they share will likely be considered suspect, as they are prone to manipulate them for political purposes. In particular, states may release selective pieces of intelligence, express greater degrees of certainty about their intelligence's veracity than is warranted, or even outright lie about the intelligence they possess in order to achieve political ends.

Example: Nuclear Intelligence and Iraq

The diplomacy around suspected Iraqi nuclear activity in 2002 and 2003 illustrates both kinds of responses to the disclosure dilemma in the nuclear domain and, in particular, the credibility problem states face when making intelligence-based claims without disclosing sources and methods. In the lead up to the Iraq War which began on March 20, 2003, the United States attempted to drum up support for the invasion of Iraq and to build a "coalition of the willing." The primary justification given publicly for going to war was that Iraq possessed weapons of mass destruction. Building support for the war thus rested on the United States's ability to convince other states that Iraq really did have this capability. These efforts culminated in Secretary of State Colin Powell's presentation to the UN Security Council, in which he tried to demonstrate this, playing audio tapes and showing satellite photographs that he said revealed Iraq's "disturbing patterns of behavior" and

New York Times. November 13, 2005.

its "policy of evasion and deception" (ElBaradei, 2011, 3).[11]

In actuality, the United States did not possess such intelligence. The NSA put together a file of all of its intelligence related to the issue and determined that several intercepts of calls among Iraqi Republican Guard commanders represented their most compelling evidence. However, the calls were selectively chosen and turned out to be highly ambiguous (Aid, 2010, 243). Powell also referred to intercepts of low-level emails and telexes that indicated that Iraqi companies were trying to purchase high-speed balancing machines which could be used for uranium enrichment, but could also be used for many other civilian manufacturing processes. NSA and CIA analysts have stated that the intercepts were inconclusive; however, Powell called these intercepts evidence "that Iraq front companies sought to buy machines that can be used to balance gas centrifuge rotors. One of these companies also had been involved in a failed effort in 2001 to smuggle aluminum tubes into Iraq" (Aid, 2010, 237). A U.S. intelligence official stated that the intercepts "provide no evidence that the suspected terrorist [al Zarqawi] was working with the Iraqi regime or that he was working on a terrorist operation while he was in Iraq" (Aid, 2010, 241) and a former NSA official stated, "There wasn't much there, and there certainly was no smoking gun" (Aid, 2010, 242).

Further, much of the information revealed was misleading and inaccurate. For instance, a key allegation in Powell's speech was that Iraq had tried to import high-grade aluminum whose tolerance Powell claimed "far exceeds U.S. requirements for comparable rockets." Powell concluded that the aluminum was meant to produce nuclear weapons. Yet two days in advance of Powell's speech, Powell received a memo from U.S. intelligence which stated that the U.S. used the same type of aluminum for its seventy-millimeter tactical rockets. Similarly misleading was Powell's claim that Iraq attempted to buy uranium from Niger, which was later shown to be fake (ElBaradei, 2011, 62-3).

The U.S. claimed that it did not reveal all of its intelligence because it feared "that intelligence-gathering sources could be compromised."[12] However, the U.S. did not actually possess credible

[11]Note that the UN Security Council is not playing a role here except as a forum for states to congregate.

[12]Rajiv Chandrasekaran and Colum Lynch. U.N. Officials Say Intelligence to Prove US Claims Is Lacking, The Washington Post, 27 January 2003, p. 12.

intelligence, and sharing everything it knew would expose this truth. Indeed, other nations suspected that this was the case; that is, many believed that the United States possessed considerable bias against Iraq, and therefore cherry-picked its intelligence to spin it for political purposes. Many states surmised that the U.S. actually had far less information than it claimed. For example, the IAEA obtained little intelligence from the U.S. and after the head of the organization complained about this to French president Chirac, Chirac stated, "You know why you dont get the information...It is because they dont have any" (ElBaradei, 2011, 66). Similarly, "to the inspection community, [Powell's] presentation was primarily an accumulation of conjecture, an alignment of unverified data interpreted according to a worst-case scenario" (ElBaradei, 2011, 3).

While the U.S. withheld sources for many of its intelligence-based claims about Iraq, those instances in which their sources were made public demonstrate the downsides in sharing raw intelligence. In revealing the intelligence the United States did possess, the U.S. compromised key sources and methods. The NSA tried to prevent the U.S. from sharing the intercepts, as it worried that doing so would alert the Iraqis to the U.S.'s means of gathering intelligence. Indeed it did, as two weeks after Colin Powell's presentation, Iraq took countermeasures. In particular, the founder of al Qaeda in Iraq, al Zarqawi, quit using his cell phone, which represented the cut off of a crucially important source the NSA had been monitoring. Further, the government of Iraq turned off all of Iraq's telephone services, making a point to outlaw satellite and cellular phones, which "closed off the last low-level sources of SIGINT that were then available to NSA about what was going on inside Iraq" (Aid 2010, 245). In this case, however, since the U.S. did not possess strong enough intelligence to prove its case, even revealing sources and methods was not enough to convince many members of the international community of the veracity of its claims.

International Organizations and Nuclear Intelligence

Absent an IO, states are forced to reveal their sources and methods in order leverage the political benefits of sharing a particular piece of intelligence. Often, exposing sources and methods is too

costly; this is the essence of the disclosure dilemma. The result is privately known or narrowly circulated intelligence. Opportunities for pooling intelligence are lost and the political process of building multilateral support for preventing proliferation is more difficult. In short, inefficiencies in nuclear intelligence-sharing due to the disclosure dilemma risk missed opportunities for the enforcement of non-proliferation goals. Our central contention is that states and the non-proliferation regime would often collectively benefit from intelligence-sharing, but this sharing is impractical if states lack an avenue for doing so without endangering sources and methods.

In part as a response to the problem of inefficient intelligence-sharing in the nuclear domain, in recent decades member-states and IOs themselves have begun experimenting with authorizing IOs to receive and protect nuclear intelligence. Sharing intelligence with an IO allows it to perform three basic functions. First, an IO that receives national intelligence may be able to authenticate the claim by pooling intelligence conclusions from multiple third parties. Second, if equipped with verification or similar investigatory powers, an IO that receives intelligence can proactively seek additional information from the suspected nuclear proliferator and other sources to authenticate the original claim and even expand on it. Third, an IO that receives national intelligence from multiple states can authenticate the intelligence while protecting sources and methods. An IO capable of receiving, gathering, and protecting intelligence can therefore serve an important role in vetting intelligence and supplementing it. When it works as designed, this process creates a richer information environment and makes the political process of building a multilateral coalition in favor of stronger verification and enforcement easier.

We consider the three important IOs in the nuclear domain in particular: the IAEA, UNSCOM, and the CTBT. [NOTE: this version only features IAEA.]

IAEA

Consider the case of an IO that is particularly critical to monitoring the nuclear regime: the IAEA. Non-nuclear weapons states are monitored by the IAEA to certify that their civilian activities are not repurposed toward nuclear weapons development. The institution's mandate is to prevent states

from using nuclear energy for military purposes, and to help the IO fulfill its mission, states provide it with proprietary records, which the IAEA then examines. The IAEA also inspects the declared facilities of member states to ensure that their observed nuclear materials match those that the states have declared. Finally, over 140 states have signed the 1997 Additional Protocol which also allows undeclared facilities to be inspected.

As highlighted previously, the nuclear realm is rife with intelligence-sharing problems. However these issues can be overcome if international organizations can commit to not disclosing member states' intelligence. The IAEA refers to such intelligence as "third-party information" which includes "intelligence obtained using national technical means, nuclear trade-related information on exports and supplier denials, and details of relevant domestic political developments (Ogilvie-White, 2014, 325). The IAEA has not always had the authority to receive and use this kind of private information. Yet after the discovery of Iraq's nuclear program in 1991, the IAEA became more focused on discovering clandestine activities as opposed to conducting routine inspections.[13] During this period, the IAEA shifted from its role as an "observer and analyst" to an "active seeker of information" (Ogilvie-White, 2014, 325-6). Approved measures include: the early provision of design information (1992); the introduction of a voluntary reporting scheme (1993); the endorsement of remote monitoring and environmental sampling (1995); the approval of the Additional Protocol (1997); the creation of a satellite imagery analysis unit (1999); the development of integrated safeguards (1998-2001); the launch of the IAEA nuclear trade analysis unit (2004); and the revision of the small quantities protocol (2005).

Part of the IAEA's more ambitious approach includes the use of intelligence. Thus, Director General Blix pushed for an expansion of the IAEA's capacity to analyze and act on intelligence. This transformation included new "in-house analytical capabilities in the Department of Safeguards, which now allows IAEA experts to assess a variety of information sources," which include a satellite imagery analysis team and nuclear forensic work (Ogilvie-White, 2014, 333-4). Moreover, since 1992, states have often shared intelligence with the IAEA to aid in the enforce-

[13]Hibbs, Mark. "Has the IAEA's Information Become Politicized?" *FP*. December 10, 2012.

ment of the nuclear regime. In February 1992, the IAEA announced that it believed itself to be able to obtain sensitive intelligence from third-party sources (Carmody, 1994, 268). For example, in Iraq, the IAEA used considerable third-party intelligence (Carmody, 1994, 277) including satellite imagery from the CIA (Richelson, 2007, 519-21). Similarly, the IAEA Director General, Mohamed ElBaradei, notes that an intelligence agency provided satellite photos of Iran's Dair Alzour nuclear site in Syria, and another intelligence agency shared other photos of the building containing the reactor. These photos allowed the IAEA to gather additional information and to further its investigation (ElBaradei, 2011, 225).

As a result of these efforts, member states have more frequently shared their intelligence with the IAEA in recent years. For example, ElBaradei describes intelligence agencies providing reconnaissance photographs and information about suspicious sites in Iraq, satellite photos of an industrial facility in Iran, and satellite images of the building housing a nuclear reactor in Syria (ElBaradei, 2011, 12, 104, 225). This increased information provision has allowed the IAEA to rely more heavily on national intelligence collection, though the *FP* notes that "If this data isn't rigorously vetted and handled carefully, the IAEA's technical and political credibility will be seriously compromised."[14]

When these new systems function well, the IAEA can ameliorate the disclosure dilemma in the nuclear domain in each of the three ways highlighted in the model. First, the IAEA can collect intelligence conclusions from multiple informed states. Second, it can vet the intelligence and then scrub it to remove any information that would compromise sources and methods before sharing it with member states. Third, it can collect additional information and its own intelligence based on tips and hints from member states seeking to protect their sources and methods. For instance, "A country might give the IAEA photos pointing to undisclosed activities somewhere else, but beforehand the evidence will be degraded."[15] The IAEA would then seek to authenticate this information. The IAEA and its member states are thus able to improve the quality of international public goods

[14]Hibbs, Mark. "Has the IAEA's Information Become Politicized?" *FP*. December 10, 2012.

[15]Hibbs, Mark. "Has the IAEA's Information Become Politicized?" *FP*. December 10, 2012.

in the nuclear realm based on such intelligence.

Corroboration from Multiple States

As we argue above and in the model, an important function that IOs can perform when receiving disclosures from multiple states is to pool and triangulate intelligence. In the nuclear realm, the IAEA has played this role, corroborating individual intelligence submissions from several states to arrive at a conclusion about the nuclear activity of a suspected proliferator. Recent reporting on the diplomacy surrounding Iran's nuclear activities suggests that this pooling and triangulation of intelligence has been critical to the IAEA's activities. As Hibbs notes, Iran is a "high-profile case where the IAEA is using a lot of third-party information to develop a complete picture of a country's nuclear program." Investigation of a possible military dimension in Iran's program has been largely based on "national intelligence information provided by more than 10 IAEA member states" (Shea, N.d.). In addition to intelligence from Western governments like the United States, recent reporting suggests even states like China have passed along intelligence to the IAEA regarding Iran's program.[16]

Because the IAEA helps to authenticate the nuclear regime, and does so in part by facilitating intelligence sharing, it is in the long-term interest of powerful states to have the IAEA corroborate intelligence through multiple sources. Powerful states like the U.S. understand the importance of diversifying sources; Brown (2015, 125) notes, "US officials...recogniz[ed] that the appearance of excessive US influence would compromise the basis for Agency authority with others." The IAEA itself has therefore, at times, refused to base its activities and conclusions on any single source. For example, in 1994 the IAEA was tasked with ensuring that South Africa had dismantled its nuclear weapons program. Although the United States offered intelligence to assist with this endeavor, the IAEA rejected its help due to a desire to preserve "institutional impartiality and integrity" (Brown, 2015, 108). The agency also sought to authenticate third party information because, particularly in the case of the Iraq War, it feared that if the war had been caused in part by an IAEA report based

[16] Associated Press, "Nuclear Watchdog Gets China's Iran Intel," April 2, 2008.

on inaccurate intelligence from a single third-party, it would damage the Secretariat's credibility (ElBaradei, 2011). In particular, Russia provided the IAEA with information which showed that the U.S.'s claims that Iraq had restarted its nuclear program were not true.[17]

Authenticating Intelligence

The model demonstrates that when the IO's bias is not too large, states can benefit by disclosing intelligence conclusions even if they do not reveal their sources and methods, because these disclosures allow the IO to authenticate the validity of the claims. As we noted previously, an IO can authenticate intelligence with its own technical experts both by using its existing information base and by collecting additional information. In the nuclear realm, the IAEA has received intelligence and has been empowered to authenticate the validity of such submissions, albeit in a limited way. Indeed, member states recognize the importance of this capability for the IAEA, as recent reforms have given the IAEA improved technology and expertise to facilitate its consumption of intelligence disclosures and its ability to gather supplemental information on its own (Ogilvie-White, 2014, 324).

In several key cases in the last three decades, the IAEA has gone to great pains to use its expertise and information gathering capabilities to vet intelligence disclosed to it by member-states. To do so, the IAEA has "described procedural safeguards that guarantee its objectivity and neutrality when using such information" (Coppen, 2015). It has promised that "all information provided to the Agency goes through a rigorous process of internal IAEA corroboration" (Heinonen, 2013), which helps the IAEA protect its reputation for impartiality.

For example, in 2005 the U.S. told the IAEA that it had information from a laptop computer showing that Iran had nuclear weaponization studies that pertained to uranium conversion, high explosives testing, and the alteration of a missile reentry vehicle to carry a nuclear warhead. The documents included over a thousand pages of descriptions of experiments and computer simulations. The U.S. wanted to raise international pressure on Iran, and thus needed the IAEA to

[17]Hibbs, Mark. "Has the IAEA's Information Become Politicized?" *FP*. December 10, 2012.

authenticate its claims to gain other countries' support. Countries that already trusted the U.S.–Britain, France, and Germany–had accepted the intelligence, but others remained skeptical. The *New York Times* reported, "In part, that is because American officials, citing the need to protect their source, have largely refused to provide details of the origins of the laptop computer" beyond saying that they obtained it in mid-2004 from a longtime contact in Iran who receive it form another person who was likely dead. A senior European diplomat stated, "I can fabricate that data...It looks beautiful, but it is open to doubt."[18] To try to muster international support and convince the UN to punish Iran, the U.S. showed its intelligence to the IAEA, though it did not release its source. The U.S. particularly wanted the IAEA to independently authenticate that the documents were not fraudulent, which was a concern of many states. The IAEA thus set about authenticating these claims, as ElBaradei said that the IAEA must "establish the veracity, consistency, and authenticity of any intelligence."[19]

At times the IAEA's *inability* to authenticate an intelligence-related claim shows the importance this capability. For example, considerable uncertainty existed over whether American claims about Iraq's weapons of mass destruction in 2002 were accurate after Colin Powell's presentation to the United Nations. The two IOs tasked with authenticating Iraqi nuclear compliance – the IAEA and UN weapons inspectors – investigated the U.S.'s assertions. The IOs were charged with authenticating the U.S.'s intelligence so that the international community could determine whether to believe the U.S. However, the IOs found little evidence to corroborate the U.S.'s statements and even determined that some of them were conclusively false. A critical component of the American claim about Iraq's nuclear ambitions was the Iraqi attempt to buy uranium from Niger. The IAEA received the papers allegedly demonstrating this on the day of Powell's speech and determined in a few hours that they were fake, complete with falsified signatures, inaccurate information, and logical inconsistencies. ElBaradei, reported to the Security Council that the documents were "not

[18]Broad, William J., and David E. Sanger. "Relying on Computer, U.S. Seeks to Prove Iran's Nuclear Aims." *The New York Times.* November 13, 2005.

[19]Broad, William J., and David E. Sanger. "Relying on Computer, U.S. Seeks to Prove Iran's Nuclear Aims." *The New York Times.* November 13, 2005.

authentic" (ElBaradei, 2011, 62-3). Similarly, U.S. claims of nuclear materials hidden on an Iraqi chicken farm were not authenticated after UN weapons inspectors and the IAEA carried out inspections.[20] ElBaradei states that he "could not reach a verdict on these allegations [about Iraq's program]...without first being able to verify the authenticity of the documents passed on by the U.S. intelligence. Nor would I have done so with any other country" (ElBaradei, 2011, 281).

In contrast, the IAEA's successful authentication of intelligence about Iraq's nuclear ambitions did substantiate reason for concern about Iraqi laser enrichment research. The *Washington Post* reported in January 2003 that "one of the inspectors' biggest successes so far – the surprise search of an Iraqi scientist's home where they found more than 3,000 pages of sensitive documents, many of them about a uranium enrichment research program – was the result of an intelligence tip from a foreign government."[21] Brown (2015, 172) argues, "Iraq demonstrated IAEA authority to provide verification that was relatively more objective, effective, and cost efficient. As long as the war that began over US allegations of Iraqi weapons programs continued, it served as a continuous reminder of the cost to the superpower and to the international community of ignoring IAEA authority." Further, this insistence on authentication led the IAEA to be trusted much more than it otherwise might have been. Indeed, "the Agency could assert new political authority because it had supplied the requisite policy partiality without the external help of others who also had the technical capacity" (Brown, 2015, 123).

However, an important constraint on fulfilling this role has been the IAEA's limitations regarding its technical expertise, which can make it impractical for the IAEA to authenticate certain kinds of nuclear-related intelligence. Specifically, the IAEA is technically only allowed to authenticate intelligence when nuclear material is involved so that the IAEA can take measurements and environmental samples. The IAEA cannot validate the authenticity of a document, which can restrict its ability to solve the disclosure dilemma. ElBaradei summarized the problem: "We were spied on

[20]Rajiv Chandrasekaran and Colum Lynch. "U.N. Officials Say Intelligence to Prove US Claims Is Lacking," The Washington Post, 27 January 2003, p. 12.

[21]Rajiv Chandrasekaran and Colum Lynch. "U.N. Officials Say Intelligence to Prove US Claims Is Lacking," The Washington Post, 27 January 2003, p. 12.

by the same intelligence agencies we relied upon to inform us when they detected nuclear anomalies; we were given selective intelligence information, which was often difficult to authenticate" (ElBaradei, 2011, 239). In such a case, when the IAEA cannot independently authenticate a single country's allegations, it tends not to endorse them.

Sources and Methods Revelation and the Threat of Leaks

The model shows that when deciding whether to reveal its sources and methods along with its information to an IO, a state weighs the benefits of releasing that information against the possibility that the information will be leaked. Because of this potential for leaks, "intelligence agencies are concerned that disclosure of information shared with the IAEA will allow proliferators to improve their countermeasures" due to the revelation of sources and methods (Hertzberg, 2010, 10). States may decide that they trust the IO enough to reveal some sources and methods, but not enough to reveal them fully; they can then share intelligence with an IO while partially obscuring sources and methods, after which the IAEA can follow up on the pieces of the claim for which sources and methods were not disclosed, or were partially exposed. The decision thus does not involve a simple dichotomy but rather a spectrum of specificity about the source of an intelligence-based claim. For instance, after Israel bombed Syria's Dair Alzour facility in 2007, the U.S. claimed that it had been the site of a nuclear reactor. The U.S. shared intelligence with the IAEA secretariat, after which the IAEA wanted to visit Dair Aizour and other related sites to gather additional information. The IAEA explained to the head of the Syrian Atomic Energy Commission that the IAEA "had seen satellite photos that showed equipment being moved from the destroyed site to other locations, so it was important to verify the nature of these three other sites" (ElBaradei, 2011, 223).

The IAEA cannot force states to disclose intelligence; it must try to reassure states that their information will not be leaked and then hope that states share it. As ElBaradei notes, "The IAEA is not a spy agency. Our inspectors do not engage in espionage or use deception to get at the truth. We do not have access to the databases of police forces, Interpol, or national intelligence agencies, unless these organizations choose to make relevant information available." (ElBaradei,

2011, 18). To convince states that it will protect their information, the IAEA has improved its management of national intelligence overtime; while the IAEA received little intelligence from members before 1992 (Feldman, 1997, 155), it now has a well-developed system to assess and protect information.[22] As a result, states often reveal at least some of their sources and methods.

Consider the U.S.'s decision about whether to provide intelligence to the IAEA regarding North Korea's nuclear program. In this case, once the threat of North Korea's program grew large enough, the U.S. stepped up its intelligence-sharing with the IAEA, partially revealing sources and methods to the institution. The U.S. did so because the IAEA could then certify the accuracy of this intelligence and pressure the North Koreans to come into compliance with their treaty obligations. The IAEA thus made the U.S.'s allegations credible and could better enforce the regime.

More specifically, beginning in 1976, the CIA provided satellite images of North Korean nuclear facilities to the IAEA. Sharing with the IAEA was limited, as the CIA withheld imagery from advanced systems to avoid compromising intelligence gathering. In February 1992, Director of Central Intelligence Robert Gates informed the House Foreign Affairs Committee that North Korea was hiding a nuclear program despite pledging denuclearization, saying, "We have some information that I can't go into here in this setting" (Richelson, 2007, 519-21). The CIA began to provide more intelligence to the IAEA after the international organization found discrepancies in North Korean statements about its reprocessing activities. Towards the end of 1992, the CIA informed the IAEA that its satellite imagery revealed North Korean workers constructing a new nuclear waste storage site across from an existing site. The CIA also had satellite imagery showing workers laying pipes between the two facilities without the IAEA's knowledge.

When North Korea refused an IAEA request for samples from a suspected waste site, IAEA director Hans Blix asked the U.S. to allow the presentation of their satellite images at the IAEA Board meeting in February 1993. The IAEA was previously only allowed to view the photos at the U.S. mission. The State Department supported this request, but mid-level CIA analysts opposed sharing the images for fear of compromising sources and methods of intelligence. CIA Director

[22]Hibbs, Mark. "Has the IAEA's Information Become Politicized?" *FP*. December 10, 2012.

Gates overruled the analysts, believing that American national security interests would be served by aiding the international inspection agency. After Gates left his position, CIA analysts tried to sabotage the image-sharing plan by providing photographs that were intentionally obscured to conceal American satellite capabilities. Fearing that providing inadequate imagery would prevent the IAEA from pressuring North Korea, newly inaugurated President Clinton allowed the imagery to be distorted, but not totally obscured (Richelson, 2007, 519-21).

On February 22, 1993, the thirty-five member states of the IAEA Board of Governors met in Vienna and viewed the black and white pictures of the storage facility under construction. This represented a significant milestone. While the U.S. had changed the images' resolution to "disguise their actual surveillance capability," the IAEA could still authenticate that the intelligence was accurate and the presentation impacted the diplomatic debate. "This was the first time in the history of the IAEA that the Secretariat had shared information supplied by Member State intelligence in a Board setting. Member States had historically been very uneasy about the Agency's use of any information obtained through national intelligence agencies" (ElBaradei, 2011, 43). Though North Korean representative Ho Jin Yun accused the IAEA of utilizing intelligence from a "third power," the IAEA Board ignored this statement, and approved a resolution demanding the inspections of North Korean waste storage sites (Richelson, 2007, 519-21).

Although the U.S. feared leaks by the IAEA, the probability of leaks was low enough, and the benefits of sharing the intelligence were high enough, that the U.S. disclosed its sources and methods anyway. However, when states believe that the risks of leaks outweigh the benefits of sharing, they do not share their intelligence with the IAEA. In general, intelligence sharing in this area "remains limited by the concern of states that possess relevant classified information that their sources and methods may be compromised" (Feldman, 1997, 155). Many have also noted that the IAEA's culture of transparency has inhibited its ability to protect sensitive information, leading to numerous stories of the IO mishandling of intelligence (Wiebes, 2003, 24).

Indeed, several specific instances feature the IAEA leaking information, as it did during its efforts to monitor Iran's nuclear program. In 2013, intelligence was provided to the IAEA which

purportedly demonstrated that Iran's program contained a military dimension. While the IAEA did not release this information under director-general ElBaradei because it could not be verified, director general Amano chose to do so. Former IAEA director-general Hans Blix objected to the decision, stating that the IAEA needed to authenticate the intelligence. He asserted, "It may be that they are exaggerating [the intelligence information]. There's also a danger in telling us without revealing the actual sources. One has to be very careful about that." He continued, "The IAEA must not be the prolonged arm of intelligence agencies."[23]

In a similar incident involving Iran, a 2012 IAEA report which suggested that Iran was developing its nuclear capability according to national intelligence was leaked to the Associated Press by "officials from a country critical of Iran's atomic program to bolster their arguments that Iran's nuclear program must be halted before it produces a weapon." The leak "underscores the need for the IAEA to carefully manage information from third parties, especially information that makes sensitive allegations on the basis of the intelligence findings of member states."[24]

These concerns often lead states to withhold sources and methods information, as occurred in response to Syria's nuclear program. As noted by Fuhrmann (2012, 220), "The United States, for instance, shared information with the IAEA on the al-Kibar site – but it did so more than six months after the Israeli strike. In the aftermath of the attack, the United States and Israel reportedly 'went to great lengths to prevent others from finding out where the site was,' presumably because they wanted to protect their intelligence sources and methods. The lack of information sharing stymied the IAEA's efforts to obtain a complete picture of Syria's nuclear activities."

Further, returning to the case of Iraq's nuclear program in the 2002-2003 period, the U.S. at times justified its decision to strictly limit intelligence sharing to the IAEA based on concerns about information security. "U.S. intelligence-sharing may also be limited because Hans Blix, the chief weapons inspector, rebuffed a request by the Bush administration late last year to appoint a senior U.S. official to the U.N. inspection agency to handle the flow of sensitive intelligence. After

[23]Tirone, Jonathan. "Iran Spy Data Need Checks as Amano Prepares for New Term." *Bloomberg Business*. March 7, 2013.

[24]Hibbs, Mark. "Has the IAEA's Information Become Politicized?" *FP*. December 10, 2012.

pressure to step up cooperation, the administration agreed to supply some secret information to the U.N. chief of intelligence, James Corcoran, a former deputy director of the Canadian Security Intelligence Agency who has U.S. intelligence clearances, to test the United Nations' capacity to keep a secret."[25]

Method and Plan of the Book

The previous discussion features excerpts from a book in progress. The book uses a multi-method approach to investigate how states form strategies of intelligence disclosure, how IOs work to both facilitate cooperation and allow states to project power through intelligence sharing, and the effect of this behavior on interstate relations. The book's chapters develop a game-theoretic model that formally lays out the argument, analyze survey responses from a novel survey experiment, and provide in-depth examinations of real-world dynamics. In addition to the IOs we focus on in our analyses, we also offer guidelines for determining the extent to which our claims pertain to other international bodies.

After introducing the key concepts in the introductory chapter, chapter two, gives an account of the existing literature and debates about information in IOs, including a description of the kinds of information that shape states' perceptions and contribute to compliance. We explain how intelligence works, providing a typology of the types of intelligence sources and the issues about which states collect intelligence. Further, we discuss variation in intelligence gathering by offering a brief history of intelligence collection and the ways in which intelligence capabilities differ among states. Finally, we show how intelligence could potentially foster global public goods in a variety of domains by allowing states to identify problems and detect compliance-related activities, detailing a series of examples of lost cooperation opportunities along the way.

In chapter three, we develop the theoretical argument in detail. Using the concepts discussed in the introduction to motivate the set-up and assumptions, we present a game-theoretic model featur-

[25]Rajiv Chandrasekaran and Colum Lynch. "U.N. Officials Say Intelligence to Prove US Claims Is Lacking," The Washington Post, 27 January 2003, p. 12.

ing two states and an observer. Prior to the game, one of the states may have violated international law. The other state can discern whether this has occurred using its sophisticated intelligence capabilities, but the observer cannot. The informed state may send a costless message to the observer regarding whether a violation has occurred, after which the observer must decide whether to punish the potential violator or not. Alternatively, the informed state may reveal its sources and methods at a cost, in which case the observer can discern whether a violation has occurred. Revealing sources and methods thus constitutes a trade-off: such revelation allows the observer to impose a strong punishment against the violator, but the informed state could lose the ability to detect violations in the future, as states adjust to the disclosure. By contrast, if the detecting state holds back its intelligence, it cannot impose a severe penalty, but maintains the ability to detect future violations. We then show how an IO can authenticate intelligence claims and ameliorate this problem by ensuring that shared intelligence, including any sources and methods, will be protected. This allows the intelligence-sharing state to build support against the violator and still maintain the ability to uncover a violation the following period. We also address how the possibility of leaks from the IO influence state behavior. The chapter works up to a series of empirical predictions that distinguish our argument from alternative explanations.

Explaining states' strategies of information disclosure and the impact of international institutions is the manuscript's central ambition, and chapters three through seven investigate these core ideas in detail in a variety of domains including nuclear weapons, chemical weapons, the laws of war, peacekeeping, and Cold War politics. In each chapter, we describe the precise nature of the disclosure dilemma and then substantiate the model's prediction that international institutions can help to solve it. Taken together, the chapters provide case study analyses highlighting how and when a variety of IOs were able to do so, which occurs when they are able to provide credible vetting of information and provide high levels of security that the information will be protected. They also show common limitations and difficulties IO encounter in this endeavor. Our theory implies that the more IOs are endowed with these capacities, the more states share such intelligence and the better the IOs are able to function to achieve their mandates.

Chapter eight then tests the model's predictions through an analysis of responses from a unique survey experiment. The results show that the public trusts information that has been vetted by IOs when it would not trust the same information if it were provided by individual states unless sources and methods are revealed. Since states often decide whether to reveal sources and methods based on domestic political considerations, our findings lend strong support to the theory.

We conclude by discussing the argument's scope conditions, along with its normative, policy and scholarly significance. We lay out guidelines for determining when the theory applies to particular institutional settings, which depends on the nature of the information required for an institution to fulfill its mandate and the institution's ability to protect sources and methods. In addition, we consider several issue areas in which IOs do not currently protect sources and methods, and imagine what such an IO would look like in order to be effective. We close with implications for both scholarship and normative questions about transparency and accountability in the international system.

References

Aid, Matthew M. 2010. *The secret sentry: the untold history of the national security agency*. Bloomsbury Publishing USA.

Andrew, Christopher. 1979. "Governments and secret services: A historical perspective." *International Journal* pp. 167–186.

Andrew, Christopher M and C Andrew. 1995. *For the President's Eyes Only: Secret Intelligence and the American Presidency from Washington to Bush*. HarperCollins London.

Battaglini, Marco. 2002. "Multiple referrals and multidimensional cheap talk." *Econometrica* pp. 1379–1401.

Betts, Richard K. 1977. "Paranoids, Pygmies, Pariahs & Nonproliferation." *Foreign Policy* pp. 157–183.

Blanton, Thomas. 2007. The Struggle for Openness in the International Financial Institutions. In *The Right to Know: Transparency for an Open World*, ed. Ann Florini. Columbia University Press pp. 243–78.

Boeyink, David E. 1990. "Anonymous sources in news stories: Justifying exceptions and limiting abuses." *Journal of Mass Media Ethics* 5(4):233–246.

Brown, Robert L. 2015. *Nuclear Authority: The IAEA and the Absolute Weapon*. Georgetown University Press.

Carmody, Susan. 1994. "Balancing Collective Security and National Sovereignty: Does the United Nations Have the Right to Inspect North Korea's Nuclear Facilities." *Fordham Int'l LJ* 18:229.

Chesterman, Simon. 2006. "Shared secrets: Intelligence and collective security." *Lowy Institute Paper* (10).

Coe, Andrew and Jane Vaynman. 2015. "Collusion and the Nuclear Nonproliferation Regime." *Journal of Politics* .

Coppen, Tom. 2015. "Developing IAEA Safeguards: An Institutional Perspective on the State-level Concept." *Journal of Conflict and Security Law* p. krv004.

Crawford, Vincent P and Joel Sobel. 1982. "Strategic information transmission." *Econometrica:*

Journal of the Econometric Society pp. 1431–1451.

Dessein, Wouter. 2002. "Authority and communication in organizations." *The Review of Economic Studies* 69(4):811–838.

Ehring, Lothar. 2008. "Public Access to Dispute Settlement Hearings in the World Trade Organization." *Journal of International Economic Law* 11(4):1021–1034.

ElBaradei, Mohamed. 2011. *The age of deception: nuclear diplomacy in treacherous times*. Macmillan.

Feldman, Shai. 1997. *Nuclear Weapons and Arms Control in the Middle East*. MIT Press.

Fuhrmann, Matthew. 2012. *Atomic assistance: How Atoms for Peace programs cause nuclear insecurity*. Cornell University Press.

Fuhrmann, Matthew and Sarah E Kreps. 2010. "Targeting nuclear programs in war and peace: A quantitative empirical analysis, 1941–2000." *Journal of Conflict Resolution* .

Geser, Hans. 1992. "Towards an Interaction Theory of Organizational Actors." *Organization Studies* 13(3):429–451.

Grant, Ruth W and Robert O Keohane. 2005. "Accountability and Abuses of Power in World Politics." *American Political Science Review* 99(01):29–43.

Gupta, Aarti. 2008. "Transparency Under Scrutiny: Information Disclosure in Global Environmental Governance." *Global Environmental Politics* 8(2):1–7.

Heinonen, Olli. 2013. "IAEA Safeguards - Evolving to Meet Todays Verification Undertakings." *Conference Paper, Belfer Center for Science and International Affairs* .

Hertzberg, Michael. 2010. *Shining a Brighter Light on Dark Places: Improving the IAEA's Use of Intelligence Through Cooperation with NATO*. CSIS.

Hymans, Jacques. 2006. "The psychology of nuclear proliferation." *Identity, emotions, and foreign policy. New York, USA: Cambridge University Press* .

Johnson, Loch K. 2007. *Handbook of intelligence studies*. Routledge.

Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press.

Koenig-Archibugi, Mathias. 2004. "International Governance as New Raison d' Etat? The Case of the EU Common Foreign and Security Policy." *European Journal of International Relations* 10(2):147–188.

Krishna, Vijay and John Morgan. 2001. "Asymmetric information and legislative rules: Some amendments." 95(02):435–452.

Krishna, Vijay and John Morgan. 2008. "Cheap talk." *The New Palgrave Dictionary of Economics* 1:751–756.

Manning, Dean. 2000. *Srebrenica Investigation: summary of forensic evidence-execution points and mass graves*. United Nations International Criminal Tribunal for the former Yugoslavia.

Miller, Nicholas L. 2014. "The secret success of nonproliferation sanctions." *International Organization* 68(04):913–944.

Monteiro, Nuno P and Alexandre Debs. 2014*a*. "The Strategic Logic of Nuclear Proliferation." *International Security* 39(2).

Monteiro, Nuno P and Alexandre Debs. 2014*b*. "The Strategic Logic of Nuclear Proliferation." *International Security* 39(2):7–51.

Moranchek, Laura. 2006. "Protecting National Security Evidence While Prosecuting War Crimes: Problems and Lessons for International Justice from the ICTY." *Yale J. Int'l L.* 31:477.

Ogilvie-White, Tanya. 2014. "The IAEA and the International Politics of Nuclear Intelligence." *Intelligence and National Security* 29(3):323–340.

Pribanic-Smith, Erika J. 2012. "On the Condition of Anonymity: Unnamed Sources and the Battle for Journalism." *Journalism History* 38(1):57.

Richelson, Jeffrey. 2007. *Spying on the bomb: American nuclear intelligence from Nazi Germany to Iran and North Korea*. WW Norton & Company.

Sagan, Scott D. 2011. "The causes of nuclear weapons proliferation." *Annual Review of Political Science* 14:225–244.

Sagan, Scott D. 2012. "Why do states build nuclear weapons? Three models in search of a bomb.".

Scheffer, David. 2012. "All the Missing Souls." *A Personal History of the War Crimes Tribunals* .

Shea, Thomas E. N.d. "The Verification Challenge: Iran and the IAEA." *Arms Control Today*. Forthcoming.

Solingen, Etel. 2009. *Nuclear logics: contrasting paths in East Asia and the Middle East*. Princeton University Press.

Solingen, Etel. 2012. *Sanctions, statecraft, and nuclear proliferation*. Cambridge University Press.

Thayer, Bradley A. 1995. "The Causes of Nuclear Proliferation and the Utility of the Nuclear Non-Proliferation Regime." *Security Studies* 4(3):463–519.

Walsh, James Igoe. 2010. *The international politics of intelligence sharing*. Columbia University Press.

Walsh, James Joseph. 2001. Bombs unbuilt: power, ideas and institutions in international politics PhD thesis Massachusetts Institute of Technology.

Wiebes, Cees. 2003. *Intelligence and the War in Bosnia, 1992-1995*. Vol. 1 LIT Verlag Münster.

Wulfemeyer, K Tim. 1982. "The Use of Anonymous Sources and Related Ethical Concerns in Journalism: A Comparison of the Effects of the Janet Cooke/" Washington Post" Incident on the Policies and Practices of Large Newspapers and Television Stations.".