The limits of reputation on compliance[1]

RACHEL BREWSTER*

Assistant Professor, Harvard Law School, Cambridge, MA 02138, USA

Introduction

Scholarly work at the intersection of international relations theory and international law (IL) has been going on for over a decade now. Of the many strands that have emerged from this dialogue, the most popular discussion concerns 'rational' accounts of how IL may influence state action. Andrew Guzman's recent book, *How International Law Works*, is an important contribution to this discussion. Guzman integrates international relations and IL theory to explain why and how IL matters to states. Unlike Jack Goldsmith and Eric Posner's book, *The Limits of International Law*, which accepts the idea of rational states but finds that the ultimate effect of IL is minimal, Guzman argues that IL makes a notable difference, primarily in the functioning of the state's reputation. Only states with good reputations receive the benefits of joint cooperation, he argues, and law is the means by which states invoke reputational collateral. The use of reputation to explain compliance is not new in the IL field, but it possesses particular significance in Guzman's account. Unlike other IL theorists, Guzman views reputation as a major, if not the primary, cause of states' decisions to follow IL.

Guzman argues that reputation can pull states toward compliance when the *real politick* tools of reciprocity or retaliation are insufficient. Viewing reputation as an important cause of compliance is appealing because it allows Guzman, and other authors in the rationalist vein, to bridge the divide between IL skeptics, who argue that states act in their own interests, and IL advocates, who are optimistic about the potential for high levels of compliance in spite of states' incentives to defect. If strong enough, reputational concerns can make any state's compliance with a wide range of international agreements pass a cost-benefit analysis.

---

[1] This essay draws on a separate article, 'Unpacking the state's reputation', *Harvard International Law Journal* (in press).

* E-mail: rbrewster@law.harvard.edu

While reputation is certainly an important consideration in international relations generally and IL specifically, Guzman may be reading too much into it. His book claims to present a theory of reputation, but leaves important gaps in the causal analysis of how the 'state's reputation' can sustain significant levels of compliance with IL. In discussing the book's approach to reputation, I focus on three elements. First, I ask to whom the relevant reputation belongs. The book maintains that the relevant reputation belongs to the state. This assignment creates a problem for causation analysis, however, because governments (who make the compliance decisions) may not internalize the state's reputation fully or consistently. Second, I examine the particular problems with using reputational sanctions to enforce IL obligations dealing with public goods, an area where the book claims that reputational sanctions should be particularly effective. Third, I discuss the methodological problems in Guzman's approach to the link between reputation and compliance. His book provides only a loose means of assessing reputational costs, even as a conceptual matter. Without a means of assessment, any claim about the power of reputation remains non-falsifiable and therefore has less theoretical force. Reputation becomes an error term that makes rationalist IL claims invariably correct. While we need not accept Goldsmith and Posner's view that IL has little influence on state action, Guzman's conception of reputation is too blunt an instrument to provide an alternative. If we are to accept the importance of reputation in explaining compliance with IL, we need a more nuanced view of how reputation works.

Whose reputation?

Reputation is important in Guzman's account because it helps the audience predict the state's future behavior. The book is clear that it is the state's reputation for compliance that is relevant, noting 'we can define a state's reputation for compliance with international law as judgments about an actor's past response to international legal obligations used to predict future compliance with such obligations' (Guzman, 2008: 73). Reputation is effective in promoting compliance because it provides the audience with information. If past actions are not particularly informative of future behavior, then reputation will be significantly less effective in promoting compliance.

The book is also explicit that the relevant reputation involves compliance with IL, not with other views of the state, such as a reputation as a 'good global citizen'. Guzman's clarity here is useful and important because the two are not identical. (Helfer, 2005) A state may have an excellent reputation for maintaining its obligations but be popularly viewed as a poor

global citizen for its failure to take on legal obligations, such as joining the International Criminal Court Charter or the Kyoto Protocol. The two reputations may even push the government in opposite directions: such as the decision to use military force abroad to halt genocide when such action would be 'illegal', because the United Nation's Security Council has failed to authorize the use of force.[2]

When discussing the effect of reputational concerns on compliance decisions, the first question to ask is whose reputation we are considering. Guzman and other IL scholars use the state as the primary unit of analysis. Guzman does not justify this choice, which is fair as it is the convention in the field. (Simmons, 2000; Hathaway, 2002, 2005; Goldsmith and Posner, 2005; Tomz, 2007) Nonetheless, such analysis needs an explanation given that our focus is on the decision-making of governments. Whether the reputation belongs to the state or to the government makes a significant difference in the causal analysis of whether a government will comply with IL.

There are two distinct points I want to make here. First, the government will not internalize fully the reputational costs and benefits to the state in its compliance calculus. This is true even if the audience (other governments and non-state actors) view the state as the relevant unit for reputational analysis. Second, if the audience expects the state's behavior to change from government to government, then the costs of defection (or the benefits of compliance) are different than what we would expect under the state model. For instance, the costs of defection are presumably discounted faster and the reputations are easier for a new government to repair if the audience does not believe that the actions of the old government are predictive of the new. Both of these effects undermine the costs of violations of IL and lessen the ability of reputational concerns to sustain high levels of compliance with ILs.

Internalizing the reputational costs to the state

Guzman's concept of reputation is centered on the state, and he models the state as if it were an individual. This means that the state is the same 'actor' over time and can control its reputation through its actions like an individual can. Using the state as the primary unit of analysis is not just an assumption to simplify the model. A government acts very differently than

---

[2] An independent report assessing the legality of the NATO (North Atlantic Treaty Organization) bombing of Serbia described the NATO governments' actions as 'illegal but legitimate'. This term makes explicit the gap between government actions that are legal under international law and government actions that are generally viewed as beneficial to the international system (Independent International Commission on Kosovo, 2000).

the idealized version of a rational (and very long-lived) individual does in ways that are important for understanding the government's motivation to comply with IL. While the state will last into the indefinite future, the government making the decisions will almost always have *shorter and varying time horizons*. Reputation thus will not have an equal influence on all states and will have a lesser influence than if we conceive of the state as an individual.

First, government's time horizons are notably shorter. Where an individual may internalize (and discount) future costs to its reputation, governments are not necessarily concerned with the future costs to the state. If the government's term expires in 2 years (or, say, the next two rounds in the prisoner's dilemma game context) then government decision-making is based on the value of the next 2 years, not all the costs that the state will bear in the future. This shortened time horizon lessens the impact of reputational costs. Although we might predict that a rational individual would comply with IL in wide range of situations, a rational government may not choose the same course of action.

Second, governments will not have *consistent* concerns about the future as an individual would. A government that anticipates interaction with international partners for years to come will value its reputation more than a government that is ending its time in office. Thus, governments may discount reputation costs less the further away the government is from an election. For instance, a US president in the first year in office may care more about the state's reputation internationally than he will in a re-election year. So even if we acknowledge that government leaders have shorter time horizons than the state, we still cannot treat reputational concerns as a constant that will have the same influence on a government throughout its time in power.

Including political parties in the model of the government does not mitigate the time horizon problem. A leader's concern for the fate of the party in future elections might extend the leader's time horizons beyond his time in office. Political parties can exist longer on the nation's stage than any individual leader, but parties must win elections as well. Like individual leaders, the desire to remain in power gives political parties shorter time horizons than an idealized state.

Focusing on the political party rather than an individual government leader might help explain why the president would not 'waste' the state's reputation in his second term in office (an end-game problem), but it does not eliminate the shorter time horizon effect. The same is true if governments are uncertain of the electoral cycle, as each government faces electoral pressure that shortens its time horizons. Many parliamentary elections, for instance, are called by the prime minister or if the current

government receives a vote of no confidence in the legislature. Thus the current government may not know when the elections will take place but knows that they will come.

We also cannot assume that reputational concerns matter more for (or had a consistent influence on) leaders who managed to hold onto power for a long time. From the perspective of history, we now know past governments' tenures, but this does not mean that the government knew their own end dates and applied long-time horizons to their reputational calculations. For instance, while Gerhard Schroeder managed to remain Chancellor of Germany for 8 years, from 1998 to 2005, this should not lead us to think of Schroeder's time horizons on an 8-year calendar. Most commentators believed (and Schroeder might as well have thought) that his government was unlikely to win re-election in 2002, although we now know that the government survived until 2005.

Similarly, reputational concerns also vary with governments that do not face electoral constraints. Even with dictatorships, the possibility of regime change exists. A dictator with a strong domestic support base may expect that he will remain in power indefinitely and thus care about the state's reputation. However, a dictator that faces a crisis of domestic opposition may face a grim and quick end and thus may have a very short-time horizon.

Reputation, therefore, is not a constant that we can add or subtract from a state's payoffs when conducting a cost benefit analysis. This is true *even if the audience believes that the relevant unit for reputational analysis is the 'state' and not the government*. A government may neither fully internalize the costs and benefits of an action to the state's reputation, nor will it place a consistent value on the state's reputation throughout its term in power. Rather, reputational analysis is highly contingent on domestic politics. Using the state as the relevant unit in reputational analyses systematically exaggerates and biases the importance of reputation in government decision-making.

The extent of reputational costs

The extent of the reputation costs and benefits that the government (and the state) face depends on the audience's beliefs about what drives state behavior. This comes in two forms. First, how does a change of government within the state influence the audience's perception of the state's likelihood of compliance with IL? Second, how does a violation in one issue area influence the audience's perception of the state's likelihood of compliance with IL in other issue areas?

Changing governments. Reputational concerns are effective in promoting compliance to the extent that the audience uses reputation to

predict the state's future actions. Guzman and others generally ignore domestic politics and assume the audience's perceptions of the state's future actions are drawn from the state's past actions. But the audience has access to more information. Audience members, at a minimum, can observe elections at a low cost, and thus, may expect changes in state behavior. The informational value of past state actions is thus less informative. (Scott and Stephan, 2006) As a consequence, the reputational costs to the state of its past government's (or new government's) actions may not be particularly high. If the audience expects a new government to act differently from the past government, then the costs of violating (and the benefits of complying with) IL to the government *and to the state* will be smaller.

To know how the audience determines the informational value of past state actions for predicting the current government's actions, we need to have an idea of how the audience views that state's domestic political system. For instance, many commentators have argued that the Bush Administration has damaged the United States' reputation for compliance with IL by violating the Geneva Conventions with regards to the Guantanamo Bay detainees. Accepting this statement as true, how much will the next administration have to invest in foreign relations to restore the United States' reputation?

If the loss of reputation is due to the audience's assessment of what the Bush Administration considered acceptable policy (such as detaining suspected terrorists at Guantanamo Bay), then the United States' reputation should be relatively easy to fix, if a new administration, which holds different views of what are reasonable policy options, is voted into office. If the loss is due to the audience's assessment of the range of policies that the American electorate will support, then the United States' reputation should be harder to fix, but repairable over time if voters consistently elect leaders with more moderate policy views. Of course, the loss of reputation could be due to the audience's assessment of the structure of the government, namely that the executive has a tremendous amount of discretion in international affairs, especially in wartime. Here, the reputational loss might be impossible to repair in the short term. The United States' reputation might only be repaired by a change to the structure of the constitutional system or a long history of consistent executive practice. Interestingly, where the reputation truly belongs to the state, rather than the government, reputational concerns have *less* influence on the decision-maker because the state's reputation is relatively immutable. (Tomz, 2007) Here, the state's reputation is likely an indicator, not a cause, of compliance.

Issue areas. The reputational costs a state faces from violations of IL also depend on the audience's perception of how broadly to view a violation

of IL. For instance, is the violation of a trade agreement relevant information regarding the state's likely compliance with a security treaty? This point was previously made by Downs and Jones (2002), and Guzman addresses this issue at length. While not fully responding to Guzman's analysis on this point here, I want to raise briefly how even the scope of the issue is wrapped up in the audience's view of the state's internal politics.

Guzman acknowledges that if reputational losses are cabined into one issue area, then reputation may not be an effective means to enforce wide-ranging compliance with IL. Let us say that a government, which violates a human rights treaty, suffers *only* a reputational loss in the human rights issue area. This is not much of a deterrent to future violations, because the government would continue to be invited to join trade agreements, security treaties, and other cooperative ventures – just not human rights treaties.[3] For reputational concerns to support the system of IL, there must be overlap with other issue areas. But as Downs and Jones (2002) note, it is not clear what informational content the violation of a human rights agreement has for other issue areas. Here, Guzman responds to the scope issue by noting that the audience will probably draw the conclusion that a violation in one area indicates that the state has a somewhat greater likelihood of violation across many areas. Analogizing reputation to ripples on a lake, Guzman notes that '[t]he reputational consequences are greatest at the point of impact, and at points further from the specific violation, the reputational consequences diminish' (Guzman, 2008: 103).

A violation in one issue area, however, need not lead the audience to conclude that a violation in any other area is more likely. In fact, the audience may well think that the government is more likely to comply with IL in other issue areas, depending on the cause of non-compliance. For instance, the election of a Green Party to power might indicate that the new government is more likely to abide by environmental treaties but less likely to abide by a trade agreement that restricts environmental regulation.

Consequently, the conclusion that Guzman draws about reputation across issue areas – that we can, at a minimum, agree that a state which breaches an international obligation can be assumed to have a lower discount rate across a wide range of issues – does not necessarily hold. *Compliance with international rules is a function of domestic support for the goals of the treaty regimes as well as respect for international legal obligations*. IL is not apolitical or homogenous good; treaties are policy

---

[3] As an empirical matter, even this is doubtful. Governments with the worst human rights records are still regularly invited to join new human rights treaties (Blum, 2008).

outcomes as well as legal agreements. Non-compliance may signal a shift in the policy goals, not a general lack of respect for legal rules.

Public goods

The cabining of reputation to certain issue areas is particularly important for Guzman's claims about efficacy of reputational concerns in public goods agreements. One of Guzman's novel claims is that reputational concerns are particularly effective in the multilateral context and 'capable of explaining the multilateral cooperation that we observe, even in these public goods areas' (Guzman, 2008: 69).

Most rationalist IL theorists rely heavily on reciprocity and retaliation to explain compliance in public goods treaties, while acknowledging that both reciprocal non-compliance and (in issue) retaliation are not ideal mechanisms to enforce public goods treaties. Guzman is optimistic about the potential for reputational concerns to sustain cooperation in the multilateral public goods context, because there are many states that can impose reputational sanctions on a violating state. All the injured states will impose the reputational sanction because it is costless to them (i.e. the injured states simply update their view of the violating state's likelihood of complying in the future) without exacerbating the public goods issue. In addition, these sanctions are very costly to the violating state as it may be excluded from a wide variety of cooperative activities. The costs of reputational sanction are high enough, Guzman argues, to enforce even public goods treaties. Yet, we have reason to doubt Guzman's optimism. Reputational sanctions are unlikely to be an effective means of enforcing public goods agreements, such as those addressing climate change or nuclear non-proliferation, without a strong overlap with a different issue area.

For example, the problem of global climate change is a true public good. All states share the environment, and pollution from one state affects the global level of green house gases. Governments, thus far, have negotiated two multilateral treaties addressing climate change. The first, the Framework Convention on Climate Change, was a treaty with wide membership but few obligations to reduce greenhouse gas emissions. The successor treaty, the Kyoto Protocol, requires developed member states (the Annex I states) to reduce emissions to 1990 levels of greenhouse gases. Developing countries (the Annex II states) do not have to meet any specific emissions levels. The higher costs of compliance with the Kyoto Protocol, compared with the Framework Convention on Climate Change, led some states, most notably the United States, to decline to sign onto the Kyoto Protocol (von Stein, 2008). This lack of participation by the world's most polluting nation was widely viewed as a serious problem in

addressing climate change. In a public good scenario, state participation is of the utmost importance. Indeed, reputational concerns about compliance could theoretically deter governments from joining the treaty if the emissions goals are difficult to meet (or encourage the government to negotiate a regime that allows higher levels of emissions) (Bellinger, 2007).

In addition, reputational sanctions are not an effective means of enforcing the agreement among the states that have joined. The 'commitment period' of the Kyoto Protocol begins in 2008 and runs through 2012 so there has not been time for a state to violate the agreement. Given current emission levels, however, it is easy to imagine that some states will fail to meet their emissions goals. Yet we do not expect that complying members of the Kyoto Protocol are likely to exclude defecting members from joining other environmental treaties or make it more costly for defecting states to rejoin. The best means for keeping participation going in the regime may actually be to offer additional aid or lower the obligations for states that have difficulty meeting the emissions goals (Chayes and Chayes, 1993). Excluding non-complying states (either directly or by raising the costs of joining) from future environmental agreements would be worse for the environment than the violation of the Kyoto Protocol. Exclusion is self-defeating when wide participation is necessary to address a public goods problem. The same applies to threats of exclusion from the Nuclear Non-Proliferation Treaty.

For a government's concern with its reputation for compliance to deter violations of a public goods agreement, there must be more than the Kyoto Protocol on the table. If reputational effects are limited to climate change agreements, then the threat of exclusion lacks bite. Of course, a state might be concerned with its reputation as a 'good global citizen' and thus want to comply with the Kyoto Protocol or other public good agreements. Yet this is not the 'reputation for compliance' that Guzman is discussing. What makes a state a 'good global citizen' has only weak connection to a reputation for compliance.

For Guzman to be correct, the government's concern with its compliance reputation must apply across issue areas, such as exclusion of a violating state from cooperative activity in a non-public goods scenario. But, as discussed in the last section, Guzman does not explain why violations of a public goods treaty carry across issue areas. Indeed, Guzman acknowledges reputations can be compartmentalized and that the reputational effects across issue areas can be quite weak. Thus reciprocity and retaliation might not be an entirely effective means of sustaining compliance with public goods treaties, but reciprocity and retaliation may be *more* effective than reputational sanctions.

Unfalsifiable hypotheses

Even with its conceptual problems, the idea of the state's reputation is certainly doing some work in promoting compliance with IL. But, not having a clear conceptual framework that provides some parameters for the potential influence of reputational concerns creates a methodological problem. If Guzman were simply arguing that reputational concerns make some difference in governmental decision-making, then this methodological problem would not be significant. Guzman proposes, however, to build a theory of compliance around the concept of the state's reputation and so the methodological problem is much more significant.

For Guzman to have a *theory* of compliance, his theory must be capable of being falsified. We have to know, at least in some abstract way, when reputational concerns (either alone or with other variables) will lead a government to comply with IL and when they will not. The book advertises reputational costs as being quite high but acknowledges several limits – information, uncertainty, issue area scope, regime, and so on. We can read his book as supporting both very broad and very limited vision of reputational costs and benefits. Thus, when analyzing a government's decision to comply with IL, the value of reputation can be very high or approaching zero.

Consider the Bush Administration's decision in 2002 to raise steel tariffs in violation of WTO (World Trade Organization) law. Guzman's theory would support accounts of this violation that argued the government's action had both a very large and a very small effect on the reputation of the United States. If we believe that the audience broadly construes a state's violation of an international rule in one area as indicative of probable non-compliance elsewhere, then a violation of any IL can carry a high cost. In this view, the United States has demonstrated its willingness to violate an international trade agreement, which might indicate its general disregard of *all* international agreements. The Bush Administration's action could thus have long-term negative impact on the ability of the United States to form trade, security, or human rights agreement for many administrations to come.

Alternately, we might use Guzman's method to argue that such a violation has a very narrow impact. From this perspective, the violation only concerns the Bush Administration's reputation with regards to other trade agreements, or more narrowly, those trade provisions that deal with politically sensitive sectors, such as steel, agriculture, and textiles. In short, the book provides a conception of reputational costs that could be high or low in almost any situation.

Consequently, the book has difficulty meeting its goal of laying out a theory from which rationalist international theorists can proceed. Without a conceptual means of assessing reputational costs and benefits, claims about

the power of reputation remains non-falsifiable. In effect, reputation becomes an error term that allows us to explain nearly all examples of compliance or non-compliance as 'consistent with the state's interests'.

Conclusion

Guzman's book offers an innovative account of IL that is based on notions of state interest and yet optimistic about the possibility of achieving high levels of compliance. The book does so by developing the idea that states will comply with IL rules out of concern for their reputations. In doing so, Guzman adds a new life to discussions in both international relations and IL. If reputation is going to become a major causal mechanism in explaining how IL works, however, we need a more developed conception of how reputation works.

References

Bellinger, J.B. (2007), 'The United States and international law', reprinted in *Digest of United States Practice in International Law*, New York: Oxford University Press.

Blum, G. (2008), 'Bilateralism, multilateralism, and the architecture of international law', *Harvard International Law Journal* **49**(2): 323–379.

Brewster, R. (2009), 'Unpacking the state's reputation', *Harvard International Law Journal* **50**(2) (in press).

Chayes, A. and A.H. Chayes (1993), 'On compliance', *International Organization* **47**: 175–205.

Downs, G.W. and M.A. Jones (2002), 'Reputation, compliance, and international law', *Journal of Legal Studies* **31**: S95–S114.

Goldsmith, J.L. and E.A. Posner (2005), *The Limits of International Law*, New York: Oxford University Press.

Guzman, A.T. (2008), *How International Law Works: A Rational Choice Theory*, New York: Oxford University Press.

Hathaway, O. (2002), 'Do human rights treaties make a difference?', *Yale Law Journal* **111**: 1935–2042.

—— (2005), 'Between power and principle: an integrated theory of international law', *University of Chicago Law Review* **72**: 469–536.

Helfer, L. (2005), 'Exiting treaties', *Virginia Law Review* **91**: 1579–1648.

Independent International Commission on Kosovo. 2000. *The Kosovo Report*. New York: Oxford University Press.

Scott, R.E. and P.B. Stephan (2006), *Limits of the Leviathan: Contract Theory and the Enforcement of International Law*, New York: Cambridge University Press.

Simmons, B. (2000), 'International law and state behavior: commitment and compliance in international monetary affairs', *American Political Science Review* **94**: 819–835.

von Stein, J. (2008), 'The international law and politics of climate change: ratification of the United Nations framework convention and the Kyoto Protocol', *Journal of Conflict Resolution* **52**: 43–68.

Tomz, M. (2007), *Reputation and International Cooperation: Sovereign Debt Across the Centuries*, Princeton, NJ: Princeton University Press.