The Reputational Consequences of International Legal Commitments

Geoffrey P.R. Wallace
Rutgers University
Niehaus Center for Globalization and Governance
geoffrey.wallace@rutgers.edu

***Research in progress, please do not cite without author’s permission***

Paper prepared for the International Relations Faculty Colloquium at Princeton University, February 9, 2015.

Acknowledgements: Earlier versions of this paper were presented at the 2013 annual meeting of the International Studies Association and the 2014 annual meeting of the American Political Science Association. Financial support was provided by the National Science Foundation (NSF) through Time-sharing Experiments for the Social Sciences (TESS). For their helpful comments on various parts of the research design and argument, I thank Ryan Brutger, Stephen Chaudoin, Emilie Hafner-Burton, Sarah Kreps, Jack Levy, Giovanni Mantilla, Paul Poast, Elizabeth Saunders, TESS principal investigators Jeremy Freese and Penny Visser, Michael Tomz, Sophia Jordán Wallace, as well as participants in the Survey Experiments in Peace Science Workshop at the annual meeting of the Peace Science Society (International). All errors remain my own.
Abstract

International legal scholars frequently claim states use international law as a commitment device that raises the reputational costs of noncompliance. According to this logic, violating a legalized commitment damages a state’s reputation as a reliable partner, which consequently reduces the willingness of other actors to cooperate with the violator in the future. Testing reputational mechanisms has been complicated by selection effects both in the design and joining of an agreement, as well as a state’s likely anticipation of the consequences of subsequent noncompliance. This paper uses an experiment embedded in a national survey to offer a more reliable estimate of the effect of legalized commitments on a country’s reputation for cooperation. Consistent with a reputational account, support for future cooperation with a foreign country declines markedly when that country violates a legalized agreement it previously made compared to a similar situation where it did not undertake a legalized commitment. However, the results also suggest the reputational consequences of international law are conditioned by the level of legalization in a commitment, the issue area at stake, as well as asymmetries in possible spillover effects across issues. The study thus has implications for the reasons countries employ international agreements and the consequences of international law on public opinion and state behavior.
Introduction

How does international law work to encourage states under anarchy to abide by their commitments? Skeptics question whether international law exists in any meaningful sense (Bolton 2000; Mearsheimer 1994-1995). Reasoning from the domestic legal analogy of laws enforced from above by higher authorities leads to claims such as that, “Law is to be found within nations rather than above them. There is no world state and therefore no world law” (Fromkin 1981:75). International organizations have undoubtedly gained greater influence across a range of issues in recent years (Hawkins et al. 2006), yet the absence of truly supranational institutions is frequently taken as proof of the limits of international law. Even if most commitments are regularly observed, countries may nevertheless face few genuine constraints from reneging on their promises when it suits their interests (Morgenthau 1985:295).

A growing chorus of international legal scholars counter that the absence of omnipotent international institutions does not mean that international law is without any tools of enforcement (Simmons 1998). A number of horizontal enforcement mechanisms instead exist that increase the costs of violations and thus facilitate compliance with international legal commitments. States, alongside a growing chorus of domestic, international, and transnational actors, thus monitor and punish violations in an anarchic, yet increasingly legalized, international system. Foremost among these decentralized sources of enforcement is the role of reputation on a country’s decision over whether or not to comply with its commitments. Legalizing a commitment is thought improve compliance by raising the ex post costs of violating promises made, in particular by making it more difficult or expensive for the violator to find others willing to cooperate in the future (Simmons 2010:275-7). Reneging on a treaty promise in turn carries

---

1 Though the distinction of domestic legal systems from their international counterpart should not be exaggerated, since compliance with domestic laws frequently depends on much more than purely the coercive apparatus of the state (Tyler 1990:19-27).
much greater costs for a country’s trustworthiness as a reliable partner than if the country had engaged in the exact same behavior but in the absence of a prior legalized commitment (Guzman 2008:34-5). Reputation thus allows international law to function in an effective manner even in the absence of a world government or some form of higher enforcement authority.

Many works have put forward variants of a reputational logic for compliance with international agreements (Tomz 2007; Henkin 1979; Keohane 1984), yet reliably estimating the reputational consequences of international law remains problematic. While several studies point to higher levels of compliance with formal legal agreements as support for reputational arguments (Simmons 2000; Fortna 2004:21-2), the motives behind this behavior could also be due to other unrelated mechanisms. Moreover, as many legal scholars recognize, empirical research on the consequences of international law face severe problems of selection effects (von Stein 2005; Shaffer and Ginsburg 2012:14-5). Since joining an agreement is voluntary, it may be difficult to determine whether any effects of international law are due to the treaty itself, or rather baseline characteristics that led a state to ratify or reject the treaty in the first place (Downs et al. 1996). Inferential problems are further compounded when attempting to assess specific compliance mechanisms like reputation. Since states are strategic actors, they are likely to take into account reputational dynamics not only at the stage of designing and joining to a legal agreement, but also at later points when faced with decisions over whether or not to abide by their commitments. Several promising studies have shown support for the reputational costs of violating legal commitments (Gibler 2008; Nelson 2010), but these findings remain vulnerable to the charge that they fail to adequately take into account important counteracting selection processes.
I seek to overcome the methodological limits in existing observational studies of international law by leveraging the use of an experimental design embedded in a public opinion survey. In the experiment, participants are presented with different types of scenarios where a foreign country either complies with or violates different types of legal commitments. Following the relevant literature on international law, reputational costs or benefits are primarily conceived as the consequences for a state’s ability to find willing partners for future cooperative endeavors. Participants are then asked how much they support or oppose the United States concluding an international agreement with the foreign country. Randomly assigning participants to different treatment or control groups for the existence or type of legal commitment overcomes problems posed by selection effects, and ensures any differences in responses are due to the legal treatment rather than confounding baseline factors. Examining differences in support for future cooperation across experimental groups allows for a more direct estimate of reputational mechanisms for compliance with legal commitments. By focusing on mass attitudes, the design also contributes to the recent “domestic turn” in the study of international law, which seeks to identify and test internal determinants of compliance (Simmons 2009:125-55).

The remainder of the paper proceeds as follows. The first section puts forward the logic underlying the reputation mechanism for international law, and generates several testable implications. I then describe the survey instrument and research design. The next section demonstrates that legal commitments do raise the reputational stakes involved, but these effects are conditioned by the issue area under dispute, the decision to violate, and the relative legalization of the commitment. The final section concludes with a summary of the findings and discussion of avenues for further research.

---

2 For a general overview of the merits of experimental methods for the study of international law, see (Chilton and Tingley 2013).
Theory

Reputation has been used to explain decision making across a variety of areas in political science (e.g. Kydd 2005; Walter 2009), but also in related fields including economics, sociology, and psychology (Kreps and Wilson 1982; Raub and Weesie 1990; Nisbett and Ross 1980). While an actor can try to take certain steps to shape its reputation, ultimately it is a function of others’ beliefs. “A man’s reputation is not a quality that he possesses, but rather the opinions which other people have about him” (Bailey 1971:4). At its core, a reputational approach argues that an actor’s past behavior provides others with a reliable guide for how they will act in the future. The credibility of current commitments thus depends in large part on whether or not an actor kept their past promises. This may then translate into actors making choices that may be extremely costly to them in the short term, which only make sense when seen from the perspective of seeking to maintain or bolster one’s reputation for the future. As Schelling (1966:124) notes in his analysis of crisis diplomacy, “If the question is raised whether this kind of ‘face’ is worth fighting over, the answer is that this kind of face is one of the few things worth fighting over.”

Reputational arguments in political science have mostly frequently advanced in the study of military threats and crisis diplomacy (Fearon 1997), but are also relevant for international law and international cooperation more generally (Axelrod 1984:150). To encourage others to cooperate in the first place often requires countries to develop a reputation as a trustworthy partner that keeps their promises even if difficult times (Simmons 2000). Since many matters of international affairs involve coordination problems where actors often have a short-term incentive to defect, on matters ranging from implementing trade barriers to cheating on arms control agreements, developing a good reputation can help alleviate concerns over violations (Stein 1982). Even in situations where cooperation still takes place, having a strong reputation
for cooperation can allow actors to bargain for better deals or achieve terms at lower costs since they are viewed as more reliable (Tomz 2007:58). On the other hand, once a country develops a poor reputation, it can be very difficult to overcome even with much effort over a long period of time (Jervis 1989:6).

Although reputation is generally viewed as an asset states seek to cultivate, far less is understood about how reputations form and change. One way in which international law is thought to make commitments more credible is by raising the consequences for a state’s reputation of its subsequent decision over whether or not to comply or violate the agreement (Schelling 1960:37). Violating any agreement could well hurt a country’s reputation as a trustworthy partner, but the more specific claim here is that reneging on a legal commitments carries with it much higher reputational stakes. As Guzman (2008:59) summarizes, “A formal treaty represents the most serious form of commitment not because it is more costly to enter into than other commitments but because it is understood to be a maximal pledge of reputation.” International law thus makes commitments more credible, and by extension compliance more likely, by raising the reputational consequences of noncompliance.

Reputational arguments are not without their critics. There are several key requirements for reputation to function as a compliance enhancing device for international legal commitments (Guzman 2008:91-100). Actors must be able to assess the payoffs to others from complying with or violating a commitment; observe the actual behavior of others; reliably distinguish between what counts as compliance and defection; and, perhaps most importantly, use past behavior as a basis for evaluating future trustworthiness. Such criteria are most likely to be achieved in small settings involving actors with similar beliefs and backgrounds, which is far from the reality for most groups of states in the current international system (Scott and Stephan 2006:89 fn.8).
Despite some of the problems posed by trying to apply reputational arguments to the heterogeneous international community, some of the largest criticisms are launched against the final criterion linking past actions to current beliefs. A number of studies suggest that states do not readily use prior behavior to assess the credibility of current or future commitments, and thus reputational concerns have few consistent effects on the decision making of states (Press 2005:8-9; Huth and Russett 1984; Mercer 1996:6-8). Earlier violations may have been a result of events beyond the control of the state with nothing to do with their true willingness to comply, and thus have no real impact on their future reputation as a trustworthy partner (Goldsmith and Posner 2005:102). Moreover, the interdependence of actions and beliefs emphasized by reputational accounts may be exaggerated, since it is far from clear that past behavior in one issue automatically transfers over to a state’s reputation for cooperation in different policy domains (Posner and Sykes 2013:33; Goldsmith and Posner 2005:102; Downs and Jones 2002). Rather than speak of a single reputation for cooperation, states may instead have many diverse, fragmented reputations, which together undermine any notion that international law can generate consistent reputational consequences.

Even when accepting that reputations may play a role, it does not necessarily follow that developing a reputation for complying with international legal commitments will be the only, or even primary, reputation with which a state will be concerned (Keohane 1997:497). Proponents may technically be correct that “[i]n an interdependent and interconnected world, a reputation for reliability matters” (Chayes and Chayes 2009:230). However, states may also value to a greater extent other reputations, such as those for resolve (Huth 1997), honesty (Sartori 2005), or even irrationality (Abreu and Gul 2000). The choices required to enhance these competing forms of
reputation may come into conflict or even be contradictory with desires to uphold legal commitments.

Part of the difficulty in sorting through such divergent claims is that empirical tests of reputation models in international relations are relatively rare (Mercer 1996:32). Challenges for evaluating reputational claims in research on international law are especially acute. In the first place, neither the design nor decision to join agreements are random, but rather strategic decisions in their own right by states weighing the costs and benefits of committing to a given treaty (Goodliffe and Hawkins 2006:363-4). Just as selection effects pervade the general study of the consequences of international law (von Stein 2005), a parallel set of pitfalls endures when evaluating specific mechanisms such as reputation. Self-selection dynamics means that it may be difficult to distinguish with any reasonable confidence whether the decline in a state’s reputation was due to noncompliance with a prior legal commitment, or rather other baseline traits that led the country to violate the agreement in the first place.

Similar strategic processes likely confound the compliance or noncompliance decisions of states. As proponents of the managerial school note, a certain share of violations may be involuntary in the sense that the country does not have the actual capabilities to comply with the terms of the agreement (Chayes and Chayes 1993:193-5). In such instances, it would be unfair to tarnish the reputation for cooperation of such states, since violations were not due to some inherent disposition of the actor towards violations (which reputation is intended to capture), but rather a sincere lack of resources. If anything, recommendations would be to encourage even closer partnerships with violations as a problem to be managed; only by engaging with such involuntary defectors and improving their capacities could cooperation and mutual benefits be achieved.
In many cases, however, the decision to comply with, or violate, a legal commitment is a strategic decision where states likely take into account all of the likely costs and benefits, reputational or otherwise. Observing the absence of reputational consequences after violations does not necessarily negate the role of reputation, since such states may have strategically selected to only renege on those commitments where the reputational stakes could be most minimized. On the flipside, it does not automatically follow that observing compliance is proof that the threat of reputational sanctions is hard at work. Louis Henkin’s famous pronouncement on the propensity of states to abide with their promises may factually be correct—“Almost all nations observe almost all principles of international law and almost all of their obligation almost all of the time” (Henkin 1979:47). Yet the reasons for such patterns of compliance are varied, including habit, normative attachment, external coercion, or only choosing commitments an actor would comply with in any event, which may have little or nothing to do with reputational concerns.

One point largely shared by both sympathizers and critics of reputational accounts of international law is the dearth of systematic empirical research on this question (Simmons 2010; Goldsmith and Posner 2005:103). Although Guzman offers a nuanced theory of the factors affecting reputation formation and change vis-à-vis international law, he acknowledges that he offers “no way to provide an estimate of the magnitude of reputational sanctions” (Guzman 2008:85). Posner and Sykes (2013:33) instead outline a more critical view of the reputational function of international law, but similarly in the end admit, “No one really knows how important reputation is for states, however, and thus how much it accounts for compliance with international law.”

---

3 The line dividing voluntary and involuntary violations can be somewhat blurred, since states may have incentives to publically claim they are unable to comply with a particular commitment in order to avoid the costs of reneging (reputational or otherwise), or to obtain a better deal.
In the remainder of this paper I seek to offer a more reliable test of the reputational consequences of international law. Because reputation is fundamentally about the beliefs of others regarding a country’s propensity to abide by its promises, I examine public attitudes toward the trustworthiness of a foreign country and, in turn, support for engaging in future cooperation. Studying the public has become a fruitful area of inquiry, as scholars increasingly identify the role of domestic factors for compliance in international relations (Simmons 2009; Dai 2005; Gaubatz 1996; McGillivray and Smith 2000; Lipson 2003). Several recent studies have found that international law can transform the beliefs of domestic publics and their support across a range of policies (Tomz 2008; Wallace 2013). While this research shows that international law shapes the beliefs of citizens, the findings only offer indirect support for claims that reputation plays a contributing and important role. The next section outlines a more direct assessment of how international law alters the reputational stakes involved in the eyes of the public.

**Data and Method**

I designed and fielded a survey experiment involving the U.S. public to evaluate the reputational consequences of legal commitments. Randomization alleviates the problems posed by selection effects and nonrandom assignment afflicting most observational studies of international law. Random assignment ensures treatment and control groups are similar on average across all observed and unobserved baseline characteristics, thereby increasing the confidence in any resulting causal inferences concerning the effects of international law. Querying mass publics offers the additional benefit of directly evaluating domestic sources for the reputational effects of international law.
The survey was conducted through Time-sharing Experiments for the Social Sciences (TESS).\(^4\) The survey was administered by the research firm Knowledge Networks (KN) and fielded from December 8 through December 19, 2011 to a random sample of 4,781 American adults drawn from the KN panel.\(^5\) Of those invited to participate, 2,929 agreed to do so, producing a final completion rate of 61.3%. KN offers a probability-based panel that covers the entire online and offline U.S. population aged eighteen years or older. The panel thus provides the opportunity to field online surveys to a nationally representative sample of the U.S. population.\(^6\)

Because reputations are based on other actors’ beliefs, the survey examines U.S. public attitudes toward a foreign country.\(^7\) The survey refers to a hypothetical country to ensure respondents are not influenced by idiosyncrasies concerning any particular country or region. After a general introductory prompt, all respondents are presented with some form of scenario involving a foreign country’s degree of commitment to a policy and its subsequent conduct. The experimental component of the survey involves randomly providing respondents with additional pieces of information regarding the scenario based on three separate treatments.

The first treatment, Issue Area, concerns the specific foreign policy issue involved to gauge whether international law generates different reputational consequences depending on the type of stakes involved. Issue Area is a binary treatment referring either to a security or economic scenario. The security scenario involves a contested island in another region where

---

\(^4\) Data collected by Time-sharing Experiments for the Social Sciences, NSF Grant 0818839, Jeremy Freese and Penny Visser, Principal Investigators.

\(^5\) Knowledge Networks is now part of the company GfK.

\(^6\) Participants are provided Internet access if necessary. For further details on KN’s sampling, recruitment, and fielding methods, see http://www.knowledgenetworks.com/ganp/irbsupport/.

\(^7\) To ease the discussion of the design and results, I describe the country as foreign in order to clearly distinguish it from the respondent’s home country of the United States. However, in the actual survey I refer to countries “in another region of the world” to avoid any negative connotations associated with the term “foreign,” which may bias responses against the country in question.
natural resources were recently discovered, and with the potential for states to use military force to resolve the dispute. The scenario focuses on territory to highlight the high stakes under contention, since such disputes tend to be the most dangerous and likely to escalate militarily (Huth 1996). The frequency of territorial disagreements also contributes to making the situation more realistic compared to a less common or completely abstract scenario. The use of an island dispute also lessens the likelihood respondents will attribute the scenario to a specific country or region, since similar disputes had previously, or recently, took place across several regions, including the South China Sea, the Middle East, and Latin America at the time the survey was fielded.8

The economic scenario, in turn, deals with trade relations between countries. The scenario goes on to note that countries in the region were facing economic problems, which increased the risks of raising trade barriers against outside products. The use of a general trade scenario was similarly chosen to guard against respondents attributing the situation to a particular country or region, since commercial relations and disputes are fairly commonplace especially given the rising membership in the World Trade Organization and the rapid spread of regional trade agreements (Mansfield and Reinhardt 2003).

Although dealing with different issue areas, all other elements in the general design of the two types of scenarios were kept as similar as possible to make sure that any differences in responses were due to the issue involved rather than any other possible distinctive features. Both scenarios involve a hypothetical country in another region of the world that finds itself in a situation where it may have incentives to engage in aggressive behavior (either military or

---

8 To assess whether the scenario led to an emphasis on a particular region which could bias responses, I asked participants in a prior pilot study which region they believed most closely resembled the scenario presented. The results indicated that no particular region predominated. Individual participant beliefs about the likely real-world region also did not significantly affect any of the subsequent results concerning the effects of international law.
economic). Both scenarios also do not directly involve the United States as a direct participant in either the territorial dispute or economic crisis, meaning respondents are assessing the foreign country reputation based on its commitments and conduct toward other states in the region.\(^9\)

The second treatment, *Type of Commitment*, deals with whether the foreign state made a commitment not to pursue a given policy depending on the specific scenario. For the security scenario, the pledge was not to resort to military force to take control of the contested island, while in the economic situation the promise was to refrain from raising trade barriers against outside products. Both scenarios thus involve a “negative” commitment, in other words promises *not* to take particular actions, as opposed to a “positive” commitment where a country pledges to engage in a given behavior. Holding constant the negative nature of the promises across both scenarios ensures that similar types of commitments are being evaluated.

Both positive and negative types of commitments are common across a variety of international agreements. For instance, military alliances can take the form of defensive alliances involving a positive commitment to come to the aid of a partner under attack, as opposed to nonaggression pacts that are more akin to a negative commitment where a country pledges not to attack other signatories. Similarly, provisions in human rights treaties dealing with physical integrity rights often involve promises not to resort to particular actions, such as in the Convention against Torture (CAT). On the other hand, the various rights contained under the International Covenant on Economic, Social and Cultural Rights (ICESCR) promote commitments for states to take active steps to improve health, labor, and educational standards.

\(^9\) Of course, an additional possible treatment of interest would be to examine whether the reputational consequence of international law are greater for legal commitments that a foreign country made directly with the respondents’ own country, in this case the United States, versus those presented in the scenarios in this instrument which involve the United States simply as an observing third party. This would be consistent with findings from psychological research showing that past events an actor experiences firsthand have a more significant impact than those experienced vicariously (Jervis 1976:239-46).
Each of these broad categories of commitments is often evident even in the same treaty. The 1949 Geneva Conventions proscribe the use of a wide range of violent tactics against vulnerable groups like prisoners and civilians, but also require states and their armed forces to take a variety of active steps to improve their military training and conduct in the midst of war. Similarly, the General Agreement on Tariffs and Trade (GATT) instituted the requirement for members to grant “general most-favoured-nation treatment” to all other contracting parties (Article 1), along with promises not to implement a variety of protectionist policies, such as quotas and other quantitative restrictions (Article 11).\(^\text{10}\)

Positive and negative commitments may have different reputational implications, since compliance with the former usually necessitates a country taking some observable actions, while the latter instead requires a country to refrain from engaging in a particular deed. Nevertheless, existing research suggests that examining negative commitments should provide a fairly conservative estimate of the general consequences of violating international legal promises. A study in the related area of crisis bargaining finds that the costs to leaders of engaging in behavior inconsistent with their prior promises are significantly greater for breaking positive commitments compared to negative commitments (Levy et al. 2014). Of course, examining the impact of positive versus negative commitments would be an interesting question in its own right, but is beyond the scope of the present study.

The Type of Commitment treatment involves three separate experimental conditions: the control group involving no commitment, a soft law treatment, and a separate hard law

\(^{10}\)Moreover, the line between positive and negative commitments can sometimes become blurred, and depends on what conditions or actions are considered the status quo. For instance, GATT Article 11 on eliminating quantitative restrictions contains both a positive requirement to reduce and eventually eliminate existing restrictions, but also the commitment noted above not to institute or raise any new quotas or comparable protections.
treatment. Groups receiving the no commitment condition were told the foreign country had not previously made any international agreement dealing with promises either over the use of force for the security scenario, or trade barriers in the economic scenario. I use the generic term “international agreement” so as not to lead respondents to infer the country refused to join a specific type of legal agreement, especially since the level of legalization is varied in the other two treatments. The no commitment condition is also considered the control group given that the corollary in most observational studies of international law, especially those dealing with reputation, is to compare the behavior of non-ratifiers to those who ratified the relevant agreement or agreements (Simmons 2000; Hathaway 2002; Hafner-Burton and Tsutsui 2005; Neumayer 2005).

By contrast, groups receiving the hard law treatment condition were told the country “had previously signed an international treaty where it made a formal promise” not to use military force, or raise trade barriers, depending on the issue area in the scenario. International treaties are expected to raise the stakes for a country’s reputation, meaning the reputational consequences should be greater for a hard law commitment compared to the absence of any commitment.

11 In a follow-up study, I also included an additional more pure baseline control, where subjects in this group were given no prompt regarding the prior commitment decision of the foreign country. Results showed little difference in the responses for those in this purer control group versus those subject told the foreign country had made no prior legal commitment. Of course, this does not necessarily imply that respondents in the pure control group assumed the foreign country had made no prior commitment. Nevertheless, it does suggest that for the purposes of determining the consequences of soft or hard law commitments that the precise choice of control group (no commitment versus no prompt at all) should not substantively change the results.

12 Technically, signing an international treaty generally does not carry the same legal obligation as formally ratifying or acceding to the agreement, though signatories are still expected to refrain from attempting to defeat the object and purpose of the treaty. See Article 18 of the 1969 Vienna Convention on the Law of Treaties, available at <http://untreaty.un.org/ilc/texts/instruments/english/conventions/1_1_1969.pdf>. However, one concern was that using more technical terms, such as “ratify,” may have led to greater confusion on the part of respondents. By comparison, signature is more straightforward yet still carries the desired connotation of making a legal commitment to the agreement. If anything, the current phrasing would suggest an underestimate of the hard law treatment effect, should some respondents infer a weaker level of commitment for signing a treaty compared to formal ratification. Nevertheless, to ensure that the particular language used to describe the commitment to the treaty was not driving responses, I conducted a subsequent study that used the original “signed” term, as well as a more generic “joined” alternative, and the results did not significantly change.
To further assess whether differences in the level of legalization affects the reputational consequences involved, a second soft law treatment condition was also included. The soft law treatment varies the level of obligation in the agreement, where respondents in this group were told the foreign country “had taken part in a regional meeting that created guidelines recommending countries” either not use military force, or raise trade barriers, depending on the issue area for the scenario. Although both treatments involve a form of legal commitment, the expectation is that the higher level of obligation in the hard law treatment should generate greater reputational consequences than the less binding recommended guidelines of the soft law prompt (Abbott and Snidal 2000). I chose to focus on obligation because this remains the most closely studied dimension of legalization and the one with the greatest implications for the reputational consequences of international law (Morrow 2000; Simmons 2010). Of course, future work could examine the reputational consequences of the other legalization dimensions involving the level of precision or delegation in an international commitment (Abbott et al. 2000).

The third and final treatment, Compliance, is binary and concerns how the country subsequently acted with regards to its type of commitment in the relevant security or economic scenario. Respondents receiving the violator treatment were told the country either decided to use military force, or raise trade barriers, depending on the type of scenario. Those receiving the compliance control were instead told that the country did not use military force, or did not raise trade barriers. Technically speaking, for the Type of Commitment control group, the foreign country could not violate or comply since in this experimental condition the country had not made any previous type of legal promise. Nevertheless, the terms violator and complier are used

---

13 One additional concern is that any differences in responses between the hard law and soft law treatments may not be due to the level of obligation in the commitment, but rather the level at which the agreement was negotiated – international for the treaty and regional for the recommended guidelines. However, a subsequent survey conducted found that referring to the recommended guidelines as arising from an “international,” as opposed to “regional,” meeting does not significantly change the results.
here as a shorthand to refer to the specific subsequent behavior in which the country engaged. The actual instrument presented to survey participants only describes the decision later taken by the foreign country (e.g. raising or not raising trade barriers), and does not employ specific terms such as “violate” or “uphold” that could unduly bias responses.

The primary theoretical expectation is that countries that violate a legal commitment should suffer greater reputational costs than if they engaged in the same behavior in the absence of a prior commitment. The reputational consequences are also expected to be greater for the more formal hard law treaty commitment compared to the less binding soft law promise. Taken together, including three separate treatments, two binary (Issue Area and Compliance) and one trichotomous (Type of Commitment), implies a three-way factorial design (2 x 3 x 2) with twelve total experimental conditions after all treatments are fully crossed. The complete set of experimental conditions with the number of respondents in each group is summarized in Table 1 below. The full text for the instrument is included in supplementary appendices.

[Table 1 about here]

After reading the particular scenario and additional prompts, respondents are then asked to answer a brief series of questions regarding support for future cooperation with the foreign country. Since reputational arguments on international law primarily focus on other actors’ expectations of a country’s future prospects for compliance and thus desires for further cooperation, the dependent variables concern public support for concluding an international agreement with the foreign country. For the security scenario, respondents are told “The United States is currently thinking about signing an agreement with this country that would involve cooperation over military issues.” Respondents are then asked to indicate how much they would support or oppose the United States concluding a military agreement with the foreign country on
a five-point Likert scale ranging from strongly oppose to strongly support. Those receiving the economic treatment are instead first told “The United States is currently thinking about signing an agreement with this country that would involve cooperation over trade issues,” and then asked how much they would support or oppose concluding a trade agreement with the country. The type of future cooperation was thus chosen to reflect either the security or economic nature of the relevant scenario. The question wording also refers to a generic “agreement” so that respondents are not necessarily cued in supporting or opposing a particular type of future legal commitment, though this could also be an interesting element to investigate in future research.

The level of public support for any future cooperation is thus expected to vary systematically based on each of the three treatments related to the issue area under dispute, the type of prior commitment made, and the subsequent compliance behavior of the foreign country. In order to assess the existence of any spillover effects for the foreign country’s reputation, respondents were asked a follow up question regarding their support for concluding a future agreement in the other issue area not covered by their scenario. This means that respondents in the security scenario were then asked how much they would support or oppose a trade agreement between the United States and the foreign country, while those in the economic scenario were asked to indicate their level of support for a future military agreement.

Because respondents were randomly assigned to one of the twelve conditions containing different versions of each of the three treatments, the experimental design increases the likelihood that each group differs on average only with respect to the relevant treatments. As expected, tests indicate that groups were comparable across all observed characteristics that

---

14 The full set of answer options are as follows: Strongly oppose / Somewhat oppose / Neither support nor oppose / Somewhat support / Strongly support.
could likely affect individual support for future military or economic cooperation.\(^\text{15}\) The nature of KN’s sampling procedures also ensures that the empirical results are unlikely to be a function of any peculiarities in the sample, but are rather fairly representative of attitudes held across the wider U.S. public.\(^\text{16}\)

**Empirical Results**

Turning to the analysis, random assignment alleviates the need to employ complex statistical models using a battery of controls that characterizes much of the observational research on international law. In what follows, I thus choose to report relatively straightforward cross-tabulations and difference-in-means tests. Nevertheless, further analysis reveals the results are robust to including standard socio-demographic controls commonly associated with individual foreign policy preferences, and are available from the author upon request.

To provide an initial sense of patterns of support for international cooperation, Table 2 below illustrates the general effects for the issue area and compliance treatments, aggregated across all of the legal commitment experimental conditions. The figures report the percentage of respondents supporting future U.S. cooperation on a military agreement for the security issue area, and for a trade agreement in the economic scenario, by each version of the compliance condition.\(^\text{17}\) 95\% confidence intervals are included in parentheses. Across all conditions respondents are less supportive of U.S. cooperation over military affairs compared to economic

\(^{15}\) I conducted an extensive series of balance tests comparing the distribution across the relevant treatment and control groups for all available baseline covariates that likely affect foreign policy preferences. Full results available in supplementary appendices.

\(^{16}\) To check that this was indeed the case, I compared various socio-demographic characteristics across respondents in the sample to the U.S. population. For a reliable point of comparison, I employed the December 2011 update for the Current Population Survey (CPS), which is the same month that the experimental survey was fielded. The sample was largely consistent with the CPS benchmarks with an average deviation of only around 3%.

\(^{17}\) Percentages are calculated by combining responses across the two levels of support (strongly support/somewhat support) and opposition respectively. Treatment effects do not change substantially when including respondents who neither supported nor opposed cooperation with the relevant type of international agreement. All analysis conducted using Stata 13.
affairs, and the effect is statistically significant. The general difference in support for cooperation across issue areas likely reflects greater public concerns over sovereignty and security interests involved in matters of high politics compared to more regular commercial relations. Similarly, respondents do appear to use past behavior as a guide for assessing future conduct and consequent support for future cooperation. Violations by foreign countries are much less likely to garner public support for future U.S. cooperation compared to countries that complied with past promises. The size of the difference between compliers and violators is larger in the security scenario (around 30%) compared to the economic scenario (around 20%), but in both cases the effect is statistically significant.

[Table 2 about here]

Although both issue area and compliance have substantial overall effects, the main question of concern for reputational theories of international law is whether and to what extent the specific type of legal commitment made by states plays affects attitudes toward future international cooperation. Figures 1 and 2 below report the reputational consequences of legal commitments in the security and economic scenarios respectively. Figure 1 measures the percentage point change in U.S. public support for future military cooperation with the foreign country resulting from a hard or soft law commitment not to resort to military force over the contested island outlined in the security scenario, relative to the control group baseline (where the country had not made any prior legal commitment). Negative values refer to a reputational cost for the foreign country in terms of reduced willingness of the U.S. public to favor future cooperation, while positive values suggest a reputational benefit of enhanced cooperative

---

18 Of course, economic relations can often have fundamental consequences for a country’s security position (Kirshner 1995; Hirschman 1980:16). Nevertheless, the results here suggest the public is much more concerned about the direct security implications involved in military cooperation compared to economic cooperation, which leads to lower overall levels of support for the former.
possibilities down the road. 95% confidence intervals are indicated by vertical lines for each point estimate.

[Figure 1 about here]

[Figure 2 about here]

Looking first at those who received the violator prompt (that the foreign country eventually decided to use military force over the islands) confirms that legalized commitments do indeed carry greater reputational stakes. Respondents told the foreign country had violated a hard law commitment previously made through a formal treaty were over 15% less likely to favor future military cooperation compared to the situation where the country resorted to military force in a similar manner, but had made no previous legal commitment promising not to do so. The effect of hard law on violators was also significant at the 95% confidence interval does not include zero, or no apparent effect. Treaty commitments are thus far from costless, but instead raise the subsequent costs of noncompliance relative to the decision of failing to make any prior legal promises.

However, the results suggest that not all legal commitments are treated equal in the eyes of the public, where the level of legalization shapes the relative reputational stakes involved. Foreign countries that violate the softer law non-binding guidelines recommending the non-use of military force also suffer a decline in public support for future military cooperation, but the effect is much smaller (around 5%) and fails to be significantly different from zero. Legal commitments exhibiting a lower level of obligation thus do not appear to generate the same reputational consequences. The decline in U.S. public support for military cooperation with a foreign state that violates a soft law commitment is statistically indistinguishable from the control condition where the country engaged in the exact same behavior but without making any
prior legal commitment. Although the confidence intervals for the hard law and soft law treatments in the violator condition overlap somewhat, the difference is nonetheless statistically significant and point to hard law commitments generating the greatest reputational costs compared to promises involving lower levels of legal obligation.

Looking at the treatment effects under the compliance condition (where the foreign country chose not to use military force over the island) reveals that foreign countries are unlikely to gain much of a corresponding reputational premium by adhering to their commitments. Public support for future military cooperation does rise when a foreign country complies in both the hard and soft law treatments, yet neither effect is statistically significant. The effect for the hard law (9%) treatment remains larger than the soft law treatment (4%), yet both are smaller in substantive terms compared to the corresponding values from the violation condition and neither is distinguishable from zero. The lack of reputational benefits from compliance may be a function of the public’s baseline expectation that countries generally comply with the promises they make. The foreign country’s adherence to a prior commitment not to use military force may provide little additional information to the U.S. public regarding the country’s likely propensity to comply with future agreements, helping to account for the absence of any strong increase in support for further military cooperation. Moreover, even though the experimental design randomizes the foreign country’s (non)compliance with the use of military force, respondents receiving the compliance condition may still be attributing certain self-selection dynamics to the intentions, and ultimately the decision making, of the foreign country. Since past work argues that states tend to only join agreements with which they intend to comply (Downs et al. 1996), then the U.S. public may be unfazed when observing compliant behavior and thus reluctant to reward the country with a more trustworthy reputation.
Taken together, the results point to the primary reputational function of international law as raising the ex post costs of non-compliance, while complying with existing agreements appears to offer few tangible reputational benefits. However, since many international agreements are formed primarily to address concerns over cheating and defection, the experimental findings provide strong micro-level support for the function of international agreements, especially binding treaties, to raise the reputational stakes for parties and make violations less attractive.

Figure 2 analyzes the reputational consequences of international law in the case of public support for trade cooperation based on the economic scenario where the foreign country was deciding whether or not to raise trade barriers against outside products due to economic woes at home. As in the security scenario, legal commitments influence U.S. public support for future international cooperation, yet shifting to an economic scenario shows some important differences based on the issues under dispute. The reputational costs of violating an economic legal commitment (raising trade barriers) on public support for future trade cooperation are magnified compared to the security scenario. Both the hard law (25%) and soft law (17%) treatment effects for violators indicate a decline in public support for future trade cooperation relative to the control group (no prior legal commitment), and are much larger compared to the corresponding values from the security scenario. Furthermore, the soft law treatment in the economic scenario is now also statistically distinguishable from zero, and similar in size to the hard law effect from the security scenario.

The greater reputational costs for violating economic commitments compared to military assurances on support for cooperation in each respective issue area perhaps reflects an understanding that security concerns are more likely to trump legal commitments when matters
of high politics are under dispute. With likely lower baseline expectations for compliance over security-related issues, the reputational costs for violations over the use of force are correspondingly less severe. By contrast, the likely greater expectations of compliance with economic commitments means that more information is revealed about a country’s willingness to abide by its legal promises, which consequently has more significant implications for its reputation as a trustworthy commercial partner. The earlier results from Table 2 offer some indirect evidence that prior expectations of compliance may shape subsequent reputational costs. The generally higher level of public support for cooperation over economic matters compared to security relations is consistent with the notion that the public may anticipate a lower risk of defections by foreign partners on economic commitments, though other unrelated motives may also be present. The greater reputational stakes involved in the commercial scenario provide a further rationale for why economic relations are amongst the most legalized areas of international law, especially over questions dealing with trade (Abbott 2000; Goldstein and Martin 2000).

Despite the difference in the size of reputational costs for violations, complying with economic legal commitments yields similarly lackluster reputational benefits to those found in the security scenario. While the hard and soft law treatments are positive in the compliance condition relative to the control group (no legal commitment against raising trade barriers), both are relatively small (6% and 4% respectively) and statistically indistinguishable from zero. The results from the economic scenario confirm that the main function of international law for a state’s reputation is to increase the ex post costs of defection, while offering few tangible reputational benefits from complying with prior legal commitments.

---

19 As noted earlier, the U.S. public may be less supportive of cooperation involving security issues because of greater worries over the sovereignty of the United States on matters of national security irrespective of the baseline expectations of compliance for agreements in either sphere of activity.
The analysis thus far has focused on support for future cooperative agreements within each policy sphere – military cooperation for the security scenario, and trade cooperation for the economic scenario. Figures 3 and 4 below assess whether commitments and compliance in one issue area can spill over and affect a country’s reputation in another issue area. This is a question where critics are most skeptical that commitments in one issue can generate reputational consequences on the credibility of unrelated policy promises (Downs and Jones 2002). Yet evaluating spillover effects with standard observational data is even more daunting given the number of possible intervening factors and inferential obstacles. The experimental design employed here thus offers a preliminary, but more reliable, test for reputational spillover effects resulting from international legal commitments. Results are based on examining the follow-up question asked to each respondent, which examines support for future cooperation in the other policy scenario – in other words, support for trade cooperation for those receiving the security scenario; and support for a military agreement from those exposed to the economic prompt.

[Figure 3 about here]

[Figure 4 about here]

In a similar manner to the prior figures, Figure 3 reports the percent change in U.S. public support for future military cooperation based on the prior type of legal commitment and compliance behavior of the foreign country in the economic scenario. The figure thus tests whether commitment and compliance decisions in the economic realm have any impact on U.S. public support for military cooperation. The results suggest that economic relations have little effect on public attitudes toward agreements involving matters more closely related to national security. A foreign country that raises trade barriers is slightly less likely to garner U.S. public support for a future military agreement when violating a hard law commitment (5%) compared to
a soft law promise (4%), but neither effect is statistically different from the control group baseline of no prior legal commitment of any form. A complying country that refuses to raise barriers does receive a more marked boost in support for military cooperation from a hard law commitment (just under 8%), while the soft law treatment actually has a small negative impact (3%), though again neither treatment effect is statistically significant.

When assessing the value of an international military agreement, the U.S. public appears to place greater weight on a foreign country’s past commitments that bore more directly on security affairs, while largely discounting any behavior related to past promises involving economic relations. Violating economic commitments can indeed generate a substantial impact on public support for future cooperation, yet these reputational costs are limited to agreements sharing similar economic concerns, while attitudes toward military cooperation remain relatively unchanged. The results from Figure 3 seem to confirm the claims made by some skeptics that, even if international law does heighten reputational stakes, those consequences are unlikely to carry over to other issue areas (Goldsmith and Posner 2005:102). States may rest a little easier knowing that although they may suffer some reputational losses from violating legal commitments, those costs will at least be limited to the issue area at hand.

Figure 4 shows that while spillover effects may not be universal, they can be quite substantial depending on the circumstances. Figure 4 reports the percent change in U.S. public support for future trade cooperation based on the prior type of legal commitment and compliance behavior of the foreign country in the security scenario. The results show that violating a hard law security commitment not only has an immediate effect on U.S. public support for a future military agreement, but the consequences for the foreign country are even more dire for future possible trade cooperation. Violators of the hard law treatment in the security scenario
involving the use of force over a contested island incur more than twice the percentage drop in support for a future trade agreement (33%) compared to the original direct decline in support for military cooperation (16%). The reputational spillover effects are in more impressive when viewed in relative terms – violating a hard law security commitment generates a 50 percent change in public support for trade cooperation compared to a situation where the violating country engaged in the exact same behavior but without having made any prior legal commitment.\textsuperscript{20} As expected, violating a soft law security commitment has a smaller reputational carryover for support for future trade cooperation (just under 10%), though the effect just fails to be statistically indistinguishable from the control group (no prior security commitment). By contrast, complying with a security commitment has little benefits for future prospects of economic cooperation, which is similar to the original direct effects on military cooperation. The hard law and soft law treatment effects for the compliance condition where the country did not resort to military force are minimal (both 2%) and neither is statistically significant relative to the control group.

Nevertheless, the experimental results from Figure 4 show that domestic publics are highly influenced by the commitment and compliance decisions made by foreign states in matters concerning national security. The large spillover effects from security commitments into support for economic cooperation may be explained by the signals such conduct communicates to domestic public regarding the trustworthiness of a foreign country. Even if expectations of compliance may be lower for security commitments, a country showing itself to willingly renege

\textsuperscript{20} The absolute percent support for trade cooperation with a foreign country that resorted to force in the military scenario, yet had made no prior legal commitment, was 71%. By contrast, the equivalent level of support for a country that employed force over the island, but had signed a treaty promising to refrain from resorting to military actions, was just 37%. The relative reduction of around half in the public’s willingness to support future trade cooperation with the foreign country thus shows the significant effect of reputational spillover effects from violating security commitments.
on a promise involving national security may then be viewed as even more prone to violating commitments involving lower politics, such as trade. Because security issues are often more salient in the public’s eye, the prospects for reputational spillover effects may be greater for security commitments compared to economic promises even if the latter have more pronounced within-issue consequences.

The impact of reputation on public attitudes toward international cooperation may in general be more significant for economic affairs, evidenced by both the direct and spillover effects on future trade cooperation. Past research has shown that security factors more generally play an important role in either facilitating or inhibiting cooperation over trade and other commercial relations (Gowa 1994:6; Thacker 1999). The survey results offer a further domestic pathway through which the public expresses much stronger support for cooperation on economic fronts for states that have proven to be more reliable in keeping their legal promises in matters of national security. By contrast, support for military cooperation may instead follow a slightly more pragmatic logic, where reputation can play some role at least for hard law security commitments though with little impact from unrelated commitments.

Conclusion

The results from the analysis of a survey experiment confirm that international law can generate significant costs for violators even in the absence of a higher enforcement authority. Legal commitments raise the reputational stakes of subsequent noncompliance, where foreign violators find much less support from the U.S. domestic public for engaging in future cooperation. The findings point to both the strength and limits of reputational accounts for the operation of international law. Reputational stakes appear much larger for violating harder law commitments compared to less legally binding promises, while the consequences are more
apparent for economic cooperation than for security relations. Spillover effects are also evident though asymmetric, with security commitments having substantial implications on the prospects for future economic cooperation, but not vice versa. On the other hand, complying with more legalized commitments, either in the security or economic arenas, does not translate into appreciable surplus benefits to a country’s reputation for cooperation. The primary reputational function of international law thus appears to be to make current commitments more credible by raising the ex post costs of defection. Although firmly within the realm of international law, legal commitments contain a certain implicitly coercive element, which encourages greater compliance through the threat of damaging a country’s ability to engage in future cooperation rather than positive inducements of producing new and potentially lucrative cooperative opportunities.

The experimental design of this study also suggests several avenues for future research. While the survey instrument focused on evaluating the core claim that international law increases the reputational stakes of noncompliance, other elements of reputational theories could also be examined through experimental manipulation. The level of obligation in a commitment proved to have substantial reputational consequences, but the other legalization dimensions of precision and delegation should also be examined. In particular, the implications of delegating greater authority to international organizations continue to be debated, especially concerning the relative enforcement power of such higher authorities (Brown 2010; Mattli and Slaughter 1998; Moravcsik 1998:501). Even assuming the direct enforcement powers of higher international authorities are far from absolute, violating commitments involving higher levels of delegation may be expected to generate greater reputational consequences. International institutions may
thus not operate simply as a top-down source of enforcement, but rather complement existing horizontal sources encouraging compliance.

The basic experimental design developed in this paper may also be extended to assess other related claims concerning reputational mechanisms. Additional possible treatments that could be investigated include the reputational consequences arising from any of the following sources: the frequency of violations; the length of time elapsed since a country last violated an agreement; whether or not the violation was voluntary or involuntary; the compliance decisions of other states faced with a similar situation; whether or not the public’s home country was directly affected by the violator’s policies; whether reputational stakes are higher for certain regime types than others; and the extent to which the reputational consequences (either positive or negative) of conduct toward past commitments are leader-specific or carry over to new leaders of a country. Much work thus remains to evaluate the development of a country’s reputation in the shadow of international law, which can offer a fuller theoretical and empirical understanding of the functioning of international legal commitments in an anarchical society.

---

21 For instance, see (Guzman 2008:91-111).
Tables and Figures

Table 1: Experimental Conditions for International Law and Reputation Survey

<table>
<thead>
<tr>
<th>Issue Area</th>
<th>Type of Commitment</th>
<th>Compliance</th>
<th>Number of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security</td>
<td>Hard Law</td>
<td>Violate</td>
<td>240</td>
</tr>
<tr>
<td>Security</td>
<td>Hard Law</td>
<td>Comply</td>
<td>237</td>
</tr>
<tr>
<td>Security</td>
<td>Soft Law</td>
<td>Violate</td>
<td>244</td>
</tr>
<tr>
<td>Security</td>
<td>Soft Law</td>
<td>Comply</td>
<td>244</td>
</tr>
<tr>
<td>Security</td>
<td>None</td>
<td>Violate</td>
<td>251</td>
</tr>
<tr>
<td>Security</td>
<td>None</td>
<td>Comply</td>
<td>245</td>
</tr>
<tr>
<td>Economic</td>
<td>Hard Law</td>
<td>Violate</td>
<td>244</td>
</tr>
<tr>
<td>Economic</td>
<td>Hard Law</td>
<td>Comply</td>
<td>242</td>
</tr>
<tr>
<td>Economic</td>
<td>Soft Law</td>
<td>Violate</td>
<td>243</td>
</tr>
<tr>
<td>Economic</td>
<td>Soft Law</td>
<td>Comply</td>
<td>233</td>
</tr>
<tr>
<td>Economic</td>
<td>None</td>
<td>Violate</td>
<td>245</td>
</tr>
<tr>
<td>Economic</td>
<td>None</td>
<td>Comply</td>
<td>261</td>
</tr>
</tbody>
</table>

Table 2: General Effects of Issue Area and Compliance Treatments on Support for International Cooperation

<table>
<thead>
<tr>
<th>Issue Area</th>
<th>Comply</th>
<th>Violate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security</td>
<td>66 (61, 70)</td>
<td>35 (31, 40)</td>
</tr>
<tr>
<td>Economic</td>
<td>74 (70, 78)</td>
<td>53 (48, 57)</td>
</tr>
</tbody>
</table>

Note: 95% confidence intervals in parentheses. Figure refers to support for a military agreement with foreign country in the security scenario, and support for a trade agreement in the economic scenario.
Figure 1: Effect of Legal Commitment in Security Scenario on Support for Military Cooperation

Y-axis measures the first difference between the legal commitment treatment (soft law or hard law) and control group (no commitment) on the probability of reporting any level of support for future cooperation with the foreign country. First differences for legal commitment are calculated for situations where the foreign country either violated or complied with the commitment.

Figure 2: Effect of Legal Commitment in Economic Scenario on Support for Trade Cooperation

Y-axis measures the first difference between the legal commitment treatment (soft law or hard law) and control group (no commitment) on the probability of reporting any level of support for future cooperation with the foreign country. First differences for legal commitment are calculated for situations where the foreign country either violated or complied with the commitment.
Figure 3: Spillover Effect of Legal Commitment in Economic Scenario on Support for Military Cooperation

Y-axis measures the first difference between the legal commitment treatment (soft law or hard law) and control group (no commitment) on the probability of reporting any level of support for future cooperation with the foreign country. First differences for legal commitment are calculated for situations where the foreign country either violated or complied with the commitment.

Figure 4: Spillover Effect of Legal Commitment in Security Scenario on Support for Trade Cooperation

Y-axis measures the first difference between the legal commitment treatment (soft law or hard law) and control group (no commitment) on the probability of reporting any level of support for future cooperation with the foreign country. First differences for legal commitment are calculated for situations where the foreign country either violated or complied with the commitment.
References


