

The Spotlight's Harsh Glare: Rethinking Publicity and International Order

Allison Carnegie and Austin M. Carson*

October 28, 2014

*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, Georgia State University, Atlanta, GA (Email: amcarson@princeton.edu). The authors thank Peter Aronow, Cameron Ballard-Rosa, Camille Strauss-Kahn, Alexander Thompson, Jane Vaynman, and Keren Yarhi-Milo for helpful comments. All remaining errors are our own.

The Spotlight's Harsh Glare: Rethinking Publicity and International Order

Abstract

How does the publicity of states' dangerous or illicit activities affect the stability of international order? What does this tell us about when enforcers of international rules decide to publicize these states' violations? In contrast to the conventional wisdom that transparency strengthens efforts to build and enforce the normative-legal order, this paper argues that transparency often undermines adherence to international norms and laws. Specifically, we develop two ways by which exposing rule violations can endanger norms and formalized legal regimes: by correcting beneficial misperceptions about the overall level of non-compliance in the international community, and by sharpening the threat to others posed by deviance. These considerations thereby lead enforcers of international rules to selectively publicize transgressions. Focusing on the domain of nuclear non-proliferation, we demonstrate that concerns over the health of the non-proliferation regime factored heavily into the United States' calculations about whether to reveal other states' efforts to obtain nuclear weapons. We first formalize this argument and then empirically test the implications derived from the model using in-depth case study analyses in which we rely in part on recently declassified U.S. intelligence assessments and policy discussions. We conclude that enforcers fail to disclose infractions when this publicity would be unsuccessful in deterring the violator's actions. In such a case, bringing this conduct to light would threaten international order by undermining the established rules and by leading other states to engage in deviant behavior in response.

Keywords: international institutions, international norms, nuclear weapons, publicity, transparency.

“PUBLICITY IS JUSTLY COMMENDED AS A REMEDY FOR SOCIAL AND INDUSTRIAL DISEASES. SUNLIGHT IS SAID TO BE THE BEST OF DISINFECTANTS; ELECTRIC LIGHT THE MOST EFFICIENT POLICEMAN.”

Louis Brandeis

How does the publicity of rule violations affect international political order? The problem of developing rules to regulate social order is one common to virtually all communities. Brandeis’s aphorism about the sanitizing power of sunlight reflects the prevailing view that exposing rule violations helps reinforce such rules and foster social order. The spotlight regulates social behavior through exposure: actual or threatened negative publicity can lead individuals, corporations, and states to comply with rules. Not surprisingly, given the scarcity of centralized coercion under anarchy, many scholars have invoked this mechanism to help explain how international political order is developed and maintained. Constructivist and institutionalist theories alike point to the importance of identifying and publicizing rule violations to build normative support and ensure compliance. These theories have strong micro-foundations in the broader fields of political science; studies of voter turnout and human rights, for example, have shown that naming-and-shaming can lead to compliance with political rules.¹

Despite the apparent consensus regarding publicity’s positive role as a mechanism for strengthening social and political order, several strands of seemingly anomalous findings in sociology and criminology suggest this association may not always hold. Decades ago, influential social theorists, including Georg Simmel and Erving Goffman, hinted at this possibility when noting how the *unobservability* of deviant behavior is part of what creates consensus and strong normative orders (Goffman, 2009; Simmel, 1950). Indeed, the order-enhancing effect of ignorance can “protect the traditional normative structure” by “reinforcing the assumption that deviation from the rules is statistically insignificant,” so that under certain conditions, “the normative system as such may suffer more from knowledge of violations than from the violations themselves” (Moore and Tu-

¹See, for example, Gerber, Green and Larimer (2010) for voter turnout and Franklin (2008); Murdie and Davis (2012) for human rights, among others. But see Hafner-Burton (2008) for an interesting argument that shaming in one area of human rights can lead states to violate human rights in another area.

min, 1949, 791). More recently, studies of rule compliance in everything from college alcohol consumption to tax reporting have found that information about the behavior of others can *reduce* rates of compliance. This raises a puzzling question: When does naming-and-shaming preserve norms and strengthen international regimes, as suggested by the conventional wisdom, and when might it actually undermine them?

To resolve this issue, this article develops and tests a novel theory that reconciles these diverging accounts. We argue that while the conventional wisdom often holds, there are important circumstances under which it breaks down. Moreover, we claim that the conditions under which the conventional wisdom fails are especially prevalent in the international setting. Therefore applying the findings of these micro-level studies to the international arena can have especially perverse consequences. In particular, we develop two ways by which exposing rule violations can endanger norms and formalized legal regimes: by correcting beneficial misperceptions about the overall level of non-compliance in the international community, and by sharpening the threat to others posed by deviance. Both mechanisms can prompt order-degrading reactions by other states. We develop a model of a particularly interesting observable implication: when these mechanisms are operative in the international politics setting, regime advocates should intentionally conceal and overlook evidence of rule violations. The theory therefore expects states to sometimes reject the normal benefits of the spotlight and engage in what we call “strategic obfuscation”—that is, they abstain from drawing attention to violations and act to maintain informational and legal ambiguity.

We evaluate the theory empirically by analyzing the politics of publicity surrounding the domain of nuclear non-proliferation. Our focus is on an especially “hard test” of the theory: whether decision-makers in the regime’s strongest advocate, the United States, intuitively grasp publicity’s dangers and act preventively to avoid exposure of other states’ nuclear weapons programs.² To shed light on how U.S. leaders viewed publicity regarding proliferation activities, we draw on unusually candid internal documents made available via declassification in the last ten years.

²While other states also help to enforce the regime (Coe and Vaynman, 2014), we focus on the U.S.’s enforcement efforts since the U.S. is the most powerful state and represents the primary enforcer.

Our primary finding is that Washington frequently worried it would exacerbate global pessimism and threat perceptions, thereby endangering the regime. Consistent with our theory, U.S. leaders engaged in strategic obfuscation when conditions favored the negative publicity dynamics we identify. When conditions shifted, easing the dangers of pessimism and threat, U.S. leaders embraced publicity and shaming in ways that mirror the conventional wisdom. We show this variation through a variety of case studies of U.S. reactions to nuclear proliferation. The nuclear non-proliferation context therefore demonstrates both the decidedly ambivalent posture states may take towards publicity and the complex and contingent relationship between publicity and international order.

The article makes three main contributions: First, it makes a strong case for revisiting a core assumption that appears in both theories of international institutions and the development of laws and norms. Norm scholarship often cites the importance of naming-and-shaming tactics in forcing holdouts to join a given regime and in reversing violations. Studies of treaty compliance have similarly emphasized the importance of making rule violations visible so that such deviance will threaten reputations, prompt retaliation, and trigger pressure in favor of compliance. Our theory suggests publicity and international order are connected in more complex ways than are suggested by this standard account.

Second, the article makes sense of puzzling cases in which regime advocates uncover yet ignore rule violations. A basic finding from the archival records we review is that leaders do not always rush to the press room when they detect deviance. In the nuclear setting in particular, regime advocates may balk at exposing nuclear programs. Publicity is thus often a political choice accompanied by a host of risks and benefits. Our theory takes an important first step in theorizing some of the factors that influence this decision to explain why states sometimes elect to respond to violations through strategic obfuscation.

Third, the article's findings contribute to our understanding of the implications of what many believe is a new era of unprecedented visibility. Social media, leaks, and democratic political institutions render state behavior more transparent than ever before. While the "sunlight as social

disinfectant” view suggests such transparency helps advance norms and rules and enforces compliance, our findings offer a more cautionary view. Removing states’ abilities to occasionally conceal and ignore rule violations may make the international normative-legal order more fragile and lead to more frequent disruptive crises. As both scholars and policy-makers consider how statecraft evolves in a relatively transparent environment, it is all the more important to carefully analyze the potentially complex relationship between publicity-related processes and international order.

The article proceeds as follows. First, we review extant research on publicity’s role in building and enforcing international political order. Next, we discuss the theoretical underpinnings of our argument by incorporating insights from research outside of political science, focusing on two distinct mechanisms through which publicity can harm informal and formal institutions. We then evaluate the plausibility of these norm-endangering mechanisms through the development of a simple formal model, and test the model’s predictions through a case study analysis. We conclude by offering thoughts on the implications of our findings and identify fruitful directions for future research.

The Spotlight’s Benefits: International Order through Publicity

A common thread in studies of both norms and international institutions is that publicity plays a productive role in helping maintain a rule-based international order. While we agree that publicity *can* work in the manner commonly supposed, we also identify conditions under which it has negative effects on international regimes. We therefore begin by reviewing the mechanisms through which publicity can promote norms and laws before discussing when this conventional wisdom is unlikely to hold. Note that, because the international system is fundamentally anarchic, international norms are generally seen as relatively synonymous with rules and laws (Finnemore and Sikkink, 1998). Since publicity is commonly assumed to impact each of these concepts in a similarly positive manner, we use these terms interchangeably.

Norms—standards of appropriate behavior for actors with a given identity—are generally ac-

cepted as a crucial component of international relations. This concept can be supported by liberalism, constructivism, and realism, though these theories may differ on the particular roles that norms play in international relations (Krasner, 1983). Indeed, scholars studying the rise and consolidation of international norms and rules, often referred to as norm lifecycles, emphasize the helpful role of publicity in propelling the emergence and consolidation of new norms. Models of norm development highlight the roles of various actors (norm entrepreneurs, transnational civil society groups, international organizations) and techniques (framing, naming-and-shaming) (Finnemore and Sikkink, 1998; Keck and Sikkink, 1998; Risse-Kappen, Ropp and Sikkink, 1999) to determine changes in norms over time, the impact of various international forums, the complex relationship between legalization and norm strength, and the potential premature cessation of norm cascades.³ These models take an almost uniformly positive view of publicity, arguing that drawing attention to norm-inconsistent behavior can stimulate domestic and international pressure and force deviant states to enter or rejoin normative regimes (Finnemore and Sikkink, 1998; Keck and Sikkink, 1998; Risse-Kappen, Ropp and Sikkink, 1999; Sikkink, 2002). Turning norm violations into political and media spectacles, it is argued, may also help reassure states in good standing that the community remains committed to upholding these principles. Applied studies of norms regulating human rights (Risse-Kappen, Ropp and Sikkink, 1999), the use of land mines (Price, 1998), and gender roles (Zwingel, 2012) have found support for these claims.

Studies of compliance with treaties and international law take a similarly positive view of publicity. Exposing violations of agreements and international laws is an important mechanism for compliance. Compliance based on state reciprocity and reputational considerations—which often play a large role in international settings (Yarhi-Milo and Weisiger, 2014)—rely on member-states’ confidence that non-compliance will be exposed and well-known (Keohane, 1984; Milgrom, North et al., 1990; Keohane, 1997). Publicity can thus serve as a deterrent for future violations (Miller, 2014). If compliance can be supported by domestic interest group pressure (Dai, 2006, 2005),

³For accounts of norm change after institutionalization see Sandholtz (2008); Krook and True (2012); for adaptation to local contexts see Zwingel (2012); and for forum effects see Coleman (2013).

state behavior must be visible to compliance advocates. Institutional design can influence the visibility of regulated behavior as well, such that more transparent structures can produce better compliance rates (Mitchell, 1998).⁴ Scholars of norms and international regimes have therefore echoed Brandeis's aphorism that sunlight serves as a social disinfectant, improving compliance by enabling decentralized punishment and energizing domestic and other pro-compliance pressure groups.

Publicity's Dangers

The prevailing view of the effects of publicizing rule violations on international order is intuitive but incomplete. Under common conditions in international politics, we argue that the publicity of rule violations can endanger rather than bolster norms and formal legal regimes. Our theory builds on several studies in criminology, economics, and sociology, which have shown that sometimes the spotlight poses a threat to the rules and conventions that govern many social environments. These studies are based on observation that one source of strength for many social rules is the *non-observability* of rule violations and community members' tendency to ignore even known violations. Influential social theorists, including Georg Simmel and Erving Goffman, note this reportedly. Regarding taboo behavior, Simmel argues that social order can be bolstered by "ignorance and a measure of mutual concealment" (Simmel, 1950, 315). Similarly, Goffman argues that the tendency for individuals to protect their social image by hiding deviant behavior creates a misleadingly strong appearance of social conformity, which, in turn, reinforces existing social rules (Goffman, 2009). These theoretical appraisals suggest a more complex relationship between deviance, publicity, and social order than that captured by Brandeis's sunshine metaphor.

In adapting this basic insight to international politics, we isolate two primary mechanisms through which publicity of rule violations can endanger social order: by revealing that fewer states

⁴The managerial school sees willful state noncompliance as less of a problem. Yet this approach also points to transparency as a key way to strengthen norms and ensure treaty compliance by coordinating, reassuring, and deterring members (Abram and Chayes, 1995).

follow the norm than previously realized, and by creating incentives for other states to violate the norm in self-defense. First, regular publicity of rule violations can prompt other states to revise previously held beliefs about the overall rate of compliance with the norm. In diverse settings from juvenile gangs to tax compliance, scholars have found that hard-to-observe violations lead community members to systematically overestimate the overall levels of agreement and compliance with the rules. This overestimation, in turn, bolsters the social and normative forces for conformity.⁵ In contrast, making violations easily observable can correct these overestimations and bring about a sense of pessimism that may weaken formal and informal pressure to conform. Interestingly, some studies have experimentally investigated the effects of the provision of accurate data on rates of community compliance. Compared to a control group that receives an uninformative manipulation, compliance with informal norms precipitously declines when new information about actual rates prompts correction of prior overestimates (Hill, Maruyama and Viceisza, 2012). Adapting this basic mechanism to world politics suggests publicity of deviant state activity may, under certain conditions, lead states to revise their beliefs about the overall rate of compliance in the international system. It is certainly plausible that states hide violations of international laws and norms just as other social actors do; if so, compliance over-reporting and overestimates of the health of regimes should be common. Publicity of new rule violations can therefore trigger new pessimism about overall rates of compliance and weaken the formal and informal pressures for states to conform.

In addition to correcting beneficial misperceptions about levels of compliance, publicity may weaken normative regimes by sharpening or even exaggerating the threat such violations pose to other states. If states tend to react to such a threat by themselves violating rules, a cascade of defections may result. This scenario is most plausible when regimes govern state activity with a high degree of interdependence, meaning one state's rule violation creates large negative externalities for other states.⁶ Negative externalities may cause one state's violation to prompt reactions in

⁵Useful overviews of this research include Prentice and Miller (1993); Bicchieri and Fukui (1999); Kitts (2003); Groeber and Rauhut (2010).

⁶A similar discussion of this concept, in the language of cooperative state behavior based on reciprocity, is found in Simmons (2010).

self-defense by others, risking a tit-for-tat or spiral process (Jervis, 1976) that further weakens the regime. Legro (1995)'s study of norms governing submarine warfare, aerial bombing, and chemical weapons in World War II demonstrates this risk. He finds that clear violations of these rules by one state created security externalities for others and thus tempted them to respond with additional violations, leading in some cases to collective abandonment.⁷ Spotlighting deviance eliminates ignorance and uncertainty that might otherwise lessen the risk of second-order violations by other states.⁸

Thus, publicizing rule violations may alter perceptions of global compliance rates and magnify the direct threat from security, economic, or environmental externalities. In what follows, we refer to the former as the *pessimism mechanism* and the latter as the *threat mechanism*.

The Nuclear Non-Proliferation Regime

To make our problem tractable, we focus on a particular international norm: the prohibition on the use of nuclear technology for the purpose of developing new nuclear weapons arsenals.⁹ The norm's origins lie in the anti-nuclear weapons advocacy of arms control of the 1950s (i.e. the test ban movements), and the norm was formalized in 1968 in the Non-Proliferation Treaty (NPT), a binding multilateral agreement (Gavin, 2006). The regime includes a formal organization, the International Atomic Energy Agency, which is tasked with providing technical assistance to non-nuclear states and verifying compliance. Since the NPT only regulates military uses of nuclear

⁷Such an outcome reflects Kahler (2000, 679)'s caution regarding international law that "cascades do not run in only one direction, as the uncertain history of both free trade and democracy in this century demonstrate."

⁸It is worth noting that domestic politics can exacerbate both mechanisms. Regarding the pessimism mechanism, publicity can lead watchful domestic constituents to revise their views of the overall rate of compliance and make violations salient to less attentive audiences. Publicity may also sharpen the perceived threat to domestic audiences. Spotlighting the rule violation of a neighbor or rival can mobilize and deepen nationalist sentiment making leaders of responding states more likely to act harshly, including by themselves defecting from the regime or abandoning the norm. The feedback loop that can lead to abandonment can be exacerbated by confusion about whether reactions to violations are themselves deviant behavior or justified punishment. See the description of uncertainty regarding responses and the creation of a "sanctioner's dilemma" in Thompson (2009).

⁹The norm is related to, but distinct from, the taboo against use of nuclear weapons in war (Tannenwald, 2009, 2007).

technology, a rule violation here is defined as pursuit of nuclear facilities and nuclear weapons-related research and development by any state besides the five classified as nuclear weapons states in the NPT.¹⁰

As a result of the treaty's importance and salience, it has attracted considerable scholarly attention (Verdier, 2008), as have decisions to violate the treaty through nuclear possession (Rublee, 2009; Lantis, 2011). For instance, scholars have presented interesting accounts of the norm's impact on specific proliferation decisions (Rublee, 2009), changes to the norm in recent years (Lantis, 2011), and the importance of nuclear weapons for pariah status in the modern system (Litwak, 2000). However, to our knowledge, we are the first to investigate the determinants of the decision to publicize violations of this norm.

The nuclear proliferation setting provides an ideal area in which to test our claims for several reasons. First, it features powerful states—the United States in particular—who sought to create and enforce a prominent and substantively important international norm. As we explain in the following section, we focus on how the theory implicates regime advocates and their reactions to detected rule violations. The non-proliferation regime is therefore an excellent setting in which the logic of our theory should apply across many countries over a long period of time. Second, the regime hosts variation in our critical independent variables across region and over time; this permits us to assess publicity and obfuscation decisions in a range of settings. Third, while it is often very difficult to identify acts of state concealment and their motivations, the non-proliferation setting features unusually candid U.S. archival material recently declassified and made publicly available. We are able to use these records to assess our claims with an uncommon level of precision.

A Simple Model of Strategic Obfuscation

We focus on an important implication of our theory: how states who support an international regime react to detected violations. This section develops a basic formal model as a heuristic

¹⁰These states include the United States, Russia, the United Kingdom, France, and China.

device to explain the conditions under which publicizing rule violations can undermine normative/legal regimes, and then uses this model to predict when regime advocates should publicize or strategically obfuscate these violations. Existing theories view publicity as either beneficial or, at worst, neutral, thus implying that such states should publicize all violations. Our model incorporates insights from the sociology and criminology literatures highlighted above to shed light on when states might choose to avoid drawing attention to rule transgressions. In so doing, we are able to generate testable empirical hypotheses, which guide the case study analysis that follows. We provide a summary of the model here and relegate the formalization to the appendix.

Actors, Utility, and Timing

The model features three states: a norm enforcer (E), state A , and state B . Prior to the start of the game, neither A nor B have violated the norm of nuclear non-proliferation. The game begins with A choosing whether to violate the norm. Because we focus on whether such a violation is publicized, A 's decision remains exogenous, such that A violates the norm with probability q . While B does not observe A 's decision, the norm enforcer does, and, if A violated the norm, it decides whether to publicize this violation.¹¹ State B does not observe A 's choice, but observes E 's decision. B forms beliefs about whether A violated the norm, which are determined by Bayes' Rule, as is standard in models of incomplete information. If E publicized the norm, A then decides whether to reverse its action in response by coming into compliance with the NPT. The proportion of A s that do not come into compliance is given by r and is also exogenously determined. B then chooses whether to violate the norm.

The states' utility functions depend on whether other states comply with the norm, along with other domestic and international considerations. Consider the incentives of the two strategic actors in the game, B and E . B derives utility from several sources. First, the benefits B receives from following the norm depend on the number of other states that follow the norm. The more states that

¹¹We assume that E never publicizes a violation if a violation did not occur. Since E must present evidence of a violation, we think it is reasonable to assume that A cannot lie, though this is easily relaxed.

comply, the more B gains from its own compliance. This specification is in line with the discussion above, as the idea that an actor benefits more from following a norm when others also follow it is central to this literature.¹² In addition, since self-defense is a frequently cited justification for states' attempts to obtain nuclear weapons, B also gains from its additional ability to defend itself if it possesses nuclear capabilities. The degree to which it benefits depends on its threat environment, i.e. whether A has nuclear weapons. Finally, we represent the myriad of other potential costs and benefits associated with acquiring nuclear weapons as the parameter c , which can be positive or negative. This could include factors such as prestige, monetary expenditure, dangers of loose nuclear weapons, and other concerns.

Similarly, E receives more utility the more states follow the norm. E cares about compliance with the norm almost by definition, as its role is to enforce the norm.¹³ Like B , E 's utility also depends on the potential domestic and international costs and benefits it may receive, though we leave this in the background to minimize notation. Further, strategic obfuscation is costly, since it requires effort to keep nuclear proliferation a secret. However, it can also provide benefits domestically and internationally, as discussed further below. E thus faces both costs and benefits associated with hiding A 's proliferation efforts.

Solving the Model

To understand how the game is played, consider two cases: first, when A is the type that reverses course once its violation is publicized, and second, when A is the type that does not reverse course when its violation is publicized.¹⁴ In the first case, E publicizes A 's violation, knowing that doing so will lead A to end its program and come into compliance. B finds out that a violation has occurred, but observes A come into compliance, and therefore does not need to acquire nuclear

¹²This is similar to a coordination game, where actors benefit from making mutually consistent decisions.

¹³To clarify, the general decision by states to enforce norms has been analyzed elsewhere (see, for example, Bernhard, Fehr and Fischbacher (2006), among many others); here we focus on a specific kind of state behavior in light of the choice to support a regime.

¹⁴For the purposes of this simple model, we also only consider pure strategy equilibria and the solution concept is perfect Bayesian equilibrium.

weapons in self-defense. Further, since *A* comes into compliance, the norm is preserved, and *B* still follows the norm since everyone else is doing so. Moreover, even if *A* is unlikely to reverse course, publicity is a rational response for *E* if the likelihood of second-order proliferation is sufficiently low. Opting not to publicize *A*'s activity may be exposed later, risking international and domestic criticism that *E* was privileging other interests over the regime's integrity. Even if naming-and-shaming is fruitless, the absence of second-order proliferation risks make it net worth the effort. Publicizing the norm violation thus enhances the norm's health by highlighting *A*'s deviant behavior and pressuring *A* to cease its violation. This represents the standard logic through which advocates of publicity argue that this mechanism operates.

However, now suppose that once *E* publicizes *A*'s violation, *A* does not come into compliance. *A* could choose not to do so for a variety of reasons. Most obviously, the punishment inflicted by the publicity may not outweigh the gains from obtaining the weapons. Further, publicity could even have perverse effects, increasing political resistance to *E*'s requests due to a rally around the flag effect (Mack and Khan, 2000).¹⁵ As a result of *A*'s non-compliance, two dynamics come into play. First, consistent with our threat mechanism described above, *B* is more inclined to violate the norm in self-defense. Since *A* represents a threat, *B* feels a greater need to counter that threat, both from a national security standpoint and due to domestic pressure. Second, consistent with the pessimism mechanism, a blow is dealt to *B*'s view of the overall rate of compliance and strength of the norm. This makes *B* more inclined to violate the norm as well. The knowledge that a violation occurred can thus weaken *B*'s incentives to remain within the regime.

When *E* observes a violation, then, its strategy depends on the cost of strategic obfuscation and its desire to deter *B* from also violating the norm. Suppose *E*'s strategy is to publicize if *A*

¹⁵While we focus on *B*'s decision in this analysis, we note that there are many factors that enter into *A*'s choice that are not modeled here. For instance, a large literature argues that publicizing violations harms *A*, imposing domestic and international reputational costs. For leaders that wish to be seen as trusted members of the international community, or face domestic penalties for breaking international norms, publicity often causes them to alter their policies due to basic cost-benefit calculations (Kirshner, 1997). Indeed, a variety of factors have been shown to increase the degree to which *A* complies including an alliance between *E* and *A* (Lam, 1990; Bergeijk, 1989, 1994), when *A* is smaller and weaker relative to *E* (Bonetti, 1998; Lam, 1990), when the target is a democracy (Nooruddin, 2002) and the post-Cold War period (Dunning, 2004).

will comply and not to publicize otherwise. If E does not publicize, B cannot detect the violation; it only knows that no violation was publicized. It thus assumes that a violation took place with probability β , which is determined by Bayes' rule. Since B is unsure whether a violation took place, it may violate the norm itself. Yet since A 's violation remains hidden, B is left with the impression of a greater normative consensus than if B knew a violation occurred, creating stronger pressure to conform. Further, B 's environment is less threatening, since it is only *possible* that B faces a direct threat from A , rather than certain. The costs of acquiring nuclear weapons may not be worth addressing this potential threat, while they may be justified in the face of a known one. In other words, strategic obfuscation avoids triggering the pessimism and threat mechanisms that otherwise might endanger an international regime.

Table 1: Choice of Strategic Obfuscation versus Publication

	A Will Comply	A Will Not Comply
High Risk of Reactive Violations	Publicize	Strategic Obfuscation
Low Risk of Reactive Violations	Publicize	Publicize

In sum, when publicity does not deter A , it can undermine the norm when the norm is less solidified and incentives to proliferate in self-defense are high. In this case, E does not publicize norm violations when doing so will prevent B from violating the norm. This discussion is summarized in Table 1.¹⁶

Empirical Analysis

We now demonstrate the plausibility of our argument by analyzing specific instances of violations of the norm against developing new nuclear arsenals. Our purpose is not to conduct a definitive theory test, but rather evaluate the model's predictions and contrast them with the expectations of the conventional wisdom depending on the underlying parameters.

¹⁶Of course, in reality, E may not know with certainty whether A will comply, but the important feature is whether E expects A to comply once it publicizes A 's program.

While not exhaustive, we believe supportive findings in what follows are especially valuable because we are conducting a relatively "hard test" of the theory. Recall our empirical analysis focuses on a specific implication of our broader theory of publicity and political order: whether certain countries (i.e. *E*) consciously manipulate information to safeguard a regime in accordance with the logic of our two mechanisms. In so doing, we are testing not just whether publicity endangers order, but also *whether states intuitively grasp this and use deception to avoid it*. It is entirely plausible that hard-to-observe violations help regimes, as our broader theory argues, but that states do not consciously and preventively cultivate this. If we find evidence of manipulation by a regime advocate, we believe this is especially strong evidence that our basic logic of publicity's effects on political order is valid.

We focus analysis on how a leading regime defender, the United States, reacted to detected violations of the nuclear non-proliferation norm. Our theory expects the U.S., or *E*, to strategically obfuscate when two conditions hold: a low likelihood that the detected state (*A*) will reverse its proliferation activities and a high danger of other states (one or more *B*s) undertaking norm-violating nuclear programs in reaction to *A*'s violations. When these conditions do not hold, we expect the U.S. to publicize *A*'s activities in line with the conventional wisdom. Decisions made by the U.S. are particularly useful to analyze in this context because it has been a lead advocate of the non-proliferation norm and its sophisticated national intelligence means have often provided privileged knowledge of rule violations.¹⁷

From the overall universe of cases of states with military nuclear programs since 1945, our specific research question narrows the relevant universe to: (1) states with credible interest (intention and minimal nuclear facilities) in a military nuclear capability; (2) states with programs active after the emergence and formalization of a norm against proliferation; and, (3) states that have not overtly demonstrated a nuclear weapons capability.¹⁸ The set of proliferation cases that meet

¹⁷On the special importance of the U.S. to the nonproliferation norm, see Lantis (2011). We discuss the importance of *E*'s information privileges in the conclusion.

¹⁸We define the second criterion as state programs that were obtained prior to 1967 both for reasons of simplicity and because the non-proliferation norm's emergence coincided with its formalization in the negotiations and signature of the NPT. States that later overtly declared nuclear arsenals may meet the inclusion criteria earlier in the life of the

these criteria, when detected by the United States, lead to a strategic choice between spotlighting or obscuring rule violations. Table 2 lists the resulting universe of cases.¹⁹

For each state in our universe of cases, we code several different variables. We begin with our main independent variables: the perceived likelihood that the proliferating state would reverse its program following publicity (i.e. the probability that *A* would come into compliance) and the perceived risks of second-order proliferation due to *A*'s activity (i.e. the risks of reactive non-compliance by *B*). The latter independent variable is coded in two parts. First, we assess the danger that other states would react by pursuing nuclear weapons in self-defense. Second, we determine how strong the normative regime was at the time during which each case takes place. The less codified the norm is—i.e., the fewer states that had agreed to comply with the NPT—the greater the danger of a single defection. Thus, for our purposes, we code the normative regime into three phases: I, II, and III. Phase I consists of the 1970s, the first decade during which the NPT was in effect. This is the most fragile period of the regime, as only 43 states initially deposited their instruments of ratification, and 93 signed on. More states continued to accede to the regime through the 1970s, such that 114 acceded by the end of the decade (State Department, 1998). Phase II consists of the 1980s, a period during which the norm continued to codify as more states joined the regime. By the end of the decade, 135 states had acceded to the regime. However, by the end of this period, additional states had developed nuclear programs such that Israel, India, and South Africa had become nuclear powers. Therefore, the threat level to the norm of an additional defection was moderate during this period. Finally, Phase III consists of the 1990s and beyond. During this phase, the regime was the most stable due to the high number of signatories and to the

program, such as North Korea in the 1990s.

¹⁹Cases included in other studies of nuclear proliferation but which do not meet our specific criteria for inclusion are Australia (lack of credible interest; program prior to regime), China (program prior to regime; declared), France (program prior to regime; declared), India post-1998 (declared), North Korea post-2011 (declared), Pakistan post-1998 (declared), Romania (publicly declared intent), Sweden (program prior to regime), Switzerland (program prior to regime), the United States (program prior to regime; declared), Ukraine (declared), the USSR (program prior to regime; declared). We exclude Iran because of its status as an open case but reflect on it more speculatively in the conclusion. Finally, we exclude Yugoslavia due to missing data; since key U.S. documents have not yet been declassified and due to the paucity of publicly available information, we are unable to code our main independent variables—the threat level perceived by the U.S. and whether the U.S. believed Yugoslavia would comply as a result of publicity.

end of the Cold War, which left the U.S. as the sole superpower, enhancing its role as the enforcer of the regime. By the end of the 1990s, 186 states had acceded to the NPT (State Department, 1998). Finally, for our dependent variable—*E*'s reaction to *A*'s violation—we code each case for whether the U.S. chose a strategy of publicity or strategic obfuscation.²⁰

Reliably coding these values requires unusually candid access to internal intelligence assessments and discussions. Our research design therefore takes advantage of the release of a trove of declassified material on U.S. assessments and deliberations about problematic nuclear aspirants.²¹ In most cases, these documents permit a truly exceptional level of insight into U.S. knowledge and strategic choices regarding proliferation. We supplement these data with additional secondary sources where necessary.

Table 2: Choice of Strategic Obfuscation versus Publicity

	<i>A</i> is Expected to Comply	<i>A</i> is Not Expected to Comply
High Risk of Reactive Violations	North Korea Iraq After 1990 Taiwan South Korea	Israel India Pakistan Iraq Before 1990 South Africa
Low Risk of Reactive Violations	Algeria Libya	Brazil Argentina

Notes: The table summarizes our findings, demonstrating which cases featured conditions that should lead *E* to strategically obfuscate (upper right box) or publicize (other three boxes).

We organize our empirical analysis according to the four scenarios in Table 2. We focus the most attention on cases in the upper right box, where our theory expects *E* to adopt strategic obfuscation, contrasting with the conventional wisdom. For the other three scenarios in which our

²⁰We code *E*'s reaction as "publicity" when it undertakes an organized effort to draw attention to and share private information regarding proliferation activity of another state. *E* may spotlight *A*'s behavior through both bilateral contacts with other governments and use of public media. We code *E*'s reaction as "strategic obfuscation" when it engages in an organized effort to minimize scrutiny of *A* and abstains from sharing private information regarding its activities. The latter creates an identifiable disjuncture between *E*'s internal perception of *A*'s potential for proliferation and *E*'s public rhetoric. We should note that strategic obfuscation is not mutually exclusive with many other policy tools, such as direct and private pressure on *A*.

²¹Especially valuable collections are with the National Security Archive, George Washington University, and the Woodrow Wilson Center's Nuclear Proliferation International History Project.

theory predicts *E* to publicize, we feature one exemplar for each in the main text and analyze the remaining cases in the supplemental appendix. We now turn to a detailed examination of cases from each box in turn.

Proliferator Unlikely to Comply and High Risk of Reactive Proliferation

When the threat of proliferation in reaction to *A*'s violation is high and *A* is unlikely to come into compliance in response to publicity and pressure, our theory expects *E* to strategically obfuscate *A*'s nuclear activity. The intuitive logic, based on both the threat and pessimism mechanisms, is that U.S. publicity would sharpen the threat to regional rivals and weaken confidence in non-compliance among the broader community of potential proliferating states. Simultaneously, because the proliferator was unlikely to reverse course, U.S. naming-and-shaming was unlikely to quell these concerns. We find that these conditions held in five important cases and, consistent with the theory, each featured U.S. efforts to strategically obfuscate.

Israel

The U.S.'s reaction to Israel's nuclear program in the late 1960s is a case of strategic obfuscation that provides strong support for the theory. U.S. leaders believed a strategy of publicly criticizing Israel's apparent weapons research would be very unlikely to prompt Israeli compliance with the nascent nuclear non-proliferation norm. U.S. leaders were also acutely concerned about second-order proliferation, both in Israel's neighborhood and further afield.

U.S. confidence that Israel was clandestinely progressing towards a nuclear arsenal grew in a political climate marked by strong concerns about larger proliferation trends. French (1960) and Chinese (1964) entry into the nuclear ranks by the mid-1960s prompted widespread fears that the proliferation of nuclear arsenals might spiral out of control.²² The Soviet Union had similar

²²President John Kennedy famously worried, "I see the possibility in the 1970s of the President of the United States having to face a world in which 15 or 20 or 25 nations may give these weapons." See transcript of a March 1963 news conference at <http://www.jfklibrary.org/Research/Research-Aids/Ready-Reference/Press-Conferences/News-Conference-52.aspx>. For recent reviews of this period, see Gavin (2006); Brands (2006).

concerns and, by the late 1960s, was actively helping negotiate what would become the NPT.²³ This period thus marked the earliest stages of the NPT—the most fragile period for the norm’s health. The number of states adhering to the norm was therefore small, such that a deviation could easily endanger the norm.

Israel’s nuclear program posed a potentially alarming problem in this context. By 1961, the U.S. had concluded that Israel would possess plutonium by 1965-6 and would have a weapon by 1966-7, and by 1969, the U.S. concluded that Israel possessed a weapon (Montgomery and Mount, 2014). Thus, just as the NPT was being negotiated and signatures gathered for entry into force, U.S. intelligence concluded the Israeli nuclear program was nearing completion of a weapons capability (Cohen, 1998, 297-99). U.S. concerns about Israel’s program were therefore as much about the program’s impact on global diplomacy as its consequences for Middle East politics. Based in part on fears of the former, and convinced that Israel could not be pressured to relinquish its program, Israeli and U.S. leaders settled on a quiet, improvised compromise: a joint commitment to conceal and avoid public acknowledgement of any Israeli arsenal or, in other words, strategic obfuscation (Cohen, 2013).²⁴

A September 1969 visit by Israeli Prime Minister Golda Meir prompted discussions within the Nixon Administration about the nuclear program and culminated in an agreement by both sides to avoid publicizing Israel’s status.²⁵ Evidence for both the pessimism and threat mechanisms in our theory is showcased during these debates. Obfuscation advocates argued that avoiding publicity of Israel’s nuclear arsenal would help reduce the chances that other states threatened by the arsenal would reject the NPT themselves and seek their own nuclear capabilities. As Cohen (2013, 50) notes, “The most serious concern in the 1960s was that the Israeli nuclear project would lead to a dangerous regional nuclear arms race.” In the short-term, U.S. leaders feared Israel’s rivals in Syria and Egypt might seek a Soviet nuclear security guarantee. U.S. National Security

²³On shared Soviet-U.S. interest in the NPT, see Gavin (2006, 112).

²⁴The interest by the two sides was not for identical reasons. While Washington sought to avoid a regional arms race and a body-blow to the NPT, Israel’s interests included other factors reviewed in Cohen (2013).

²⁵Cohen (2013, 2) refers to this meeting as the “birthplace of the bargain.”

Advisor Henry Kissinger's top advisor on the Middle East, Joseph Sisco, wrote, "Until the Arabs could develop nuclear weapons, they might seek, and get, Soviet agreement to extend a 'nuclear umbrella' to the Arab states" (Sisco, 1969). Publicity would make this worse in part because the Soviet Union would likely find it harder to decline such a request once Israel's arsenal was a known and established fact. Over time, Arab states would likely seek an independent nuclear capability. The importance of publicity is specifically mentioned during a January 1969 National Security Council meeting. After President Nixon noted that the Israeli's were "very close to a weapon," General Andrew J. Goodpaster remarked, "The main prohibition against an Israeli announcement or admission that they have nuclear weapons would be the threat that the Soviets or Chinese would help the Arabs" (NSC, 1969).

The reaction of states further afield, that were not directly threatened by an Israeli arsenal, was also a concern. U.S. leaders argued that if Israel's nuclear progress were splashed across international newspapers during deliberations over joining and complying with the NPT, many states may have resisted. An issue paper widely circulated within the Nixon Administration on progress in NPT diplomacy identified key events which might have jeopardized momentum towards global acceptance of the treaty. Israel's nuclear arsenal and the need to privately lobby Israel from creating a diplomatic crisis were top considerations (i.e. "[s]elective diplomatic pressure to avoid actions (e.g., by Israel) that would prejudice attainment of the objectives of the treaty, or the treaty itself" (Kissinger, 1969c)). A State Department Policy Planning staff memo argued "[a] known Israel nuclear capability would have far-reaching unfavorable effects," first among them "[o]ther nuclear capable countries would be more likely to opt in favor of nuclear weapons for themselves and, even if they did not decide to produce weapons immediately, would be less likely to sign the NPT" (Owen, 1969). Internal documents show that leaders outside of the Middle East were tracking Israel specifically during NPT signature and accession debates. For example, Gavin (2006, 127) describes a 1966 meeting between the West German chancellor and Kissinger in which the German leader explicitly cites the possibility of proliferation in India or Israel as constituting a key consideration in German thinking about U.S. protection or its own nuclear option. The U.S.

intelligence community echoed policymakers' concerns, arguing that the non-proliferation norm's fate would be influenced by general perceptions of proliferation trends.²⁶

Given the dim prospects for coercing Israel to relinquish its capability, the U.S. pursued strategic obfuscation in which it created "circumstances in which Israel would not 'announce' a nuclear capability and would maintain secrecy" ensured through "private, bilateral assurances that Israel would not deploy or test nuclear explosive devices" (Davies, 1969). As Henry Kissinger succinctly stated in a memo to Nixon in the summer of 1969, "public knowledge is almost as dangerous as possession itself" (Kissinger, 1969*b*).²⁷ A later memo from Kissinger reminded President Nixon that during the Meir visit he had "emphasized that our primary concern was that the Israelis make no visible introduction of nuclear weapons or undertake a nuclear test program" (Kissinger, 1969*a*).²⁸

U.S. leaders believed obfuscation could meaningfully affect the potentially negative reactions of other states. Although many already suspected Israel's intent to eventually acquire a weapon—Egypt's Nasser had publicly warned the Arab world about a possible future Israeli arsenal in 1960—U.S. leaders believed obfuscation could help "keep Israeli possession from becoming an established international fact" (Kissinger, 1969*b*). Obfuscation would keep knowledge of Israel's progress to a minimum and news of it to a trickle rather than a firehose.²⁹ Indeed, although details about Israel's nuclear weapons program appeared in leaked form in the international press as early as 1970 (Smith, 1970; Cohen, 2013), and many states believed Israel possessed an operational arsenal by the mid-1970s, Israel and the U.S. never acknowledged its existence.³⁰ A leading scholar

²⁶See references to "global permissiveness" and the "general course of proliferation" in the analysis of considerations driving proliferation in several states in National Security Archive (1974).

²⁷The decision in 1969 made permanent the secrecy regarding U.S. estimates about the Israeli nuclear program, and Israel's own secrecy, that was in operation as early as 1960 (Montgomery and Mount, 2014, 14).

²⁸See also the discussion of the same memorandum in Cohen (2013, 26-7).

²⁹Internal debates acknowledged that "[n]ews about Israeli progress could continue to seep out, as it has already begun to do, until it is generally taken for granted that Israel has this capability" (Owen, 1969). The general goal of the NPT and the U.S. policy at that time was, however, to preserve momentum in favor of the norm against acquisition in order to at least slow the rate at which this occurred. Gavin (2006, 112) quotes internal U.S. documents stating that the NPT "'would keep [the] FRG in line for an extra 5-10 years (and perhaps longer).' If Israel remained nonnuclear, 'other nations might be kept in line,' and nuclear weapons acquisition might be delayed 'at least 5-10 years.'"

³⁰Scholars have noted that Israel's opacity policy "could only be maintained with the tacit acceptance of the inter-

of Israel's program deemed that the nuclear arsenal would become the international community's "worst-kept secret" (Cohen, 2013).

India

During the same time period, U.S. leaders also feared the consequences of India's progress towards acquisition of a nuclear weapon. These concerns came to a head in 1974 when India successfully conducted a specific kind of nuclear test—a peaceful nuclear explosion (PNE)—which India claimed was intended for civilian rather than war-related purposes. The PNE prompted a major crisis for the non-proliferation regime. The central question in the wake of the test was whether the event was merely a prelude to a declared and demonstrated Indian nuclear arsenal. Consistent with the theory, U.S. leaders, citing fears of reactive proliferation and the creation of pessimism about the regime's health, rejected public censuring of India and instead embraced strategic obfuscation.

The 1974 Indian PNE was a genuinely ambiguous event in normative and legal terms. As part of its provisions encouraging civilian uses of nuclear technology, Article V of the NPT allows low-yield nuclear explosions for civil engineering purposes, and India thus consistently framed its nuclear explosion as a peaceful one (Perkovich, 1999). India had not signed the NPT, though its actions were uniformly treated as relevant to the fate of the regime it had refused to join.

One important reason U.S. leaders opted against a campaign of naming-and-shaming India for its PNE was pessimism about the prospects of India relinquishing its program, particularly because of both its utility in India's conflict with Pakistan and its deterrent effect against China (US Embassy in Seoul, 1982). Before the PNE, U.S. policy-makers doubted their efforts would dissuade India, stating that the U.S.'s current strategies of deterrence were "likely to secure such a delay for only a relatively limited period" (Department of State, 1966, 10). Other policy officials expressed their beliefs that "U.S. efforts would probably have little value in influencing a final decision [about whether to go nuclear]" (Department of State, 1967, 1). Further, U.S. leaders had privately warned India that explosion of "any nuclear device, including one labeled a peaceful

national community." See Cohen and Miller (1990, 65), Cohen (1998, 284), and Cohen (2013, 241).

nuclear explosion was tantamount to the development of a nuclear weapon” (State Department, 1974). Given that New Delhi had ignored such warnings, U.S. leaders had good reason to believe that publicity would not force India to yield.

After the PNE, senior U.S. leaders therefore expressed their dismay to Prime Minister Indira Gandhi, but abstained from public criticism and sanctions.³¹ The State Department issued instructions to the embassy in New Delhi to respond to media inquiries with a “low-key” tone (Department of State, 1974, 1). U.S. diplomats reminded India in subsequent discussions that despite “sharp Congressional pressure,” the U.S. “did not publicly criticize India over its nuclear explosion last May. Indeed, [the] Secretary ‘welcomed’ [India’s] policy statement that India intended not [to] develop nuclear weapons when he visited Delhi” (State Department, 1975, 1). U.S. leaders also allowed new economic aid and debt forgiveness to proceed following the test (Perkovich, 1999). Thus, despite privately believing that India had an active military research program and would soon seek to create an arsenal, U.S. leaders opted against refuting India’s characterization of its nuclear program as peaceful.³²

This strategic obfuscation was in part driven by fears of second-order proliferation from dynamics embodying both the pessimism and threat mechanisms. Internal assessments of the test argued, “The implications could be considerable, both with regard to South Asia and in the broad nonproliferation context. The challenge is...trying to dissuade others from following suit” (State Department, 1974). Publicly rebuking India and providing evidence of its suspected military goals risked additional defections from the regime. As one history of South Asian nuclear dynamics in this period concludes,

³¹U.S. leaders chose to downplay the incident and India’s example publicly but, behind closed doors, launched a new initiative to restrict access to sensitive nuclear technology in general. As one retrospective intelligence report notes, “[i]n the immediate aftermath of that explosion, the judgment of nonproliferation supporters was that the reaction of supplier and other governments...was unfortunately mild in their failure to censure strongly India’s action or to impose sanctions. In actuality, however, the actions supplier governments have taken to prevent other governments from duplicating India’s course have been firmer, more comprehensive, and more successful than any rational observer might have expected...Supplier governments drew together—in the form of the London Suppliers Group—to limit substantially the dissemination of those items and knowledge that contribute most directly to making a nuclear explosive” (NSC, 1985, 10). See also (Perkovich, 1999, 187-90).

³²U.S. intelligence before and after the PNE “consistently predicted” that India would proceed with weaponization (NSC, 1985, 8).

Nuclear acknowledgement...by New Delhi and Islamabad would be likely to trigger an all-out nuclear competition...Perhaps equally important, nuclear acknowledgement would also considerably increase the pressures on neighbouring and other countries to acquire their own nuclear weapons...Open acknowledgement of nuclear weapon capabilities by India and Pakistan could well revive incentives within the Brazilian military, for reasons of prestige, to revive their now dormant nuclear weapon programme. More broadly, open acknowledgement would have a greater corrosive impact on overall global non-proliferation efforts, undermining perceptions that wide-spread proliferation can still be avoided and lessening countries' readiness to pay a price to support non-proliferation efforts" (Dunn, 1991).

Similarly, the U.S. worried that "the likelihood of further proliferation (e.g., Japan and Israel) would be increased, and nuclear pressures might be set in train in Germany" (Department of State, 1966, 1).

The U.S. feared that if other states learned of a violation of the normative regime by India, the regime could begin to break down, as "India was considered a potential 'nth country' whose decision to retain or reject non-nuclear status would send out a strong signal to the entire region and other aspiring nuclear states" (Schrafstetter, 2002, 88). "By breaking the nth country barrier India has made it difficult for [the U.S.] to argue the value of the NPT" (Department of State, 1974). Even as late as 1995, the U.S. feared a chain reaction starting in South Asia. "If India and Pakistan become declared nuclear-weapon states....some of their neighbors may reconsider their adherence to the NPT" (US Department of State, 1995, 1).³³ By portraying its test as part of the "peaceful" nuclear uses exception, and in part due to the relatively muted international response, the impact of the 1974 test on the wider behavior of potential proliferating states was blunted (Joeck, 1990; NSC, 1985).

³³For analysis of the India/Pakistan 1998 nuclear tests and subsequent blow to the non-proliferation regime, see Talbott (1999).

South Africa

South Africa began weapons research in 1974, considered testing in 1977, and may have conducted a joint test with Israel in the Atlantic Ocean in 1979. In 1977 the U.S. concluded that South Africa had enough highly enriched uranium for several devices, and in 1984 it determined that South Africa had enough of this uranium for 2-4 devices (Montgomery and Mount, 2014). The program was shrouded in secrecy until it was publicly disclosed in 1993 during the transition away from apartheid. At several key junctures, U.S. officials, debating whether to publicize or ignore South Africa-related events, expressed fears of second-order proliferation along with beliefs that publicity would not roll back South Africa's program.

Ultimately, U.S. officials believed that South Africa would not reverse its activities as a result of publicity, as the U.S. had "few means of pressuring the South Africans away from pursuing nuclear weapons" (Burr and Richelson, 2013). These officials cited South Africa's "unwillingness to bend to outside pressure" (CIA, 1984, 7). Indeed, intimidation "might provoke the South Africans to conduct a test as a warning that Pretoria will not be bullied" (CIA, 1984, 2).

Moreover, the U.S. worried that publicity would undermine the non-proliferation norm. For example, after a U.S. reconnaissance satellite detected a possible nuclear test in 1979, an interagency memo discussed the difficulties raised by the potential publicity of the event. The memo urged U.S. cabinet officials to consult with other governments. Given the likelihood that the event would become public, and as the perception of it as a South Africa test solidified, the author worried that such an outcome could threaten the norm, arguing that "[t]he nonproliferation stakes could be high" (NSC, 1979, 2). Similarly, U.S. policy officials argued, "Revelation that South Africa possessed nuclear weaponry would immediately exacerbate the tensions that exist in southern Africa" (CIA, 1984, 1). The introduction of nuclear weapons might "prompt neighboring black African states....to allow deployment on their territory of tactical nuclear weapons superior, no doubt, to any South Africa might have" (CIA, 1984, 15). It would further "weaken the international non-proliferation regime and encourage the acquisition of nuclear weapons by other countries" (CIA, 1984, 3).

This threat was particularly severe since it came early in the regime (during Phase I), such that a single violation could severely weaken the norm. Statements by U.S. officials reflected this worry. For instance, the CIA reported, “South African testing of a nuclear device would weaken the international nonproliferation regime and encourage the acquisition of nuclear weapons by other countries....Several states (such as Pakistan, Argentina, Brazil, and Israel) might feel fewer inhibitions about developing nuclear weapons or openly publicizing their nuclear weapons capabilities” (CIA, 1984, 18).

To prevent a weakening of the normative regime, the U.S. thus sought strategic obfuscation. Indeed, quiet pressure from the U.S. played a large role in South Africa’s decision not to visibly demonstrate its norm-violating program. A recently declassified interagency intelligence report on the 1979 Vela test event notes, “In late 1977 the Vorster government apparently suspended preparations to test. Strong U.S. pressure and other international reactions appeared to have deflected South Africa at least temporarily from testing” (CIA, 1979*b*, 6). The 1985 NIC report concluded that, in retrospect, such pressure helped support the non-proliferation norm “indirectly by causing Third World decision-makers to postpone actions – such as South Africa’s decision in 1977 not to test its nuclear device – which might have seriously damaged the nonproliferation regime” (NSC, 1985, 14). Policy officials stated, “For the past few years at least, South Africa’s nuclear program has not been a major source of contention in Africa nor has it had adverse implications for the United States except in relation to the international nonproliferation regime....This is partly because of its low visibility” (CIA, 1984, 17). Strategic obfuscation, then, helped to preserve the international norm.

Pakistan

The U.S. estimated in 1983 that Pakistan could produce a weapon by 1985-6, and in 1986 it determined that Pakistan had reached this capability (Montgomery and Mount, 2014). Pakistan:

 categorically denied our allegations and assured [the U.S.] orally for the first time that Pakistan would not manufacture a nuclear explosive device of any kind...intelligence

indicated that this procurement activity was continuing....There [was] overwhelming evidence that [Pakistan] ha[d] been breaking [its] assurances to [the U.S.]. We are absolutely confident that our intelligence is genuine and accurate. Moreover, intelligence available since Ambassador Walters' last visit to Islamabad indicates continued Pakistani weapons activities....The Pakistanis will not acknowledge publicly when and if they acquire a nuclear capability (Schultz, George, 1982, 3).

Yet the U.S. did not believe it could reverse Pakistan's program. In 1979, it concluded, "Unless we can somehow devise a regime to head them [Pakistan and India] off, we face the certain prospect of two nuclear weapons states in this troubled area" (U.S. Department of State, 1979). In fact, the U.S. also believed that "premature publicity could scuttle the effort" (U.S. Department of State, 1979). In 1982, it maintained this belief, stating "The odds are against any of the available options resulting in a complete termination of the Pakistan nuclear weapons program" (Schultz, George, 1982, 5). The U.S. noted that a strong reason that Pakistan was unlikely to reverse course was "Pakistan's apparent determination, at a minimum, to deny India a nuclear monopoly, and, at a maximum, to develop a nuclear weapons capability and thereby establish its status as a regional military and industrial power and as a leader of the Islamic world. One measure of Pakistani determination is the degree to which it has been willing to endanger its ties with the United States and western Europe by illegal clandestine imports of sensitive technology" (Cronin, 1988, 8). Further, "Pakistan's leaders are sensitive to public criticisms that Pakistan has become a 'pawn in a superpower rivalry,' and they shy away from publicly identifying with broader U.S. policy goals in the region" (Cronin, 1988, 10). Pakistan would thus resist coming into compliance, especially as a result of public U.S. pressure.

Additionally, the threat level from Pakistan acquiring a weapon was high. Pakistan's program was developed during Phases I and II of the NPT, which meant that its public defection could prompt pessimism globally and undermine the regime as a whole. This danger was exacerbated by regional tensions surrounding Pakistan. U.S. intelligence emphasized, "Stability in South Asia will be seriously weakened as Pakistan approaches a nuclear weapons capability threatening to

India...The likely alternative is that India will establish its own nuclear force, thus making India and Pakistan the first pair of nuclear armed adversaries in the Third World” (US Embassy in Seoul, 1982, 4). The U.S. believed that if Pakistan acquired nuclear weapons, it could “jeopardize our global non-proliferation strategy, which could collapse under the weight of two additional nuclear weapons states” (U.S. Department of State, 1979). The U.S. worried that it would “be faced by the mid-eighties, or even earlier, with two internally unstable nuclear weapons states with a long history of conflict, mutual mistrust, and unresolved problems” (U.S. Department of State, 1979).

Indeed, U.S. predictions about the potential effects of publicity were dire:

Pakistan’s nuclear weapons activities, if carried to completion, will lead to a nuclear arms race on the Subcontinent. This would result in greater regional insecurity, including the possibility of pre-emption by India or Israel or even eventually a nuclear exchange...Moreover, eventual transfer of nuclear technology or weapons by Pakistan to unstable Arab countries cannot be excluded. The intelligence community expects that Pakistani nuclear weapons activities will involve safeguards violations. Such violations would gravely undermine confidence in the IAEA safeguards system, which is critical to U.S. security and peaceful nuclear cooperation (Schultz, George, 1982, 4).

The U.S. thus desired to keep Pakistan’s program a secret, noting that it could potentially succeed at “limit[ing] the buildup of stockpiles to that done covertly rather than overtly” (Griffin, 1981, 10). Rather than dismantling Pakistan’s enrichment program, the U.S.’s goals were to dissuade it from conducting a nuclear test that would publicly display of Pakistan’s capabilities (CIA, 1981). Indeed, long after the U.S. knew that Pakistan possessed a nuclear device, the Reagan administration continued to certify that it did not (The National Security Archive, 2012).

However, it is important to note that in addition to the threat posed by Pakistan’s nuclear capabilities, the U.S.’s desire to keep the program a secret was overdetermined in this case because of Pakistan’s centrality in the Cold War. Publicizing Pakistan’s program would trigger foreign aid sanctions by the U.S. Congress, which U.S. President Ronald Reagan wanted to avoid. Reagan believed that “cutting off aid to Pakistan could endanger U.S. efforts to oppose the Soviets

in Afghanistan”(Cronin, 1988, 1). Thus, while the outcome is consistent with our theory, our theoretical mechanisms were not the only ones at work in this instance.

Iraq Before 1990

Under president Saddam Hussein, Iraq pursued a secret nuclear weapons program beginning in the early 1970s. While the U.S. was aware of Iraq’s efforts, it did not believe that it could reverse Iraq’s program through publicity and pressure. Policy officials stated, “Under most foreseeable conditions, however, Iraq’s interest in developing a capability to produce nuclear weapons will continue to be sustained by security concerns, regional ambitions, and recognition that its oil wealth and strategic location give it considerable leverage over its principal suppliers...The Iraqis can be expected to persist in trying to acquire all the necessary skills, facilities, and fissile material” (Director of Central Intelligence, 1979, 2-3). These officials further worried, “Denial efforts could be quite costly but still inadequate to keep Iraq from acquiring nuclear weapons (Director of Central Intelligence, 1979, 4). Thus, the U.S. concluded, “Given Iraq’s concern over potential military threats from Iran and Israel, we doubt that international pressure would dissuade Baghdad from its goal of establishing a nuclear weapons capability” (Directorate of Intelligence, 1989, 7).

However, the risk of second-order proliferation was very high. A U.S. intelligence memo stated that if Iraq acquired nuclear weapons, “The Arab-Israeli dispute would probably be significantly affected....[since] incidents that could trigger hostilities are common and few leaders have much understanding of the consequences of using nuclear weapons” (Director of Central Intelligence, 1979, 3). This threat was particularly severe at the time due to “the weakness of the existing international nonproliferation regime” (Director of Central Intelligence, 1979, 3), especially since the program was developed during Phase I and II of the NPT. Further, the memo notes that Iraq’s acquisition of nuclear weapons could “spark emulation elsewhere in the Middle East...Baghdad’s nuclear program has reportedly already stirred Syrian interest in mounting a similar effort, and, if pressed aggressively, Iraq’s quest for weapons-applicable facilities and technology could bring about a revival of Iran’s nuclear ambitions” (Director of Central Intelligence, 1979, 15).

Though overshadowed by its policy surrounding two subsequent wars, the reaction of U.S. presidents to Iraq's nuclear program *before* 1990 was one of strategic obfuscation. Iraqi nuclear progress in the 1980s occurred amid the Iran-Iraq War and its immediate aftermath. U.S. policy under both Reagan and George H. W. Bush prioritized regional stability over public and private criticism of Iraq's nuclear and other unconventional weapons programs. The result was a stark contrast between private U.S. intelligence assessments and public U.S. behavior. On the one hand, intelligence analysis of Iraq's nuclear progress, especially with the end of the Iran-Iraq War in 1988, sounded alarms. A top working group under the Director of Central Intelligence "identified Iraq as a key country on which to focus [intelligence] community resources" (Richelson, 2007, 336,353) and a National Intelligence Estimate described "Iraqi attempts to violate the Nuclear Non-Proliferation Treaty" (Karabell and Zelikow, 1994, 16-7). On the other hand, U.S. leaders consistently ignored this information in bilateral contacts with Iraq and did not reference Iraq's nuclear progress in public comments (Karabell and Zelikow, 1994, 8-12). A senior State Department official later defended the policy of ignoring Saddam's nuclear procurement by drawing parallels to other cases of ignored violations, explaining, "Violations of the Non-Proliferation Treaty did not lead to a U.S. break with Pakistan, South Africa, or Israel." Until Iraq's military approached Kuwait's borders, U.S. leaders' policy of constructive engagement led to obfuscation regarding Iraq's activities.³⁴ The Kuwait invasion and crumbling of the Soviet Union in 1990 altered conditions, prompting President George H. W. Bush to remind U.S. troops gathering in the Persian Gulf that action was critical due to the growing atomic threat from Baghdad (Richelson, 2007, 355-6).³⁵

³⁴One Congressional critic later argued "When it came to Iraq, the general policy of thwarting proliferation was at odds with the President's policy of increasing trade with Iraq as spelled out in NSD 26...despite abundant evidence of Iraq's intentions and military programs" (Gonzalez, 1992).

³⁵Analysis of the shift in conditions created by the 1990 Kuwait invasion, and U.S. policy towards Iraq thereafter, appears in the supplemental appendix.

Proliferator Unlikely to Comply and Low Risk of Reactive Proliferation

Our theory expects *E* to publicize when the threat of reactive proliferation is low, even if *A* is unlikely to come into compliance after being named-and-shamed. As we describe above, even a small chance that *A* will reverse course merits a publicity strategy, given that the second-order proliferation dangers are minimal and there are costs from exposed strategic obfuscation. In cases featuring these conditions, the U.S. should prefer to publicize programs on the off chance the proliferator does reverse course and to avoid international and domestic scrutiny later if obfuscation is exposed. We find that in the nuclear setting, these conditions held in two closely linked cases, Argentina and Brazil. We focus on Brazil here and analyze the U.S.'s reaction to Argentina's program in the appendix due to space constraints.

Brazil

Brazil began constructing a nuclear power plant in 1971, and by 1983 the U.S. had determined that Brazil could produce highly enriched uranium by the mid-1990s. The program remained a secret throughout the 1970s and 1980s, before dismantling it in 1990. While Brazil did acquire the technology necessary to produce nuclear weapons, it did not actually build them and instead acceded to the NPT in 1998. During the years in which it maintained its clandestine program, the U.S. determined that while Brazil would not likely give in to U.S. pressure, the threat environment was relatively low. As a consequence, the U.S. did not follow a path of strategic obfuscation.

The U.S. argued that Brazil would not come into compliance as a result of U.S. publicity and pressure, stating, "We believe Brazil...will persist in its refusal to place its nuclear research activities under safeguards and in its resistance to any U.S. effort to constrain its nuclear ambitions" (CIA, 1983, 2). Furthermore, U.S. officials claimed that Brazil's unwillingness to roll back its program was because, "Brazilian leaders clearly see the eventual mastery of nuclear fuel cycle technology as necessary for the great-power status to which they aspire...Brazil strongly resists what it perceives as foreign effort to limit its access to new equipment and technologies" (CIA, 1983, 8). Intelligence officials stated that therefore, "It will be difficult for the United States—or

any other nation—to have a major impact on its nuclear policies...Brazilian leaders have a strong determination to pursue their own self-interest—especially with regard to the acquisition and development of nuclear technology—and the confidence to oppose those who stand in their way” (CIA, 1983, 12-13). Brazilian officials also echoed these sentiments. For instance, an unidentified Brazilian embassy representative stated, “We know how to resist any and all U.S. pressures...Our nuclear program will continue, at least to the extent it depends on us, against all internal and external pressures” (US Embassy in Brazil, 1973, 1-2).³⁶

Indeed, U.S. pressure and publicity did not lead to Brazil’s compliance, but generated resentment instead. In addition, it had the unintended effect of propelling Brazil and Argentina to work together. While the two states had faced a brief period of tension over hydroelectric resources which had prevented them from reaching an understanding over their nuclear issues, U.S. pressure made both so angry that they resolved their differences and cooperated in the nuclear domain instead (Hymans, 2001).

Although publicity of Brazil’s program did not lead Brazil to comply with the regime, the threat environment was also not particularly high, leading the U.S. to reject strategic obfuscation. First, the codification of the norm was in its intermediate stages; therefore the threat of a defection to the overall regime was moderate. Further, the main concern regarding second order proliferation was that Argentina would respond to Brazil’s nuclear program by trying to develop its own nuclear weapon. Yet, “in contrast to India and Pakistan, neither Argentina nor Brazil perceived the other as having the intention to introduce instability into the Southern Cone by building nuclear weapons (Hymans, 2001, 161). Instead, the two countries held talks beginning in 1967 that helped to ensure that a nuclear arms race would not occur. In 1967, one of the government’s primary goals was to develop nuclear cooperation agreements with Argentina to reassure each other that nuclear energy would be used for peaceful purposes (Mallea, 2014*b*). Though the talks had several fits and starts, they picked up in the early 1980s and continued until the dismantling of the programs. This “fluid

³⁶In 1975, Brazil signed a secret deal with the U.S. to cancel its order of reprocessing plants from Germany in exchange for arms and security guarantees but the deal was leaked to the U.S. press and Brazil subsequently canceled it. After being burned once, it would not make such a deal again (Levite, 2006).

dialogue...kept both sides comfortably aware of their counterparts' intentions and capabilities and prevented the development of a nuclear rivalry" (Mallea, 2014a).

A further consideration of U.S. officials was that the extent of nuclear interest in both Brazil and Argentina, even when their rivalry was salient, appeared to be safely short of weaponization. As one author notes, "[t]he Brazilian military sought to reach the same technological level that Argentina was on the verge of achieving: that of enrichment capacity and the resulting nuclear option. In this regard the technology itself was seen as a 'species of deterrent;' the mere capacity to match a potential Argentine bomb was presumed sufficient to deter its construction" (Reiss, 1995, 15). The U.S. therefore believed the threat environment between Brazil and Argentina was not particularly high since "relations between the two nations are marked by intermittent rivalry rather than overt hostility" (CIA, 1983, 9). In affect, the U.S. chose not to hide Brazil's program from the international community.

Proliferator Likely to Comply and High Risk of Reactive Proliferation

Our theory expects *E* to publicize *A*'s activities if *A* is likely to come into compliance after being named-and-shamed, even if the threat of reactive proliferation is considerable. The logic behind this expectation is that, while it may exacerbate the threat and raise the stakes for global proliferation trends, publicity is likely to lead the proliferator back into compliance and strengthen the regime. We find that these conditions held for four cases: North Korea, Iraq after 1990, Taiwan, and South Korea. South Korea serves as our exemplar case, which we present here; the remaining cases in this category are analyzed in the appendix.

South Korea

The US correctly ascertained that South Korea was moving towards obtaining a nuclear weapon in 1975 (Montgomery and Mount, 2014). The threat environment in this case was high both because the risk came during Phase I of the norm's development and because the U.S. feared that North Korea would then obtain nuclear weapons in response. Moreover, there was a threat that North

Korea's activities would destabilize Japan, as well. U.S. internal policy memos highlighted the danger, stating, "In the case of Korea our general concerns are intensified by its strategic location and by the impact which any Korean effort to establish nuclear capability would have on its neighbors, particularly North Korea and Japan. ROK possession of nuclear weapons would have [a] major destabilizing effect in an area in which not only Japan but USSR, PRC, and ourselves are directly involved. It could lead to Soviet or Chinese assurances of nuclear weapons support to North Korea in the event of a conflict" (National Security Council, 1975, 1).

However, the U.S. expected public scrutiny and its own private threats to lead to compliance, particularly due to South Korea's reliance on the U.S.'s nuclear umbrella. Security guarantees were therefore thought to be a particularly effective inducement to convince South Korea to relinquish its program (CIA, 1978). Policy officials explained, "We believe that direct, early, and firm approach will be most appropriate and will have best chance of success" (US Embassy in Seoul, 1975, 2). The U.S. thus rejected strategic obfuscation, advocating for "several complementary policy courses...[that would] be evolved inside of, or in consonance with, the multilateral framework." These policies included working with other nations to "inhibit ROK access to sensitive technology and equipment,...press the ROK to ratify the NPT, [and] improve our surveillance of ROK nuclear facilities" (National Security Council, 1975, 3-5). While the close political relationship with South Korea provided the U.S. leverage, it also constrained its ability to explicitly and officially direct public criticism towards Seoul. Instead, U.S. intelligence and policy officials generated scrutiny of the South Korean program through unattributed comments and leaked intelligence analysis.³⁷

U.S. predictions were indeed correct, as "under U.S. pressure, in January 1976 it suspended negotiations for a reprocessing facility; in December 1976 it suspended the whole formal program to develop nuclear weapons technology that it had inaugurated only two years earlier" (CIA, 1978, 1). The U.S. therefore made no effort to strategically obfuscate, and successfully convinced South Korea to give up its program.

³⁷See, for example, headlines in major national newspapers warning of possible French cooperation on a reprocessing plant with Seoul quoting "Administration officials" (Gelb, 1975) and similar warnings about South Korean nuclear ambitions by "high-ranking Administration officials" in (Binder, 1976a).

Proliferator Likely to Comply and Low Risk of Reactive Proliferation

Finally, *E* is especially likely to publicize *A*'s violation of the norm if *A* is likely to come into compliance after being named-and-shamed and the threat of reactive proliferation is considerable. The intuition for this scenario is that publicity is likely to lead the proliferator back into compliance and strengthen the regime. Although such a result is not particularly urgent due to the low risk of second-order proliferation, the U.S. still prefers that *A* comply with the norm and thus publicizes. We find that these conditions held in two cases: Algeria and Libya. Libya serves as our exemplar case here, while Algeria is discussed in the Supplementary Appendix.

Libya

The U.S. had monitored Libya's nascent nuclear program since the early 1980s and found in 2001 that Libya was seeking dual-use technologies. By 2002 the U.S. believed that Libya possessed enough highly enriched uranium for a weapon by 2007, though in reality Libya did not (Montgomery and Mount, 2014). However, Libya represented a low threat for the nuclear non-proliferation regime both because the program was mainly developed during Phase III of the NPT, after which the norm was well codified, and because of Libya's isolation from the international community (NSC, 1985). The program was also far from being able to actually build a bomb, limiting the perceived threat to other nations (Crawford, 2004).

The U.S. expected naming-and-shaming to roll back the program, and indeed, it did. Libya's pursuit of nuclear weapons was discovered by U.S. and British intelligence and publicized (Bowen, 2006). Further, while multilateral sanctions had already been placed on Libya,³⁸ the U.S. required Libya to renounce its nuclear program before it would remove these sanctions, stating "The United States has raised the bar to give the condition a spin it did not have when the resolution was passed. It now holds that the resolution covers weapons of mass destruction as well" (Qadhafi, 2003, 44). Libya subsequently complied with the U.S.'s demands, which the U.S. attributed to the effectiveness of this strategy. The U.S. believed that, "Libya's announcement [of compliance]...is

³⁸Due to the Lockerbie incident

a product of the President's strategy which gives regimes a choice. They can choose to pursue WMD [weapons of mass destruction] at great peril, cost and international isolation. Or they can choose to renounce these weapons, take steps to rejoin the international community, and have our help in creating a better future for their citizens" (Qadhafi, 2003, 44). In addition, in line with the conventional wisdom on the effects of publicity, the U.S. stated, "These actions have sent an unmistakable message to regimes that seek or possess WMD: these weapons do not bring influence or prestige - they only bring isolation and other unwelcome consequences" (Qadhafi, 2003).

The U.S. further surmised that Libya relinquished its weapons because, "It was clear to Gaddafi that we were willing to use all the tools at our disposal to stem the flow of WMD. Ongoing international diplomacy, coupled with economic sanctions, isolated Libya and were having a significant impact on Libya's international status and economy. The Bush administration's relentless pursuit of the WMD black market exposed Libya's and other[s'] WMD programs, and diminished their chances of success" (Qadhafi, 2003). This case thus exemplifies the common belief about the result of naming-and-shaming; the violator will come into compliance and the regime will be upheld.

Conclusion

This article began by noting a recurring challenge faced by both societies and international political actors: building and maintaining order. Publicity of deviant activity has long been seen as a critical contributor to this goal. Conventional wisdom suggests publicizing rule violations helps to build and strengthen those rules. Although there is much truth to this, we suggest that the full picture is more complex. Drawing from research originating outside of political science that links ignorance and social stability, we build a theory explaining when and how publicizing deviance in international politics endangers international political order. We focus on two mechanisms related to the risk of second-order rule violations following the publicity of deviant behavior: publicity can both trigger pessimism among states about the overall rate of compliance and sharpen the threat posed to states by deviant behavior. This suggests that the regime and norm-enhancing effects of public-

ity identified by both rational institutionalists and norm theorists has overlooked a countervailing process. In short, we argue that “sunlight” has more than just disinfectant properties; shining a light on rule violations can in fact make international political order more difficult to manage.

We formally derive expectations about how regime advocates should act in light of our theory and evaluate it in a concrete empirical context. Our core finding is that the most powerful and vocal advocate of the nuclear non-proliferation regime, the United States, regularly resorted to obfuscation rather than publicity to safeguard this regime from the second-order dangers posed by specific nuclear programs. Consistent with the mechanisms we describe, U.S. leaders strategically obfuscated the progress and nature of nuclear programs in Israel, India, South Africa, Pakistan, and Iraq (prior to 1990). In contrast, these leaders publicized other nuclear programs when reversal was likely and the prospects of second-order proliferation dim. Our findings therefore suggest the existing view of publicity’s order-enhancing effects is valid but contingent; in broadly applicable conditions, the publicity of rule violations can pose serious threats to stability.

Our theory and findings are important in light of several recent research and contemporary policy challenges. First, our study speaks to the likely impact of Western efforts to address Iran’s nuclear program, which has been a lightning rod of controversy in recent years. Western governments have named-and-shamed Iran’s uranium enrichment program while simultaneously seeking a diplomatic solution, but whether this ongoing diplomacy yields a deal that endures remains to be seen. Our study theorizes some specific drawbacks in the West’s strategy of naming-and-shaming, which has led Western leaders to describe Iran’s program in dire terms and has raised the stakes of the ultimate settlement of this slow-burning crisis. Rhetoric and information disclosures have likely exacerbated threat perceptions of regional rivals and created additional pessimism about the overall health of the non-proliferation regime. Second, recent research on naming-and-shaming in the human rights contexts suggests many such campaigns fail to convince holdouts to comply (Hafner-Burton, 2008). Our findings suggest that such failed publicity campaigns can have previously under-explored deleterious effects on the regimes overall. Third, revisiting publicity’s impact on international order is especially important given what many see as a new era of transparency in

world politics ushered in by social media and the global convergence towards liberal democracy (American Bar Association, 2011). While extant political science scholarship tends to suggest that transparency and low-cost publicity are good news, our study suggests norms and regimes may be placed under greater stress if it is no longer feasible for states to occasionally engage in strategic obfuscation.

Our findings also have important ramifications for a variety of normative regimes in international relations, though our theory features mechanisms that may not be relevant in all situations. The danger of second-order violations, for example, depends in part on whether a rule regulates conduct with externalities that lead violations to create new threats. This is a modest scope condition in our view. The domain of nuclear weapons is an admittedly potent example of such externalities but is not qualitatively different from other issue areas. As we describe previously, many if not most security, economic, and environmental rules in world politics regulate an activity which has externalities. In addition, the efficacy of a strategy of obfuscation relies on a norm enforcer possessing better information about a case of deviance than the wider international community. We believe this is a similarly modest condition. The nuclear realm showcases U.S. information advantages on some of the most closely held secrets of other governments. Many regional and global powers regularly possess information advantages – even if only temporarily – which provide them with meaningful choices between publicity and obfuscation.

Finally, this article contributes to wider political science debates in several ways. For instance, we make a clear case for revisiting an important assumption underlying the logic of international institutions as well as constructivist studies of norm emergence and consolidation. Our theory also makes sense of under-analyzed and otherwise puzzling cases of states abstaining from publicizing rule violations. Such activity has been overlooked to date, yet the increasing availability of archival materials suggests scholars can now investigate the strategic choice to *not* use the spotlight. Future research may add to our findings by using a similar approach to investigate how the dynamics we uncover operate in other domains. Or, researchers might further explore the microfoundations of theories linking publicity, state reactions, and order—perhaps by adapting experimental approaches

from economics that expose subjects to different kinds of information about rule violations.

References

- Abram, Chayes and Antonia Handler Chayes. 1995. "The new sovereignty: compliance with international regulatory agreements." *Cambridge, MA: Harvard UP* .
- Albright, David. 2002. "Iraqs Programs to Make Highly Enriched Uranium and Plutonium Prior to the Gulf War." *Institute for Science and International Security* .
- Albright, David and Corey Gay. 1998. "Taiwan: Nuclear nightmare averted." *Bulletin of the Atomic Scientists* 54(1):54–60.
- American Bar Association. 2011. "No More Secrets: National Security Strategies for a Transparent World." *American Bar Association Standing Committee on Law and National Security, Office of the National Counterintelligence Executive, and National Strategy Forum* .
- Benjamin, Milton R. 2078. "Taiwan's Nuclear Plans Concern U.S. Officials: Taiwan's Nuclear Plans Worry U.S." *The Washington Post* .
- Bergeijk, PAG van. 1994. "Economic diplomacy, trade, and commercial policy: Positive and negative sanctions in a new world order." *Vermont: Edward Elgar* .
- Bergeijk, Peter AG van. 1989. "Success and failure of economic sanctions." *Kyklos* 42(3):385–404.
- Bernhard, Helen, Ernst Fehr and Urs Fischbacher. 2006. "Group affiliation and altruistic norm enforcement." *The American Economic Review* pp. 217–221.
- Bicchieri, Cristina and Yoshitaka Fukui. 1999. The great illusion: ignorance, informational cascades, and the persistence of unpopular norms. In *Experience, Reality, and Scientific Explanation*. Springer pp. 89–121.
- Binder, David. 1976a. "U.S. Fears Spread of Atomic Arms in Asia." *New York Times* .
- Binder, David. 1976b. "U.S. Finds Taiwan Develops A-Fuel: Officials Say the Processing of Plutonium Will Give it Capacity to Build A-Bomb." *The New York Times* .
- Bonetti, Shane. 1998. "Distinguishing characteristics of degrees of success and failure in economic sanctions episodes." *Applied Economics* 30(6):805–813.
- Bowen, Wyn Q. 2006. *Libya and nuclear proliferation: stepping back from the brink*. Vol. 380 Routledge.

- Brands, Hal. 2006. "Rethinking Nonproliferation: LBJ, the Gilpatric Committee, and US National Security Policy." *Journal of Cold War Studies* 8(2):83–113.
- Braut-Hegghammer, Målfrid. 2004. "Disarming Iraq?: the United Nations Special Commission 1991-98."
- Burr, William and Jeffrey T. Richelson. 2013. "Proliferation Watch: U.S. Intelligence Assessments of Potential Nuclear Powers, 1977-2001." *The Nuclear Vault: The National Security Archive* .
- Campbell, Kurt M, Robert J Einhorn and Mitchell Reiss. 2004. *The nuclear tipping point: Why states reconsider their nuclear choices*. Brookings Institution Press.
- CIA. 1978. "South Korea: Nuclear Developments and Strategic Decision-making." *National Foreign Assessment Center* .
- CIA. 1979a. "Iraq's Nuclear Interests, Programs, and Options." *Interagency Intelligence Memorandum* .
- CIA. 1979b. "Soviet Options in Afghanistan." *Interagency Intelligence Memorandum (CIA)*. *Central Intelligence Agency Electronic Reading Room TCS 3267-79*.
- CIA. 1981. "A Broader Bilateral Relationship." *MDR Release* .
- CIA. 1983. "Brazil's Changing Nuclear Goals: Motives and Constraints." *National Intelligence Estimate* .
- CIA. 1984. "Trends in South Africa's Nuclear Security Policies and Programs." *National Intelligence Estimate* .
- Cirincione, Joseph, Jessica T Matthews, George Perkovich and Alexis Orton. 2004. "WMD in Iraq: evidence and implications." *Biosecurity and bioterrorism: biodefense strategy, practice, and science* 2(1):51–55.
- Coe, Andrew and Jane Vaynman. 2014. "Collusion and the Nuclear Nonproliferation Regime." *Working Paper* .
- Cohen, Avner. 1998. *Israel and the Bomb*. Columbia University Press.
- Cohen, Avner. 2013. *The Worst-kept Secret: Israel's Bargain with the Bomb*. Columbia University Press.

- Cohen, Avner and Marvin Miller. 1990. "Facing the unavoidable: Israel's nuclear monopoly revisited." *The Journal of Strategic Studies* 13(3):64–76.
- Coleman, Katharina P. 2013. "Locating norm diplomacy: Venue change in international norm negotiations." *European Journal of International Relations* 19(1):163–186.
- Crawford, David. 2004. "Libya Was Far From Building Nuclear Bomb; Program Was Haphazard, But Shows How Technology Was Bought Off-the-Shelf." *Wall Street Journal* .
- Cronin, Richard P. 1988. "Pakistan's Nuclear Program: U.S. Foreign Policy Consideration." *Congressional Research Service* .
- Dai, Xinyuan. 2005. "Why comply? The domestic constituency mechanism." *International Organization* 59(02):363–398.
- Dai, Xinyuan. 2006. "The conditional nature of democratic compliance." *Journal of Conflict Resolution* 50(5):690–713.
- Davies, Rodger. 1969. "Memorandum from Rodger Davies to Mr. Austin et Al." *Israel Crosses the Threshold. Review Group Consideration of Response to NSSM-40, June 26, 1969* .
- Davis, Zachary S. 1990. "Iraq and Nuclear Weapons." *Congressional Research Service* .
- Department of State. 1966. "Memorandum." *NSC Meeting, National Archives and Records Administration Record Group 59, Central Foreign Policy Files* .
- Department of State. 1967. "Memorandum for the Secretary of Defense: The Indian Nuclear Weapons Problem: Security Aspects." *National Archives and Records Administration Record Group 59, Central Foreign Policy Files* .
- Department of State. 1974. "Telegram to U.S. Mission in Geneva, Indian Nuclear Tests."
- Director of Central Intelligence. 1979. "Iraq's Nuclear Interests, Programs, and Options." *Intergency Intelligence Memorandum* .
- Directorate of Intelligence. 1989. "Iraq: Nuclear Weapons-Related Procurement Activities." *CIA* .
- Dunn, Lewis A. 1991. *Containing Nuclear Proliferation*. Vol. 263 Brassey's for the International Institute for Strategic Studies.
- Dunning, Thad. 2004. "Conditioning the effects of aid: Cold War politics, donor credibility, and

- democracy in Africa.” *International Organization* 58(02):409–423.
- Finnemore, Martha and Kathryn Sikkink. 1998. “International norm dynamics and political change.” *International organization* 52(4):887–917.
- Franklin, James C. 2008. “Shame on you: the impact of human rights criticism on political repression in Latin America.” *International Studies Quarterly* 52(1):187–211.
- Gavin, Francis J. 2006. “Blasts from the Past: Proliferation Lessons from the 1960s.”
- Gelb, Leslie H. 1975. “Arms Expert Warns on Nuclear Spread.” *New York Times* .
- Gerber, Alan S, Donald P Green and Christopher W Larimer. 2010. “An experiment testing the relative effectiveness of encouraging voter participation by inducing feelings of pride or shame.” *Political Behavior* 32(3):409–422.
- Goffman, Erving. 2009. *Stigma: Notes on the management of spoiled identity*. Simon and Schuster.
- Gonzalez, Henry B. 1992. “Bush Administration Had Acute Knowledge of Iraq’s Military Industrialization Plans.” *Congressional Record, House of Representatives* .
- Griffin, Edward G. 1981. “Pakistan: Security Planning and the Nuclear Option.” *Bureau of Intelligence and Research* .
- Groeber, Patrick and Heiko Rauhut. 2010. “Does ignorance promote norm compliance?” *Computational and Mathematical Organization Theory* 16(1):1–28.
- Hafner-Burton, Emilie M. 2008. “Sticks and stones: Naming and shaming the human rights enforcement problem.” *International Organization* 62(04):689–716.
- Hersman, Rebecca KC and Robert Peters. 2006. “Nuclear u-turns: Learning from South Korean and Taiwanese rollback.” *Nonproliferation Review* 13(3):539–553.
- Hill, Ruth Vargas, Eduardo Maruyama and Angelino Viceisza. 2012. “Breaking the norm: An empirical investigation into the unraveling of good behavior.” *Journal of Development Economics* 99(1):150–162.
- Hymans, Jacques EC. 2001. “Of gauchos and gringos: Why Argentina never wanted the bomb, and why the United States thought it did.” *Security Studies* 10(3):153–185.
- Jervis, Robert. 1976. *Perception and misperception in international politics*. Vol. 49 Princeton

- University Press.
- Joeck, Neil. 1990. "Tacit bargaining and stable proliferation in South Asia." *The Journal of Strategic Studies* 13(3):77–91.
- Jones, Brian. 2004. "Intelligence, verification and Iraq's WMD." *Verification Yearbook 2004* .
- Kahler, Miles. 2000. "Conclusion: the causes and consequences of legalization." *International Organization* 54(3):661–683.
- Karabell, Zachary and Philip Zelikow. 1994. *Prelude to War: US Policy Toward Iraq 1988-1990*. Case Program, John F. Kennedy School of Government, Harvard University.
- Keck, Margaret E and Kathryn Sikkink. 1998. *Activists beyond borders: Advocacy networks in international politics*. Vol. 35 Cambridge Univ Press.
- Keohane, Robert O. 1984. *After hegemony: Cooperation and discord in the world political economy*. Princeton University Press.
- Keohane, Robert O. 1997. "International relations and international law: two optics." *Harv. Int'l. LJ* 38:487.
- Kimball, Daryl G. 2003. "Iraq's WMD: Myth and Reality." *Arms Control Today* 33(7):2–2.
- Kirshner, Jonathan. 1997. "The microfoundations of economic sanctions." *Security Studies* 6(3):32–64.
- Kissinger, Henry. 1969a. "Discussions with the Israelis on Nuclear Matters." *Memorandum from Henry Kissinger to Richard Nixon. Israel Crosses the Threshold. National Security Archive Electronic Briefing Book No. 189* .
- Kissinger, Henry. 1969b. "Israeli Nuclear Program." *Memorandum from Henry Kissinger to Richard Nixon. Select Documents, Mandatory Review Opening (11/28/2007), Nixon Presidential Library* .
- Kissinger, Henry. 1969c. "Memorandum from Henry Kissinger to Spiro Agnew." *Foreign Relations of the United States, 1969-1976, Volume E2: Documents on Arms Control and Non-Proliferation* .
- Kitts, James A. 2003. "Egocentric bias or information management? Selective disclosure and the

- social roots of norm misperception.” *Social Psychology Quarterly* pp. 222–237.
- Krasner, Stephen D. 1983. *International regimes*. Cornell University Press.
- Krook, Mona Lena and Jacqui True. 2012. “Rethinking the life cycles of international norms: The United Nations and the global promotion of gender equality.” *European Journal of International Relations* 18(1):103–127.
- Lam, San Ling. 1990. “Economic sanctions and the success of foreign policy goals: A critical evaluation.” *Japan and the World Economy* 2(3):239–248.
- Lantis, Jeffrey S. 2011. “Irrational Exuberance? The 2010 NPT Review Conference, Nuclear Assistance, and Norm Change.” *The Nonproliferation Review* 18(2):389–409.
- Legro, Jeffrey. 1995. *Cooperation Under Fire: Anglo-German Restraint During World War II*. Cornell University Press.
- Levite, Ariel E. 2006. “Never say never again: nuclear reversal revisited.” *International Security* 27(3):59–88.
- Litwak, Robert. 2000. *Rogue states and US foreign policy: containment after the Cold War*. Woodrow Wilson Center Press.
- Mack, Andrew and Asif Khan. 2000. “The efficacy of UN sanctions.” *Security Dialogue* 31(3):279–292.
- Mallea, Rodrigo. 2014a. “The Brazilian Proposal to Renounce Peaceful Nuclear Explosions and the Argentine Response (1983-1985).” *Nuclear Proliferation International History Project* .
- Mallea, Rodrigo. 2014b. “The First Attempt at Argentine-Brazilian Nuclear Cooperation and the Argentine Response, 1967-1972.” *Nuclear Proliferation International History Project* .
- Milgrom, Paul R, Douglass C North et al. 1990. “The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs.” *Economics & Politics* 2(1):1–23.
- Miller, Nicholas L. 2014. “The Secret Success of Nonproliferation Sanctions.” *International Organization* .
- Mitchell, Ronald B. 1998. “Sources of transparency: information systems in international regimes.” *International Studies Quarterly* 42(1):109–130.

- Montgomery, Alexander H and Adam Mount. 2014. "Misestimation: Explaining US Failures to Predict Nuclear Weapons Programs." *Intelligence and National Security* 29(3):357–386.
- Moore, Wilbert E and Melvin M Tumin. 1949. "Some social functions of ignorance." *American Sociological Review* pp. 787–795.
- Murdie, Amanda M and David R Davis. 2012. "Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGOs1." *International Studies Quarterly* 56(1):1–16.
- National Security Archive. 1974. "Prospects for Further Proliferation of Nuclear Weapons." *Special National Intelligence Estimate 4-1-74. National Security Archive Electronic Briefing Book No. 240* .
- National Security Council. 1975. "US National Security Council Memorandum, ROK Weapons Plans." *History and Public Policy Program Digital Archive, Gerald R. Ford Presidential Library, National Security Adviser Presidential Country Files for East Asia and the Pacific, Box 9, Korea. Obtained by Charles Kraus.* .
- Nooruddin, Irfan. 2002. "Modeling selection bias in studies of sanctions efficacy." *International Interactions* 28(1):59–75.
- NSC. 1969. "Minutes of National Security Council Meeting." *Foreign Relations of the United States, 1969-1976, Volume E2: Documents on Arms Control and Non-Proliferation* .
- NSC. 1977a. "Argentina's Nuclear Program." *Background Paper* .
- NSC. 1977b. "Non-Proliferation Country Problems." *Background Paper* .
- NSC. 1979. "U.S. Intelligence and the South African Bomb." *Memorandum from Christine Dodson to National Security Council. National Security Archive Electronic Briefing Book No. 181* .
- NSC. 1985. "The Dynamics of Nuclear Proliferation: Balance of Incentives and Constraints." *National Intelligence Council, NIC M 85-10001. U.S. Intelligence Assessments of Potential Nuclear Powers, 1977-2001. National Security Archive Electronic Briefing Book No. 451* .
- NSC. 1988. "Middle East-South Asia: Nuclear Handbook." *Weapons of mass Destruction* .
- Office of Scientific and Weapons Research. 1986. "Key Judgments: North Korea: Potential for Nuclear Weapon Development." *CPAS* .

- Owen, Henry. 1969. "Impact on U.S. Policies of an Israeli Nuclear Weapons Capability." *Memorandum from Henry Owen to the Secretary. Israel Crosses the Threshold. National Security Archive Electronic Briefing Book No. 189* .
- Perkovich, George. 1999. *India's nuclear bomb: the impact on global proliferation*. Univ of California Press.
- Perry, William J. 1999. *Review of United States Policy Toward North Korea: Findings and Recommendations*. The Nuclear Vault.
- Policy Coordinating Committee. 1990. "Nonproliferation Initiative for Iraq." *Nonproliferation PCC* .
- Policy Coordinating Committee. 1991. "The Algerian Nuclear Program." *U.S. Department of State* .
- Prentice, Deborah A and Dale T Miller. 1993. "Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm." *Journal of personality and social psychology* 64(2):243.
- Price, Richard. 1998. "Reversing the gun sights: transnational civil society targets land mines." *International Organization* 52(3):613–644.
- Qadhafi, Seif Al-Islam. 2003. "Libyan-American Relations." *Middle East Policy* 10(1):43–4.
- Redick, John R. 1995. *Nuclear Illusions: Argentina and Brazil*. Number 25 Henry L. Stimson Center.
- Reiss, Mitchell. 1995. *Bridled ambition: Why countries constrain their nuclear capabilities*. Woodrow Wilson Center Press.
- Richelson, Jeffrey. 2007. *Spying on the bomb: American nuclear intelligence from Nazi Germany to Iran and North Korea*. WW Norton & Company.
- Risse-Kappen, Thomas, Stephen C Ropp and Kathryn Sikkink. 1999. *The power of human rights: International norms and domestic change*. Vol. 66 Cambridge University Press.
- Roy, J. Stapleton. N.d. "Algerian Nuclear Program." *U.S. Department of State*. Forthcoming.
- Ruble, Maria Rost. 2009. *Nonproliferation Norms: Why States Choose Nuclear Restraint*. Uni-

- versity of Georgia Press.
- Sandholtz, Wayne. 2008. "Dynamics of international norm change: Rules against wartime plunder." *European Journal of International Relations* 14(1):101–131.
- Schrafstetter, Susanna. 2002. "Preventing the Smiling Buddha: British-Indian nuclear relations and the Commonwealth Nuclear force, 1964–68." *Journal of Strategic Studies* 25(3):87–108.
- Schultz, George. 1982. "Overwhelming Evidence that Zia has been Breaking his Assurances." *CREST* .
- Sikkink, Kathryn. 2002. "Transnational advocacy networks and the social construction of legal rules." *Global Prescriptions: The Production, Exportation, and Importation of a New Legal Orthodoxy* pp. 37–64.
- Simmel, Georg. 1950. *The sociology of georg simmel*. Vol. 92892 Simon and Schuster.
- Simmons, Beth. 2010. "Treaty compliance and violation." *Annual Review of Political Science* 13:273–296.
- Sisco, Joseph J. 1969. "Israel's Nuclear Policy and Implications for the United States." *Memorandum from Joseph J. Sisco to the Secretary. Israel Crosses the Threshold* .
- Smith, Hedrick. 1970. "US Assumes the Israelis Have A-Bomb or Its Parts." *New York Times* 18.
- State Department. 1974. "Telegram TOSEC 794/104621 from the Department of State to the Mission to the International Atomic Energy Agency." *Foreign Relations of the United States, 1969-1976, Volume E8: Documents on South Asia, 1973-1976* .
- State Department. 1975. "Foreign Relations of the United States, 1969-1976 Volume E-8, Documents on South Asia, 1973-1976, Document 197." *www.state.gov* .
- State Department. 1998. "Signatories and Parties to the Treaty on the Non-Proliferation of Nuclear Weapons." *www.state.gov* .
- Talbott, Strobe. 1999. "Dealing with the bomb in South Asia." *Foreign Affairs* pp. 110–122.
- Tannenwald, Nina. 2007. *The nuclear taboo: the United States and the non-use of nuclear weapons since 1945*. Number 87 Cambridge University Press.
- Tannenwald, Nina. 2009. "Deciding Not to Join the Nuclear Club."

- The National Security Archive. 2007. “The Algerian Nuclear Problem, 1991: Controversy over the Es Salam Nuclear Reactor.” *The Nuclear Vault* .
- The National Security Archive. 2012. “New Documents Spotlight Reagan-era Tensions over Pakistani Nuclear Program.” *The Nuclear Vault* .
- The National Security Archive. 2014. “The United States and the Two Koreas, Part II: 1969-2010.” *The Nuclear Vault* .
- Thompson, Alexander. 2009. “The rational enforcement of international law: solving the sanctioners’ dilemma.” *International Theory* 1(2):307–321.
- U.S. Department of State. 1979. “Non-Proliferation in South Asia.” *Cable 145139* .
- US Department of State. 1995. “Implications of an Indian Nuclear Weapon Test.” *NPIHP* .
- U.S. Department of State. 1998. “Road Map/Sanctions.” *The Nuclear Vault* .
- U.S. Department of State. 1999. “Script of Talking Points for William J. Perry.” *The Nuclear Vault* .
- U.S. Embassy in Algeria. 1991. “Nuclear Cooperation between Argentina and Algeria, Cable 01357.” *The Nuclear Vault* .
- US Embassy in Brazil. 1973. “Memo.” *National Archives and Records Administration Record Group 59, Central Foreign Policy Files* .
- US Embassy in Seoul. 1975. “ROK Plans to Develop Nuclear Weapons and Missiles.” *Gerald R. Ford Presidential Library* .
- US Embassy in Seoul. 1982. “National Intelligence Estimate, NIE-4-82, Nuclear Proliferation Trends Through 1987.” *Wilson Center Digital Archive* .
- Verdier, Daniel. 2008. “Multilateralism, bilateralism, and exclusion in the nuclear proliferation regime.” *International Organization* 62(3):439–76.
- Vertic. 1998. “Verifying Iraq.”
- Washington Post. 1976. “Reports Say Taiwan Producing A-Fuel.” *Washington Post* .
- Yarhi-Milo, Keren and Alex Weisiger. 2014. “Revisiting Reputation: How Do Past Actions Matter In International Politics.” *International Organization* .

Zwingel, Susanne. 2012. "How Do Norms Travel? Theorizing International Womens Rights in Transnational Perspective1." *International Studies Quarterly* 56(1):115–129.

Appendix

Model Proof

To review the discussion above, the model features three states: a norm enforcer (E), state A , and state B . Prior to the start of the game, neither A nor B have violated the norm of nuclear non-proliferation. The game begins with A choosing whether to violate the norm, which constitutes A 's type. Because we focus on whether such a violation is publicized, A 's decision remains exogenous, such that A violates the norm with probability q . The norm enforcer observes A 's decision and, if A violated the norm, decides whether to publicize A 's decision. State B does not observe A 's choice, but observes the norm enforcer's decisions. B 's beliefs about whether A violated the norm are determined by Bayes' Rule. If E publicized the norm, A then decides whether to reverse its action in response by coming into compliance with the NPT. The proportion of A s that do not come into compliance is given by r . B does not know A 's type, though E does. B then chooses whether to violate the norm.

We assume that E can never publicize a violation if a violation did not occur. Since E must present evidence of a violation, we think it is reasonable to assume that A cannot lie, which can be easily relaxed. Our solution concept is perfect Bayesian equilibrium, and we consider only pure strategies. A strategy for E is whether to publicize a violation $P = \{p, \sim p\}$. A strategy for B is to select whether to violate the norm, which maps E 's decision about whether to publicize onto B 's decision about whether to violate. An equilibrium thus consists of E 's decision about whether to publicize, B 's decision about whether to violate, and B 's beliefs β which are determined by Bayes' rule.

B receives utility from following the norm, denoted $\Omega(\cdot)$, which depends on the number of other states that follow the norm, and is assumed to be a concave function. The more states that follow the norm, the more utility B receives from also following it. Since this simple model consists of two states, B 's utility from following the norm in effect depends on whether A and B itself follow it. We normalize $\Omega(0) = 0$. Further, if A has nuclear weapons, B receives utility from violating

the norm due to its resulting enhanced capacity to defend itself against A , denoted d . B also faces a domestic benefit b and cost c associated with obtaining nuclear weapons. B 's utility function is thus specified as $U_B(v_A, v_B) = (1 - v_B)\Omega(2 - v_A - v_B) + v_A v_B d + v_B(b - c)$. Thus, if A violates the norm, and B decides to as well, B 's utility reduces to $d + b - c$.

The norm enforcer receives utility from other states following the norm, denoted $\Theta(2 - v_A - v_B)$, which we assume is a concave function. The norm enforcer's utility also depends on the potential domestic and international costs and benefits it may receive, which we leave in the background to minimize notation. Finally, the norm enforcer incurs a small cost from strategic obfuscation, $S = \{s, 0\}$, where $S = s$ if strategic obfuscation occurred, and 0 otherwise. The norm enforcer's utility function is thus $U_E(v_A, v_B, S) = \Theta(2 - v_A - v_B) - S$.

We solve the game for the Perfect Bayesian Nash equilibrium. To begin, note that if E observes a violation of the norm and A is the type that comes into compliance after E publicizes, E 's unique best response is to publicize the violation. After E does so, B will not violate the norm. Since B did not find it profitable to violate the norm prior to the start of the game, and nothing has changed since then, it is still not profitable for B to violate it. If E does not publicize the norm, A will not reverse its action, and E will incur a cost. Thus, publicizing in this case results in E 's preferred outcome.

Because E will always publicize a violation if, after it does so, A comes into compliance, we must only consider two potential pure strategies for E : publicize all violations by A , or publicize all violations when A will come into compliance, and strategically obfuscate otherwise.

Consider the first possibility; that E publicizes all violations. In this case, once B observes that E publicizes A 's violation, if A does not come into compliance, B violates the norm if $d + b - c \geq \Omega(1)$. Thus, B is more likely to violate it when d is higher, which occurs when A 's nuclear weapons pose a graver threat to B 's security, and when $\Omega(1)$ is lower, which occurs when the norm is not well codified. If A does not come into compliance, but B does not violate the norm, E receives utility equal to $\Theta(1)$, while if B responds with a violation, E receives $\Theta(0)$.

Now consider E 's second possible strategy: publicize all violations when A will come into

compliance, and strategically obfuscate otherwise. When no publication occurs, B then believes that a violation occurred with probability $\beta = \frac{q(1-r)}{1-qr}$ and violates the norm when $\beta(d + b - c) + (1 - \beta)(b - c) \geq \beta\Omega(1) + (1 - \beta)\Omega(2)$, or, simplifying, when $\beta(d) + b - c \geq \Omega(2) - \beta[\Omega(2) - \Omega(1)]$. Thus, B is more likely to violate the norm when it receives more utility from self-defense, when the norm is less well codified (i.e. when $\Omega(\cdot)$ is lower), and when β is higher. Plugging in the value of β , the results further indicate that B is more likely to violate when A is more likely to have violated the norm and when A is less likely to come into compliance after publicity. Note also that in this scenario, if A violates, but B does not, E receives utility equal to $\Theta(1) - s$, while if B violates, E receives $\Theta(0) - s$.

To determine E 's equilibrium strategy, we compare the pay-offs E receives in each case. When A will come into compliance, the payoffs are identical, so we need only focus on what occurs when A will not do so. Under the first strategy, if the parameters are such that B violates regardless, then E should always publicize violations. Yet this never occurs, or B would have violated at the start of the game. Thus, if the parameters are such that B responds to publicity when A does not comply with a violation, but does not violate when E strategically obfuscates, then E 's second strategy is optimal since $\Theta(N - \sum_1^{N-2} v_i - 2) \leq \Theta(N - \sum_1^{N-2} v_i - 1) - s$.³⁹ If B never violates regardless, E 's first strategy is optimal, since when A will not come into compliance, E receives $\Theta(N - \sum_1^{N-2} v_i - 1) \geq \Theta(N - \sum_1^{N-2} v_i - 1) - s$.

It remains to specify off-path beliefs. If E follows the first strategy but deviates to strategic obfuscation, B believes that a violation has not occurred with probability 1. However, since E always publicizes when B never violates, this does not affect the equilibrium.

Thus, if B 's self-defense needs are low and few states have violated the norm, such that it is the type that never violates the norm, E will always publicize a violation, since doing otherwise would incur a cost and B actions are the same regardless. However, if B is the type that reacts to a violation by violating the norm itself, E hides the violation. Thus, the greater the threat that A poses to B , and the fewer states that follow the norm, the greater the likelihood that publicizing

³⁹Given that s is sufficiently small.

would cause *B* to violate the norm when it otherwise would not, and the less likely *E* is to publicize.

Lemma 1. *When publicizing the norm does not lead *A* to reverse its violation of the norm, the greater the threat that *A* poses, and the fewer states that follow the norm, the less likely *E* is to publicize the norm.*

Additional Case Analysis

The article omits analysis of several cases featuring conditions that should lead *E*, or the United States, to publicize and pressure the proliferating state (*A*). We provide analysis of the conditions and the U.S.'s reaction here.

Argentina (Proliferator Unlikely to Comply and Low Risk of Reactive Proliferation)

The U.S. determined in 1974 that Argentina would have a nuclear weapons capability by the early 1980s. Indeed, Argentina began its nuclear weapons program in earnest in the early 1980s, but this was then abolished in 1983. In 1985 Argentina announced it would not develop nuclear weapons due to domestic and economic constraints (Montgomery and Mount, 2014). During this time, the U.S. concluded that publicity and pressure would not lead Argentina to reverse its decision to develop the weapons, but that the threat environment was low enough not to warrant strategic obfuscation. It therefore applied both bilateral and multilateral pressure on Argentina (Redick, 1995).

In the 1970s, the U.S. thought that perhaps publicity and pressure would succeed in persuading Argentina to comply, arguing in 1977, “They are reluctant to pay the high political price of a deferral of reprocessing. Privately, the Argentines have noted that this might be possible if Brazil also agreed to deferral. Hence we believe that there is a chance for our nuclear strategy to succeed as far as the Argentine side of the Brazil/Argentina equation is concerned” (NSC, 1977*a*, 3). However, initial U.S. efforts to persuade Argentina to give up its nuclear program instead led Argentina to dig in its heels. For instance, “In the late 1970s...the United States applied persistent pressure on Buenos Aires and Brasilia to modify their nuclear plans and policies. Both states reacted with

strongly nationalistic opposition to this pressure at the time and would do so in the future” (US Embassy in Seoul, 1982, 22). While some accounts conclude that Argentina actually never intended to obtain nuclear weapons at all, they argue that U.S. naming-and-shaming “backfired and led to an even greater expansion in Argentine nuclear capacities” (Hymans, 2001, 155). The U.S., learning from the failure of these efforts, concluded, “the various disincentives—including strained relations with neighboring states and with Western industrialized states, the potential for a long-term cutoff of foreign nuclear technology, uncertainty about the implications of possible Brazilian reactions—probably would not appear unmanageable to Buenos Aires” (US Embassy in Seoul, 1982, 21), and thus would not lead to compliance. Indeed, even once Argentina gave up its weapons, this decision was due to internal domestic change rather than US pressure (Hymans, 2001).

However, the threat environment was low enough that strategic obfuscation was unwarranted. First, the codification of the norm was in its intermediate stage (Phase II), such that a defection would represent only a moderate threat to the norm’s health overall. Instead, the main fear was that Argentina’s activities would lead Brazil to violate the rules as well. U.S. intelligence officials stated, “We believe, however, that if Brasilia became convinced that Buenos Aires had made a firm decision to develop, test, or deploy nuclear explosives, Brazil would initiate a serious effort to achieve its own nuclear weapons capability” (CIA, 1983, 1). Yet this possibility was seen as relatively unlikely to occur, as the U.S. also concluded that Brazil would not likely develop a nuclear weapon due to domestic and economic constraints (Montgomery and Mount, 2014). Indeed, nuclear weapons were not considered legitimate domestically in Brazil, further reducing the possibility that Brazil would actually develop them. Instead, “nuclear arms advocates faced a national normative consensus against the bomb, a norm originating in the international nuclear nonproliferation regime” (Reiss, 1995, 30). In addition, “Brazilian officers did not perceive a clear military threat, either nuclear or conventional in nature, from Argentina during this period. Brazilian officials recognized that only extreme and uninfluential nationalist sectors in Argentina favored construction of atomic weapons. Hence Brazil reportedly never engaged in research necessary to develop employable weapons. Nor is there any evidence that atomic weapons were ever integrated

into Brazilian military planning or strategic doctrine” (Reiss, 1995, 15). Rather than attempt to develop nuclear weapons, Brazil cared more about simply possessing the technology as a deterrent, as “the mere capacity to match a potential Argentine bomb was presumed sufficient to deter its construction” (Reiss, 1995, 15). Finally, the U.S. believed that while the Argentina and Brazil cases were somewhat threatening to the normative regime, the threat was low and much more manageable than the threats coming from the Middle East and Southeast Asia (US Embassy in Seoul, 1982). The U.S. thus chose not to strategically obfuscate.

Iraq After 1990 (Proliferator Likely to Comply and High Risk of Reactive Proliferation)

Despite having ratified the NPT, Iraq sought nuclear weapons to counter Iran throughout the 1980s, actually using chemical weapons against Iran starting in 1982 (Jones, 2004). During the Gulf War, the coalition of states that opposed Iraq’s invasion of Kuwait was threatened by the possibility that Iraq could wield chemical or biological weapons against it, which also increased the salience of the nuclear issue (Jones, 2004). The war’s ceasefire agreement therefore required Iraq to destroy its nuclear capabilities and then provide verification of this action. It further stipulated that Iraq would be monitored to verify that it did not restart its nuclear weapons’ programs. The UN Security Council resolution 687 codified these demands and gave the responsibility for ensuring they were met to the IAEA (along with the UN Special Commission (UNSCOM) for chemical and biological weapons) (Jones, 2004).⁴⁰

The IAEA was tasked with ongoing monitoring and verification (OMV) of Iraq’s efforts, which constituted the strictest international nuclear investigation ever undertaken. Meanwhile, the U.S. determined in 1991 that Iraq would have highly enriched uranium by the late 1990s and that it

⁴⁰It is important to note that Iraq after 1990 differs from some other cases with similar underlying conditions in one important way: its relationships with the United States. Iraq after 1990, North Korea, South Korea and Taiwan all were cases in which the proliferating state was seen as likely to comply by the U.S. but their programs raised serious risks of second-order proliferation. Close political relations with U.S. leaders in the latter two cases, however, led Washington to exert public scrutiny via leaks and supplement publicity with private pressure (rather than more overt coercion). For Iraq after 1990 and North Korea, where relations with Washington were far more acrimonious, a more traditional naming-and-shaming style was adopted. See the analysis of South Korea and Taiwan elsewhere for more on these points.

could develop a crash bomb within six months to a year (Montgomery and Mount, 2014). In 1999, the U.S. believed Iraq could develop a weapon in 5-7 years, and thought it had begun a program (though it had not) in 2002 (Montgomery and Mount, 2014). Inspectors from these bodies discovered and disseminated knowledge of Iraq's efforts to obtain nuclear weapons, which included an advanced enrichment program. The IAEA produced several reports detailing Iraq's efforts (Albright, 2002). These activities were heavily publicized, including extensive media coverage of UN inspectors' activities, and even the destruction of weapons stockpiles. According to a prominent NGO, this gave "unprecedented publicity to an activity—verification and monitoring—that was hitherto perceived as mysterious and arcane" (Vertic, 1998).

Indeed, verifying and publicizing Iraq's activities was established as the desired strategy to deal with Iraq's program from the outset. A recently declassified memo states,

We propose the following steps with our allies and other countries that have substantive relations with Iraq. [1] Share information on Iraq's nonconventional weapons capabilities and related procurement networks with the MTCR, Australia Group, and other allied countries. [2] Urge these countries to contact Iraq to voice concern over its programs and their negative effect on regional tensions, and to strengthen controls on exports that could contribute to Iraq's nonconventional weapons development. [3] Urge these countries to implement or strengthen controls over the provision of defense services that could contribute to nonconventional weapons programs in Iraq and other countries. [4] The U.S. itself could seek, and urge other potential supplier countries to seek legislation similar to that recently passed by the Federal Republic of Germany, specifically making it a crime for citizens to assist foreign nuclear or CBW programs. (In the U.S. this would supersede current sanctions for violating the licensing provisions of the Atomic Energy Act, Arms Export Control Act or Export Administration Act, and supplement a recent law imposing criminal penalties for support of foreign BW programs. [5] Use "public diplomacy" to expose the activities of foreign firms and individuals supporting Iraqi weapons programs (Policy Coordinating Committee,

1990).

Further, the threat environment was high; though the Iraq pursued the weapons during Phase III of the norm, after which time it was well codified, regional self-defense concerns were large. The U.S. believed that “Baghdad’s nuclear program has reportedly already stirred Syrian interest in mounting a similar effort, and, if pressed aggressively, Iraq’s quest for weapons-applicable facilities and technology could bring about a revival of Iran’s nuclear ambitions. The actual or impending acquisition of nuclear weapons by a country with Iraq’s interoperate credentials would be particularly upsetting to the Saudis and other Persian Gulf Arabs. The Saudis have long regarded Iraq as a major rival for influence in the Arab world and especially in the Gulf, but they are well aware they cannot match Iraqi military might.” (CIA, 1979*a*, 15). Additionally, according to a 1992 Congressional Research Report, “An overt or secret Iraqi attempt to acquire nuclear weapons would be the first violation of the NPT by a non weapons member. Were this to happen and no sanctions be applied, other members of the treaty might well question its value...Also, such an Iraqi venture could trigger an open or secret arms race in the Middle East, which could be seen as destabilizing and increasing chances that these weapons might be used in this region” (Davis, 1990).

As anticipated by many policy-makers at the time, the sanctions regime and inspections led to a high degree of cooperation. While Iraq initially resisted the IAEA’s efforts, General Hussein Kamel Hassan’s defection led to more complete disclosure. From 1995-6 in particular, the IAEA, UNSCOM, and domestic intelligence organizations worked together closely and achieved many of their goals. In particular, Iraq supplied over 50,000 pages of secret papers, 17 tones of managing steel, and enough carbon fiber for 1000 centrifuges. Numerous accounts agree that the program “served as a powerful deterrent to Iraq when it was in place” (Albright, 2002), and that it was very effective in dismantling Iraq’s nuclear program (Braut-Hegghammer, 2004; Cirincione et al., 2004; Kimball, 2003), though Iraq often made unilateral changes to avoid negative domestic publicity. While the inspectors then left Iraq in 1998, as the UN consensus broke down, up to that point, their efforts had largely succeeded.

Taiwan (Proliferator Likely to Comply and High Risk of Reactive Proliferation)

In 1974 the U.S. correctly assessed that Taiwan was conducting a nuclear program with the intent of developing nuclear weapons. In 1987 it discovered that Taiwan was trying to acquire a reprocessing plant (Montgomery and Mount, 2014). It was reported, “A US interagency team recently returned from making a comprehensive inspection of nuclear facilities in Taiwan. There is strong (though circumstantial) evidence when intelligence reports are combined with the team’s discoveries in Taiwan that the ROC illegally diverted and clandestinely reprocessed spent fuel...The IAEA did not detect this probable diversion, despite an intensive inspection last July...Any such diversion would be a clear violation of the agreement between the IAEA and the ROC on the TRR reactor. It would also be a violation of the spirit of the NPT and arguably of the letter of the Treaty as well” (NSC, 1977*b*).

Further, the threat of second order proliferation was high in this case. The proliferation risk came during phases I and II of the norm, so that a defection could potentially harm the norm’s health. Further, regional self-defense concerns loomed large. The Washington Post reported that the program “could cause major difficulties with mainland China, multiply the nuclear worries of Japan and of less developed Asian states, and accelerate a worldwide drift toward nuclear proliferation” (Washington Post, 1976).

However, the U.S. thought it could reverse Taiwan’s program through public scrutiny and private pressure. An internal policy memo reported, “In Taiwan, however, where there already is a perception of a declining US commitment, fear that secret nuclear weapons development would further accelerate this decline will act to discourage such activities” (US Embassy in Seoul, 1982, 5). This agreement with the U.S. was made public (Albright and Gay, 1998), and key accounts conclude that “U.S. determination to prevent Taiwan from obtaining nuclear weapons had paid off” (Albright and Gay, 1998, 60). As a result, “The United States increased public and private pressure on Taiwan to end all nuclear weapons-related activities. Washington threatened to cut off all fuel supplies, demanded the return of all plutonium of U.S. origin, and hinted that Taipei’s actions threatened to weaken the U.S. security guarantee and could result in a freezing of weapons

sales to the island” (Hersman and Peters, 2006, 544). As in the case of South Korea, close political relations constrained how directly and openly U.S. officials could criticize Taipei. Instead, U.S. intelligence and policy officials generated scrutiny of the Taiwan nuclear program through a regular stream of unattributed comments and leaked intelligence analysis in major news outlets.⁴¹

Thus, the U.S. relied on both bilateral and international pressure, and did not engage in strategic obfuscation in the sense that it did not attempt to cover up the program. In response, Taiwan first abandoned the nuclear program in 1976, resumed the program in 1987, and gave it up again in 1988, after the U.S. presented intelligence on Taiwan’s activities to Taiwan. Taiwan then agreed to ban all research related to nuclear weapons (Campbell, Einhorn and Reiss, 2004).

North Korea (Proliferator Likely to Comply and High Risk of Reactive Proliferation)

North Korea began building the Yongbyon nuclear site in 1962, joined the IAEA in 1974, and agreed to safeguards in 1978. However, its graphite-moderated reactor went critical, and it began to build another one in the late 1980s. One of its lines for reprocessing fuel was almost finished in 1994, and in 1992 the U.S. determined that North Korea could produce a weapon in anywhere from a few months to a couple of years. In 1998 the U.S. discovered a nuclear complex at Kumchang-ni, and in 2001 it found that North Korea possessed a centrifuge enrichment program (Montgomery and Mount, 2014).⁴²

The danger of reactive proliferation to this program was relatively high. The overall threat to the norm varied overtime since North Korea’s program spanned Phases I, II, and III of the NPT, but regional tensions made second-order proliferation an especially large concern. The U.S.

⁴¹See, for example, concerns about Taiwan’s nuclear ambitions expressed by “top U.S. officials” in the Carter Administration in Benjamin (2078) and, two years earlier, confirmation by “Administration officials” that leaked intelligence analysis describing diverted nuclear fuel by Taiwan was correct in Binder (1976*b*).

⁴²It is important to note that North Korea differs from some other cases with similar underlying conditions in one important way: its relationships with the United States. Iraq after 1990, North Korea, South Korea and Taiwan all were cases in which the proliferating state was seen as likely to comply by the U.S. but their programs raised serious risks of second-order proliferation. Close political relations with U.S. leaders in the latter two cases, however, led Washington to exert public scrutiny via leaks and supplement publicity with private pressure (rather than more overt coercion). For Iraq after 1990 and North Korea, where relations with Washington were far more acrimonious, a more traditional naming-and-shaming style was adopted. See the analysis of South Korea and Taiwan elsewhere for more on these points.

concluded, “Exposure could lead South Korea—with its superior nuclear technology—to develop nuclear weapons as a response” (Office of Scientific and Weapons Research, 1986, 4). A U.S. policy memo summed up the U.S.’s concerns as follows: “Acquisition by the DPRK of nuclear weapons...[would] undermine the conditions for pursuing a relaxation of tensions, improved relations, and lasting peace [in the region]. Acquisition of such weapons by North Korea could also spark an arms race in the region and would surely do grave damage to the global nonproliferation regimes covering nuclear weapons” (Perry, 1999, 3-4). Indeed, U.S. policy reports noted that “the most important factor in South Korea’s future nuclear decisions will be...the imminence of the North Korean threat” (CIA, 1978, 2).

However, despite the high threat level, the Clinton Administration initially believed that North Korea would comply with a bargain reducing sanctions in exchange for concessions regarding the nuclear program. The Administration also enlisted the help of other nations, holding the Four Party Peace Talks between the U.S., North Korea, South Korea, and China, which “provided the framework for the removal of the TWEA sanctions” (U.S. Department of State, 1998, 2) and stated that “the Four Party process is the best and most appropriate mechanism to make additional progress” (U.S. Department of State, 1998, 2). This forum was viewed “as a key means to securing greater cooperation from Pyongyang” (The National Security Archive, 2014) and was expected to lead to “an era of decisively improved relations between the US and the DPRK” (U.S. Department of State, 1999, 3).

The Clinton Administration thus relied on multilateral pressure, particularly during the Four Party Talks, to reduce tensions and discuss the nuclear situation. Even during bilateral talks, the U.S. frequently emphasized that other nations were aware of the proceedings. For instance, the U.S. would state that it had also “consulted with the governments of China and Russia, and with the European Union” (U.S. Department of State, 1999, 1).

After a period of engagement and apparent progress in reversing North Korea’s nuclear ambitions, North Korea terminated the agreement shortly before President Bill Clinton left office. The George W. Bush administration continued the general strategy of publicly naming-and-shaming

North Korea with two changes: sharpened rhetoric (i.e. North Korea as a member of the “axis of evil”) and a shift away from negotiated agreements to coercion.⁴³ This shift in tactics did not achieve meaningfully different results. Attempts by the U.S. and other governments to safeguard the regime and regional rivals from a North Korean nuclear arsenal were thwarted with its nuclear test in 2006.

Algeria (Proliferator Likely to Comply and Low Risk of Reactive Proliferation)

The U.S. discovered Algeria’s construction of a nuclear reactor in 1991 and found that it had acquired a reactor—Es Salam—from China. Further, the U.S. noted, “Algeria’s record on timely completion of safeguards negotiations with the IAEA on the [Argentine reactor] is not a good one” (Policy Coordinating Committee, 1991, 2). Despite domestic turmoil, however, the U.S. expected publicity to lead to compliance. It stated, “Algeria, which has not signed the Nuclear Non-Proliferation Treaty (NPT), probably can be persuaded to accept limited safeguards, which will help to ensure that the country’s low proliferation potential is kept to a minimum” (NSC, 1988, 1). The U.S. thus did not seem to believe that strategic obfuscation was necessary.

Further, the U.S. also determined that the threat level remained low, noting that if a “significant nuclear proliferation risk” emerged, it would explore “other options” (Policy Coordinating Committee, 1991). And while “the Moroccans were quite concerned about the Algerian nuclear program....there was little real basis for this worry” (U.S. Embassy in Algeria, 1991). Further, Algeria’s program was developed during Phase III of the NPT, such that the threat to the overall norm’s health was small.

However, the U.S. did not have much of a chance to react to Algeria’s activities, as news about Algeria’s program was leaked to the *Washington Times* in the spring of 1991 (The National Security Archive, 2007).⁴⁴ The U.S. had only acquired intelligence of the reactor’s construction

⁴³For example, the National Security Strategy and the Nuclear Posture Review declared it a belligerent state (Montgomery and Mount, 2014).

⁴⁴The sourcing for this story and evidence of displeasure within the Bush Administration regarding the disclosure does not suggest this was an intentional leak, as with past news stories about countries of concern like South Korea and Taiwan.

months before the story broke. It stated, “We do not have sufficient information from which to conclude that the [Algerian Government] has decided to pursue a military nuclear program” (Policy Coordinating Committee, 1991, 3). Thus, while it seems that the U.S. chose not to strategically obfuscate, it is difficult to determine what would have happened had the leak not occurred.

However, after the leak, Algeria responded to the publicity and pressure by coming into compliance, as the U.S. expected. According to a recently declassified cable, “To reinforce direct U.S. diplomatic pressure on Algeria (and China), we have sought to enlist seven other countries—France, the UK, Germany, Italy, Spain, Portugal and Japan—to exert influence on Algeria to ensure the Algerian nuclear program is devoted exclusively to peaceful purposes and is fully safeguarded....We have approached the seven countries noted above with information currently available to us and have requested that they supplement it to whatever extent possible. We have made these approaches in diplomatic channels and, in parallel, in intelligence channels” (Roy, N.d., 4) The U.S. also “encouraged the IAEA to offer to visit the Ain Oussera [village] facility well before its completion” (Roy, N.d., 4), sought China’s help in convincing Algeria to follow nuclear nonproliferation, and pressured Switzerland not to sell a hot isostatic press that could be used for nuclear development purposes to Algeria. Despite a civil war raging in Algeria, it complied by agreeing to sign the NPT in 1993, and then signing it in 1995. Thus, “Washington, Beijing, and the international community brought Algiers into the NPT-system within a few years of the controversy” (The National Security Archive, 2007).

Summary of all Cases (Table 3)

Table 3: Summary of Cases

Proliferator	Will it Comply?	Regional Threat	Norm Stage/Risk	Strategic Obfuscation?	Consistent with Theory?
Israel	No	High	I/High	Yes	Yes
India	No	High	I/High	Yes	Yes
South Africa	No	High	Mostly I/High	Yes	Yes
Pakistan	No	High	I,II/Medium-High	Yes	Yes, but overdetermined
Iraq (Pre-1990)	No	High	I,II/Medium-High	Yes	Yes
Brazil	No	Low	Mostly II/Medium	No	Yes
Algeria	Yes	Low	III/Low	No	Yes
Argentina	No	Low	Mostly II/Medium	No	Yes
Libya	Yes	Low	Mostly III/Low	No	Yes
North Korea	Yes	High	I,II,III/Medium	No	Yes
Iraq (Post-1990)	Yes	High	III/Low	No	Yes
Taiwan	Yes	High	I,II/Medium	No	Yes
South Korea	Yes	High	I/High	No	Yes

Notes: The table displays the hypotheses derived from the model.