Frivolous But Rational: Moribund Hard Law in International Institutions*

Christopher Marcoux† Johannes Urpelainen‡

September 18, 2011

Abstract

States often create international institutions that impose legally binding rules on member states, and then do not even attempt to enforce these rules. Why? In this article, we present a game-theoretic model of moribund hard law in international institutions. We show that if some states face domestic pressure to negotiate a hard law treaty, their incentive to insist on hard law is maximized when less enthusiastic states expect that enforcement is not forthcoming. As a form of informal governance, moribund hard law is thus frivolous but rational.

*We thank Felicity Vabulas for useful comments on a previous draft.

†Visiting Assistant Professor of Political Science, New College of Florida. 5800 Bay Shore Road, SSC102. Sarasota, FL 34243, USA. +1-941-718-0794. cmarcoux@ncf.edu.

‡Assistant Professor of Political Science, Columbia University. 412 West 118th Street, 712 International Affairs Building. New York, NY 10027, USA. +1-734-757-0161. ju2178@columbia.edu.

1 Introduction

Neoliberal institutionalists have argued that treaty provisions are rationally designed to allow collective action among states (Abbott and Snidal, 1998; Keohane, 1984; Koremenos, Lipson, and Snidal, 2001). In many cases, though, this prediction seems to diverge from empirical reality. States negotiate treaties with seemingly stringent "hard law" provisions that are never used for anything. Human rights provisions in trade treaties go unenforced (Hafner-Burton, 2005), trade bans are ignored (Clapp, 1994), and a state's failure to comply with environmental rules does not result in sanctions (Börzel, 2000).

Why do states incorporate *moribund hard law* into treaty texts?[1] Previous research does not offer a compelling answer to this question. Yet it is directly relevant to the study of international institutions. Theoretically, it is an essential building block of any rigorous analytical model of treaty formation and compliance. Without a clear understanding of when formal rules matter, and when they are simply ignored, we cannot understand how international institutions really work (Stone, 2011). If some treaty provisions are not even intended to change state behavior, they should not be treated as equivalent to those that aim to induce compliance.

Empirically, such an understanding would be useful for two reasons. First, without a theory of moribund hard law, our empirical tests of treaty design produce biased results because "false positives" – treaty provisions that serve some purpose other than actually changing state behavior – are not accounted for. Some treaty provisions that states never intended to exert influence on their own behavior are interpreted as a rational response to international cooperation, even though their real purpose lies elsewhere.

Second, without a theory of moribund hard law, we also cannot test theories of treaty compliance. Some treaty provisions were simply not created to induce compliance, so they should not be included in studies of treaty compliance. The existence of these provisions is associated with compliance failure, and this is so because states wanted it to be so. Conventional approaches to treaty compliance would misconstrue this association as treaty ineffectiveness.

We develop a strategic model of treaty design that emphasizes bargaining interactions among states (Fearon, 1998). In our model, two states – North and South – negotiate on the design of a

[1] In this article, hard law is said to be "moribund" if states choose not to enforce compliance with legally binding provisions despite the existence of a sanctioning mechanism. For examples, see our data analysis below.

treaty. Of the two states, South worries that rigorous enforcement of hard law could be very costly to it.[2] Our model thus applies to a wide variety of treaties: the primary scope condition is that some member states did not support hard law, but instead preferred the flexibility of soft law. For example, developing countries could oppose stringent protection of intellectual property rights or enforcement of competition policy (Sell, 1995).

Building on this foundation, we pay particular attention to the role of domestic and international audiences, such as activists working for non-governmental organizations (NGOs). Recent scholarship has shown that NGOs are an increasingly central player in international affairs (Keck and Sikkink, 1998; Price, 2003; Wapner, 1995). We argue that in addition to the substantive cooperation problem at hand, states consider audience reactions to treaty provisions. A state that is not really interested in stringent environmental regulations could incorporate them in treaties to appease domestic environmentalists.[3]

Our central finding is thus a paradox of moribund hard law: stringent treaty provisions are the most likely when states are the least likely to enforce them. If a state with a high compliance cost expects that stringent treaty provisions will be enforced, it has incentives to reject a treaty that incorporates hard law. But if this state expects that stringent treaty provisions will not be used, it can accept a hard law treaty. To the extent that another state obtains audience benefits from hard law, it can support them without any intention to actually use them. The expectation of failed enforcement allows treaty formation.[4]

This finding has some important implications for the study of "informal governance" in international politics (Stone, 2011). While informal expectations and norms may sometimes function as a rational complement to formal rules, and thus facilitate international cooperation, we show that informal governance can also be a domestic smokescreen. If states expect few gains from cooperation but their domestic audiences demand hard law, the inclusion of moribund law may offer a way out. In this sense, moribund hard law is frivolous but rational: it helps states deal with domestic political pressures, yet it does not force states to incur high compliance costs.[5]

[2]The labels "North" and "South" need not be interpreted literally. In many cases, industrialized countries have more to lose from rigorous enforcement of treaty obligations.

[3]While this act may not fool the environmentalists in the long run, it does reduce political pressure on the government until the failure to enforce the treaty becomes apparent.

[4]This is not to say that hard law goes unenforced for sure. Instead, we expect it to become moribund with some probability. In some instances, states counting on failed enforcement are in for a nasty surprise.

[5]The word "frivolous," therefore, does not mean that moribund hard law is without political meaning. Akin to a

In addition to providing a variety of qualitative examples, we systematically test the argument against data on environmental regime formation (Young and Zürn, 2006). We find that if NGOs participate in negotiations, the resulting environmental treaties are more likely to contain sanctions and compliance review mechanisms. A qualitative analysis of *de facto* enforcement of treaties with enforcement provisions shows, however, that these enforcement mechanisms were often not used.

The article is organized as follows. We first review the literature and outline the research need that this article meets. We then discuss the theoretical model and the testable hypotheses. Next, we present the empirical test. The conclusion summarizes the argument and returns to its broader implications for the study of international institutions and cooperation.

2 Soft and Hard Law in International Institutions

International institutions are "explicit arrangements, negotiated among international actors, that prescribe, proscribe, and/or authorize behavior" (Koremenos, Lipson, and Snidal, 2001, 762). Substantial variation exists between such institutions, though, along the legal obligations that they impose (Abbott et al., 2000). Some treaties are best characterized as *hard law*, for they impose precise and binding legal obligations on signatories. Others are best seen as *soft law*, for they provide recommendations and guidelines that cannot be interpreted as legally binding (Abbott and Snidal, 2000; Shaffer and Pollack, 2010).

The conventional wisdom in international cooperation theory on the choice and design of such obligations adopts a functionalist perspective. It thus emphasizes their collective rationality (Abbott and Snidal, 1998; Keohane, 1984; Koremenos, Lipson, and Snidal, 2001). The choice between soft and hard law, the argument goes, is determined by such factors as the transaction costs of formal ratification and the perceived need for flexibility versus credible commitment (Lipson, 1991). According to neoliberal institutionalist theories, states design international institutions to minimize the transaction costs of cooperation. Thus, treaty provisions are designed in a way that allows a group of states to realize joint gains from cooperation. When the transaction costs of negotiating and designing such provisions do not exceed the collective benefits, we should see treaty provisions that address international cooperation problems in ways that appeal to the negotiating parties.

frivolous lawsuit, moribund hard law is politically meaningful but does not result in formal rule enforcement.

While the basic idea of collective rationality is relevant to treaty negotiations, in reality treaties and institutions are negotiated in a strategic bargaining setting (Fearon, 1998; Urpelainen, 2011; Young, 1989). Thus, the existence of joint gains from cooperation will not inevitably lead to their realization. States may fail to negotiate a treaty that is potentially beneficial due to a variety of bargaining failures, ranging from incomplete information (Gartzke, 1999) to commitment problems (Powell, 2006). While the possibility of such bargaining failures is readily recognized in the more abstract literature on distributional bargaining, we are not aware of any systematic theoretical arguments that could explain when and how irrational treaty designs can expected to emerge.

This problem is not an irrelevant theoretical curiosity. States frequently negotiate treaties with bizarre legal obligations that no one seems to have any interest in. Consider, for example, the hazardous waste provisions included in the Lomé IV Convention between the European Community and the ACP developing countries that was negotiated in 1989. The agreement not only contains an absolute export ban for members of the European Community, something that was clearly opposed by all of the major European powers, but also a provision that prevents the ACP developing countries from importing any hazardous waste from anywhere in the world (Clapp, 2001). It is hard to imagine what the possible gains from such a fundamental violation of national sovereignty are. The provision has never been enforced since it was adopted, despite widespread and blatant non-compliance by many parties, and it was not included in the successor to the Lomé IV Convention, namely the 2000 Cotonou Agreement. Why adopt hard law provisions that flagrantly violate national sovereignty, and then not use them at all?

Mexico's experience with the North American Agreement on Environmental Cooperation is in many ways similar. Mexico accepted an environmental side agreement to secure the ratification of the North American Free Trade Agreement in the United States (Cameron, 1997), and since then evidence indicates that Mexico has done little to actually upgrade national environmental regulations in view of the requirements (Bechtel and Tosun, 2009). Curiously, though, the United States has shown little interest in enforcing these regulations. Why adopt stringent regulations for environmental cooperation if the more powerful party has no interest in enforcing them?

The extant literature provides some possible explanations for these patterns, but upon careful inspection they all turn out to be somewhat implausible or incomplete. Scholars working within the realist tradition have argued that treaties are irrelevant "window dressing" (Mearsheimer, 1994-

1995). If this hypothesis were valid, then treaty provisions would be pretty much random and subject to historical fads. But the empirical evidence clearly shows that they are not (Koremenos, 2005). Additionally, empirical scholarship has documented treaty effects that are inconsistent with the notion that they are simply window dressing (Dai, 2005; Mitchell, 1994). In a recent review on treaty compliance, Simmons (2010) shows that empirical scholarship on treaty compliance indicates that, on balance, they have effects on state behavior across various issue areas.

The window dressing hypothesis notwithstanding, another possible explanation pertains to domestic audiences. Koremenos (2011) and Stasavage (2004) argue, for example, that public negotiations sometimes result in flawed treaties because governments appease domestic interest groups with narrow interests that complicate cooperation. Indeed, activists have become increasingly visible in multilateral negotiations on a variety of issues from human rights to landmines and environmental cooperation (Price, 2003; Raustiala, 1997).

This explanation may not apply to seemingly stringent legal obligations and other forms of hard law that are never used. If states have no intention to use hard law, and if activists understand this, why would the act of including such provisions *per se* have any reputational benefits for states in the eyes of these activists? Unless the activists are irrational, in contradiction to recent empirical research on the strategic dimension of international activism (Keck and Sikkink, 1998; Ron, Ramos, and Rodgers, 2005), one would imagine that they also realize the frivolity of such treaty provisions. This is particularly problematic in view of the emerging literature emphasizing the importance of strategy for activists and other special interests (Bob, 2005; Keck and Sikkink, 1998). We show that the irrationality assumption is not required to explain why states obtain reputational benefits from hard law. Our game-theoretic model provides a detailed characterization of the conditions under which states have incentives to incorporate hard law into formal treaties, so as to "posture" in front of an NGO audience. In particular, we show how states can sometimes bluff rational audiences for reputational gains, at least in the short run.

In a recent book, Stone (2011) has argued that states often engage in "informal governance," ignoring formal rules when it is expedient. According to Stone (2011), a hegemonic state's ability to override formal rules in exceptional circumstances is a key reason why this hegemonic state is willing to comply with international law in the first place. However, this account does not explain why powerful states would propose legal rules that benefit them, and subsequently fail to enforce

them in spite of their dominant position in the international system.

A final, sociological explanation pertains to the notion of "institutional isomorphism:" maybe states use hard law simply out of mimicry and emulation, or because other states also legalize their treaties (DiMaggio and Powell, 1983; Meyer et al., 1997)? If this explanation held, then one would expect treaty designs to converge over time due to increased global pressure to use international law. But as Koremenos (2005) shows, this hypothesis can be rejected: treaty designs are not converging. Instead, treaty design is characterized by considerable variation, and this variation is not random, so it would be hard to ascribe the use of moribund hard law to uniform global pressures.

We provide a systematic theoretical explanation for the use of moribund hard law in international institutions. We show that if uncertainty surrounds states' preferences, their strategic bargaining interactions sometimes produce moribund hard law. We adopt the premise that states may have incentives to posture to domestic and international audiences, and we show exactly how such posturing incentives produce moribund hard law in certain empirically identifiable circumstances: states include legal obligations that they prefer not to use. Our theoretical model allows us to generate empirically falsifiable hypotheses on the inclusion of moribund hard law, and extensive empirical examples and a systematic data analysis provide evidence in support of these contentions.

3 Frivolous But Rational: The Argument

We now present and solve an analytical model that allows us to explain why frivolous treaty provisions often stem from strategic bargaining interactions between rational parties. We first present the model, then characterize the equilibria and flesh out the primary empirical implications, and finally summarize some model extensions of interest. A full mathematical analysis can be found in the appendix.

As we build the model, we rely on two analytical scope conditions. First, we assume that at least one negotiating state is worried that hard law could carry a high cost (Lipson, 1991). Without this assumption, states would have no reason to refrain from using hard law. The assumption is particularly plausible in negotiations between states with very different interests, such as North-South (Sell, 1995) or East-West bargaining (Kydd, 2000).

Second, we assume that at least one negotiating state would reap reputational benefits from presenting itself as a resolute supporter of hard law to NGOs (Keck and Sikkink, 1998; Raustiala, 1997). Without this assumption, states would not have incentives to incorporate hard law into a treaty and then not enforce it. Our preliminary quantitative examination of the effect of NGO participation on enforce and monitoring mechanisms shows that, indeed, in the absence of NGO participation such mechanisms are very rarely used.

3.1 Model

The game is played by three players. *North* and *South* are states bargaining over international treaty design to address a cooperation problem, and *NGO* is an interested third party who prefers stringent legal obligations. For example, North could be a representative European state while South would be a representative African state. If the treaty focuses on environmental cooperation, the NGO could be Greenpeace.

The model is intended to apply more broadly than North-South interactions, yet we use these labels for ease of exposition. North can be thought of as a forerunner state that is interested in creating a new treaty, and South can be thought of as representing those states that would have to join to make the treaty work. Thus, North proposes hard or soft law, and South accepts or rejects. This sequence of moves seems plausible given the idea that the most enthusiastic states generally play a key role in the design of international treaties (Benedick, 1998; Oberthür and Ott, 1999; Steinberg, 2002).[6]

The exact nature of soft and hard law depends on the context. For example, trade cooperation based on exact tariff reductions and a dispute settlement mechanism would be an example of hard law (McCall Smith, 2000). In contrast, an "aspirational" trade treaty without binding obligations would exemplify soft law. In the case of environmental cooperation, binding emissions reductions accompanied by the threat of trade sanctions would exemplify hard law (Steinberg, 1997). Public commitments to sustainability without specific obligations would exemplify soft law.

Except for nature's draw of North's Bayesian type in the beginning of the game, the sequence of moves is the following:

[6] In reality, it is equally plausible that both states simultaneously support or oppose hard law. For simplicity, we assume asymmetry in the mathematical analysis.

1. North offers a treaty provision, $T = \{S, H\}$.

2. South accepts or rejects the provision, $A(T) \in \{0, 1\}$.

3. If the provision is hard, $T^* = H$, and South accepts, $A(H) = 1$, North decides whether to enforce, $E \in \{0, 1\}$.

4. If the provision is hard, $T^* = H$, and South accepts, $A(H) = 1$, nature sends a public signal $q \in \{0, 1\}$ regarding enforcement E^*.

5. The NGO selects a level of support for North, $S \in [0, 1]$.

In this sequence, North is the state with the initiative, and it must select between a "soft" S and "hard" H obligation (Abbott and Snidal, 2000; Lipson, 1991). South observes this offer, and either accepts or rejects the treaty. If South accepts a hard obligation, then North must decide on enforcing it. The enforcement decision is imperfectly observable, and so the NGO may or may not learn whether North actually enforced the treaty or not. The sequence of moves is illustrated in FIGURE 1. Given the complexity of the game, we do not include payoffs and moves by nature in this figure.

[Figure 1 about here.]

We assume, for simplicity, that North holds the initiative. An extension that we consider shows, however, that if South sets the negotiation agenda with some positive probability, the basic insights from the equilibrium analysis remain intact. South obviously has an incentive to propose soft law, but as long as North moves first with nonzero probability, all of our comparative statics remain intact.

The game has several different outcomes. If South rejects the treaty, the status quo without a treaty prevails. If South accepts a soft provision, modest soft law cooperation prevails. If South accepts a hard provision, the outcome depends on whether North actually enforces the treaty, and on the signal q. This signal either indicates that North is really enforcing the treaty or not.

In addition to the probabilistic signal, the game has one variable that is initially subject to uncertainty: North's type. North is either interested or disinterested, $t \in \{int, dis\}$. An interested North actually prefers a stringent and enforceable treaty, whereas a disinterested North prefers a

weak treaty. Let $p^{int} \in (0,1)$ denote the prior probability that North is actually interested, and suppose nature secretly reveals this private information to North before the game begins. Thus, North's type is private information.

The players have different incentives. The NGO plays the role of an audience in the game, so it has a rather passive role. At the end of the game, it selects the level of support for North, and it does so in view of the posterior probability μ^{int} that it ascribes to North being interested. Without loss of generality, suppose the NGO selects the level of support S such that optimum is $S^* = \mu^{int}$. This simple assumption states that the NGO's optimal support level increases with the probability that North is really interested. For example, Greenpeace has stronger incentives to support governments that are probably interested in environmental regulation.

This support level should be interpreted as the NGO's immediate reaction to the negotiation outcome. Over time, the NGO obtains new information about North's preferences. Indeed, successful or failed enforcement of hard law is itself a very informative signal. But such new information arrives with delay. Thus, even a state that is not willing to enforce hard law may obtain temporary reputational benefits from disingenuously negotiating a hard law treaty. For example, environmentalists may increase their support to the United States government in elections due to a new multilateral environmental treaty, even if they later criticize this very government for implementation failure.

Consider now South. It obtains a payoff zero from the status quo, and a bonus of 1 if the treaty is weak or not enforced. If the treaty is enforced, it obtains a payoff $1 - c$, where $c > 1$ so that a stringent and enforced treaty is not something South prefers.

North's payoff depends on its type. It obtains a payoff of wS from the status quo, and a payoff $1 + wS$ from a weak or non-enforced treaty. Here $w > 0$ indicates the value that North ascribes to NGO support for reputational reasons. From a stringent and enforced treaty, North obtains $1 + I + wS$ if it is interested, where $I > 0$, and $1 - D + wS$, where $D > 0$, if it is disinterested. Simply put, the only difference between an interested and disinterested North is the state's payoff from enforcing the treaty. To maintain a sharp focus on this difference, we assume that $D > w$. This assumption implies that a disinterested North has a strong preference against enforcement, and so it will not enforce for reputational reasons.

The uncertainty is focused on one player for simplicity. In the section on model extension, we ex-

plain why allowing uncertainty about South's preferences would not change any of our predictions. The focus on a single state is warranted because in many situations one of the players is potentially capable of enforcement while the other is not. This is certainly the case with many actual North-South treaties given the power asymmetry that often characterizes North-South interactions (Bechtel and Tosun, 2009; Najam, 2005).

Nature's signal is $q = 1$ with probability one if North enforces. This assumption implies that if North enforces, it can provide evidence to domestic and international audiences regarding this act of enforcement. But if North does not enforce, the signal acquires a value $q = 0$ with some probability $G \in (0, 1)$. Notably, this assumption implies that North sometimes "gets away" without enforcing hard law. This is plausible, as effective enforcement mechanisms may deter state parties from violating treaties, and thus no visible act of enforcement can necessarily be expected (Urpelainen, 2010). Thus, North could always argue that it has found a clever way to enforce the treaty, yet this mechanism is invisible to third parties.

3.2 Equilibria

The game has a number of different equilibria. In each equilibrium, the two types of North select their design offer optimally, and South updates its beliefs regarding North's type before accepting or rejecting. North then decides on enforcement based on its type, and finally the NGO selects the political support level for North. We then summarize the equilibrium in a series of analytical claims that are proven in the mathematical appendix.

To begin with, let us reject the possibility of a "separating equilibrium."

Claim 1 (no separating equilibrium). *No equilibrium exists wherein the two different types of North play different pure strategies.*

The intuition behind this claim is straightforward. If the two different types of North were to play different pure strategies, South could infer North's type from these strategies. This is the very idea of separation in a signaling model. The interested type would propose hard law, and thus have an incentive to actually enforce it, so South would reject to avoid the high cost of enforced hard law. From this, it thus follows that at least one type of North would have to deviate from the

equilibrium strategy. In other words, a separating equilibrium cannot exist because it would allow states with a strong preference for soft law to simply remain outsiders.

In principle, the game may have exactly two different pooling equilibria. In a "pooling up" equilibrium, both types of North offer hard law; in a "pooling down" equilibrium, both types of North offer soft law. Consider first the possibility of hard law.

Claim 2 (pooling up). *If South's compliance cost c is so low that it accepts hard law based on its prior beliefs, $p^{int}(1-c)+(1-p^{int})>0$, the game has a unique equilibrium that meets the following criteria:*

1. *Both types of North offer the hard provision, $T^*=H$;*
2. *South accepts the hard provision, $T^*=H$;*
3. *Only the interested type of North enforces the hard provision, $E^*=1$ if and only if $t=int$.*

This claim states that for hard law to be offered with certainty, South's cost from hard law must be low enough. This is intuitive, as a very high cost for South would induce it to reject a hard law offer. Consequently, the disinterested type of North would offer soft law instead.

What about the possibility of both types always offering soft law?

Claim 3 (pooling down). *If South's compliance cost c is so high that it rejects hard law based on prior beliefs, $p^{int}(1-c)+(1-p^{int})<0$, while the reputational value w is sufficiently low, the game has a non-unique equilibrium that meets the following criteria:*

1. *Both types of North offer the soft provision, $T^*=S$;*
2. *South accepts the soft provision, $T^*=S$, and rejects the hard provision, $T^*=H$.*
3. *Only the interested type of North enforces the hard provision, $E^*=1 \Leftrightarrow t=int$.*

In this equilibrium, the two types of North pool down. The reason why such an equilibrium may exist is that if the interested type of North were to deviate by suddenly offering hard law, South would realize that it will be enforced, and thus South would reject the offer. If the value of reputation is low enough, even the interested type of North is unwilling to sacrifice the treaty for

a reputational benefit in the eyes of the NGO. But if the value of reputation is high enough, the interested type of North deviates and the "pooling down" equilibrium implodes.

Finally, an equilibrium in mixed strategies may also exist.

Claim 4 (mixed strategies). *If South's compliance cost c is so high that it rejects a stringent treaty based on prior beliefs, $p^{int}(1-c)+(1-p^{int})<0$, the game has an equilibrium in mixed strategies that meets the following criteria:*

1. *An interested North offers the hard provision, $T^*=H$, whereas a disinterested North plays hard and soft with positive probability;*
2. *South accepts the soft provision, $T^*=S$, while accepting and rejecting the hard provision, $T^*=H$, with positive probability;*
3. *Only the interested type of North enforces the hard provision, $E^*=1 \Leftrightarrow t=int$.*

This equilibrium is unique whenever the reputational value w is sufficiently high.

In this equilibrium, the interested type of North always offers hard law. The disinterested type, though, mixes between soft and hard law. Consequently, South is unambiguously willing to accept soft law but may sometimes reject hard law. Interestingly, this equilibrium may exist simultaneously with the "pooling down equilibrium," as is illustrated below.

These outcomes are summarized in FIGURE 2. The game has a unique equilibrium except when South's hard law cost c is high while the value of reputation w for North is low. When South's hard law cost is low, the "pooling up" equilibrium is played, with North proposing hard law and South accepting it. In this case, enforcement remains probabilistic because North's preferences are subject to uncertainty. Additionally, the NGO is unable to update its beliefs. When South's hard law cost is high and reputational is valuable, a "pooling down" equilibrium exists. In this equilibrium, hard law is never created. If South's hard law cost is high, this suffices for an equilibrium in mixed strategies to exist.

[Figure 2 about here.]

3.3 Empirical Implications

Based on the above equilibrium analysis, we are in a position to derive several empirical implications. We rely on the South's compliance cost c and the North's interest in NGO reputation q to explain the probability of hard law and enforcement thereof.

Let us, first, characterize the situations in which hard law is used for sure but enforced only occasionally.

Proposition 1 (hard law). *If South's compliance cost c is low enough, hard law is used with certainty. The size of this region decreases as the probability that North is interested increases, but the conditional probability of enforcement increases as the probability that North is interested increases.*

Consider FIGURE 3. It shows a paradoxical effect: as North's interest in stringent provisions increases, they become less probable. The reason is that South becomes increasingly worried about harmful enforcement, and is thus more willing to reject hard law. But at the same time the probability that they are used conditional on their existence increases given North's interest.

[Figure 3 about here.]

This finding has important theoretical implications for the study of international institutional design. Beginning with Keohane (1984), scholars working within the neoliberal institutionalist tradition have argued that institutional design is a rational response to cooperation problems (Abbott and Snidal, 1998; Koremenos, Lipson, and Snidal, 2001). The present analysis shows, though, that as states become less interested in enforcing hard law, the probability that they use it actually increases. Thus, the conjecture that states enact hard law to enforce commitments turns out to be highly problematic in situations characterized by distributional conflict.

This observation resolves our main puzzle: why do states worrying about enforcement accept hard law? The reason is that they ascribe a low probability to enforcement. It is certainly possible that other states ultimately choose to enforce the legal provision. But if the probability of enforcement is low enough, this is not enough to outweigh the benefits of treaty membership in the

absence of enforcement. We do *not* expect all hard law to be unenforced. Rather, we expect some portion of it to be unenforced.

By contrast, the use of hard law, conditional on it existing, becomes more likely as states become more interested in it. This implies that studies of hard law and soft law could profit from shifting their focus from institutional design to actual practice and implementation. Without such a shift in focus, it is hard to say whether hard law is enacted for reputational reasons or whether states have a real interest in it.

Bargaining over environmental and labor provisions in the North American Free Trade Agreement (NAFTA) and the US – Jordan Free Trade Agreement offers an illustration. At the 1990 meeting of the World Economic Forum in Davos, Mexican president Carlos Salinas de Gortari and U.S. president George H. W. Bush agreed to the negotiation of a U.S. - Mexico free trade agreement. Canada subsequently joined negotiations. As a condition of granting "fast track" authority to the Bush administration, the Democratic-led Congress stipulated that the treaty would have to address various environmental and labor concerns – two traditional constituencies of the Democratic party. The 1992 election of Bill Clinton amplified this concern, and both labor and environmental groups played an important role in the NAFTA debate (Audley, 1997; Cameron, 1997; Raustiala, 1997). In 1993, the North American Agreement on Environmental Cooperation was negotiated to accompany NAFTA. Mexico ratified it, and thus paved way for the ratification of NAFTA by the US Congress (Audley, 1997).

These negotiations illustrate the relationship stated in our first proposition: if South's compliance cost is low enough, then hard law is expected. Mexico's compliance cost with the NAFTA side agreement was, indeed, low relative to the treaty benefits. Mexican politicians worried that NAFTA negotiations would fail, or that the negotiations would stretch into the 1994 presidential elections. From the perspective of the Salinas administration, which had consistently advocated economic liberalization, the NAFTA negotiations represented an opportunity to improve economic growth and lock in liberal reforms. As Cameron (1997, 124) puts it, "Mexico was in a vulnerable position during these negotiations – failure to reach supplemental accords could derail NAFTA, and Mexico needed NAFTA more than ever as economic growth began to falter in 1993 and the date for presidential elections approached." Even delaying the process imposed serious costs. Though the PRI had controlled national government for decades, a combination of domestic political unrest

and economic crisis left it vulnerable. A salient policy failure would not help it remain in power. Even though Mexico had to implement some environmental reforms (Husted and Logsdon, 1997), their costs were much lower than the expected cost of NAFTA failure or delay.

Indeed, the case of NAFTA also illustrates the possibility of the imperfect enforcement of hard law. Although Mexico did implement environmental reforms, many were only partially enforced and certainly not what environmental organizations in the United States originally had in mind (Auer, 2001; Bechtel and Tosun, 2009). This is consistent with the proposition, for it illustrates that the inclusion of hard law in a treaty text may not be enforced.

The U.S. – Jordan free trade agreement offers a particularly useful illustration of the hypothesis that hard law should often not be enforced. As negotiated and signed by the outgoing Clinton administration, it also contained provisions to safeguard environmental and labor standards. However, it was ultimately ratified and implemented by the Bush administration. As a Republican, Bush was much less interested in environmental and labor standards. Thus, in spite of the inclusion of hard law in the agreement, the two countries formally agreed (1) that appropriate responses to non-trade disputes did not include trade sanctions; and (2) that they would refrain from using the formal dispute resolution procedure outlined in the treaty (Rosen, 2004). The change in the partisan preferences of the US executive prevented the enforcement of hard law, and thus the seemingly stringent treaty provisions remained moribund.

Consider next the possibility of soft law.

Proposition 2 (soft law). *If South's compliance cost c is high enough, and the value of reputation q is low enough, an equilibrium exists wherein soft law is used with certainty. The size of this region increases with the probability that North is interested, but the conditional probability of enforcement decreases with it.*

This proposition is basically the converse of PROPOSITION 1. Soft law is used in this equilibrium. The probability of this outcome declines with the probability that North is interested, but the conditional probability of enforcement increases with it. One may refer back to the previous figure to see how this works. Again, the logic is strategic: if South is sufficiently worried about hard law, then the only possible equilibrium treaty is one that is based on soft law. As North's

interest in enforcement grows, its ability to entice South to participate declines.

The theoretical implications are notable. If North is almost certain to prefer hard law, and has little interest in the NGO's views, then soft law is used. This is somewhat paradoxical, as it implies that states with a keen interest in hard law often end up relying on soft law. The key to resolving this paradox is to recognize the importance of asymmetric interests, and understand that if a state becomes increasingly willing to enforce hard law, then states that do not prefer enforcement have stronger incentives to avoid hard law than previously.

This logic is nicely illustrated by the Convention on Biological Diversity. Largely through the efforts of non-governmental organizations, such as the International Union for Conservation of Nature (IUCN), the issue of biological diversity quickly rose to prominence on the international agenda. Multilateral negotiations culminated with the adoption of the Convention on Biological Diversity (CBD) in May 1992. Though this international agreement was formalized in a legal document, it nonetheless may be described as an example of the use of soft law. The CBD did not set forth specific targets for biodiversity conservation, nor did it prescribe specific regulations for member governments to implement.

According to PROPOSITION 2, if South's compliance cost is high enough, and the North's valuation of NGO reputation is low enough, an equilibrium exists wherein soft law is used with certainty. Our first step, then, is to assess South's cost of compliance. In fact, these compliance costs were potentially quite large.

The most immediate costs to South were those related to sovereignty. As Tolba (1998) explains, when the issue of biological diversity reached international prominence, it was defined as "common heritage of mankind." Initially, attempts to preserve biodiversity centered on states' responsibility to preserve this common, global heritage. But biological diversity is not distributed equally throughout the international system. It is greatest in warm and tropical climates, and these are home primarily to developing countries. The situation is exacerbated by the fact that, historically, industrial development has been one of the primary threats to the maintenance of biological diversity. The developed states of the global North began with a relatively smaller endowment of biological diversity and, by the late twentieth century, they had lost a substantial amount of what diversity they once had as a consequence of industrialization. As a practical consequence, the immediate policies required to conserve biological diversity would fall disproportionately on the

global South (Sell, 1995). Even before negotiations had formally started, a UNEP working group concluded that any feasible international treaty would need to "provide realistically for the transfer of resources, allowing . . . implementation by the poorer countries that are also the custodians of much of the biological heritage of the earth" (Tolba, 1998, 137).

The South also faced the possibility of substantial opportunity costs from biodiversity conservation. Biological diversity has both "productive" and "consumptive" value (Mitchell, 1999; Tolba, 1998). It is useful in a productive sense because genetic diversity is important for biotechnology. However, any enjoyment of the productive use of biodiversity comes directly at the expense of its consumption value. Therefore, even if South receives side payments that cover the immediate costs of preserving and administering forest resources, that side payment will have a net negative value if it requires South to forgo development projects that may provide immediate economic gains.

This leads directly to distributional costs surrounding the productive uses of biological diversity. Biotechnology firms in the North seek access to genetic resources in order to innovate. For investment in biological diversity to be profitable, firms sought assurance that any international treaty to preserve biodiversity would also safeguard intellectual property rights. This issue brought North and South into direct conflict. North was less interested in preserving biological diversity if it required private investors to share valuable intellectual property; throughout much of the CBD negotiations, biotech industry representatives urged that financial transfers be restricted to recipient states that would protect the patents of products developed using protected genetic resources. South was less interested in preserving biological diversity if it resulted in foregoing valuable opportunities for economic development. As (Tolba, 1998) recounts, there were some novel attempts during negotiations to find a middle way through this impasse. One proposal would have required private firms to share intellectual property with host countries, but would require host countries to prohibit domestic firms from competing on international markets. While compromise would mitigate the most politically explosive situations – e.g., private firms in the North selling medicines developed from local resources at exorbitant prices – it would also eliminate much of the economic incentive for South to preserve the productive value of biodiversity (Tolba, 1998).

The Convention on Biological Diversity ultimately left most of these disputes intact. Rather than establish a supranational regulatory regime governing biodiversity, it essentially created a forum for continued negotiation. Article 6 directs contracting parties to take action to conserve

biological diversity, but according to a nationally created plan of action. Regarding the controversial issue of intellectual property, the CBD directs firms to negotiate the terms of foreign investment with host countries on a case by case basis. While Article 15.2 directs contracting parties to "facilitate access to genetic resources for environmentally sound uses by other Contracting Parties," Article 15.4 notes that "access, where granted, shall be on mutually agreed terms," To ensure that access to genetic resources only occurs on mutually agreeable terms, the Article 15.5 requires investing firms to obtain prior informed consent from a host country before accessing genetic resources. In short, while the CBD represents an agreement regarding the significance of biological diversity, it clearly upholds the sovereignty of contracting parties, and creates no costly obligations for South other than to collect information on the extent of biodiversity (Article 7) and to develop an internal plan for conserving biodiversity (Article 6).

Why did North not gamble on a hard law arrangement? PROPOSITION 2 states that such a gamble is not profitable if North's interest in cooperation is large while the value of NGO reputation is relatively low. Given that the CBD was negotiated in a general multilateral summit on the global environment, a failure to reach agreement would have jeopardized the entire summit. Although non-governmental organizations were engaged in the process, their preferences were not unanimous. Environmental groups pressed for stronger measures to protect biodiversity. However, influential industry groups in the United States and elsewhere pressed for an agreement that respected international property rights (Sell, 1996).

Consider, finally, the possibility of mixed strategies. When may we expect negotiation failure due to an unacceptable hard law offer?

Proposition 3 (mixed strategies). *If South's compliance cost is high enough, an equilibrium exists wherein hard law is used with some probability and South rejects the treaty with some probability. The size of this region increases as the probability that North is interested increases. The probability of a hard law proposal increases with the probability that North is interested and South's cost. The conditional probability of actual enforcement does not depend on the probability that North is interested but decreases as South's cost increases. The probability that South accepts the proposal decreases as the reputational value q increases.*

This proposition shows an interesting logic: the probability of hard law increases with the probability that North is interested, whereas the probability of enforcement decreases. South's cost, surprisingly, has similar effects.

This proposition shows that as long as South's cost is high, even in an equilibrium with mixed strategies the use of soft law is possible. As South's cost increases, surprisingly the enactment of hard law is more probable. At the same time, though, the conditional probability that it is used decreases. These findings are very similar to those in PROPOSITION 1, and thus testify to the robustness of the underlying theoretical logic.

This process is exemplified by the failure to reach agreement on a compliance system at the sixth Conference of Parties (COP) to the UN Framework Convention on Climate Change (UNFCCC), held in The Hague in November 2000. Three years earlier, parties to the UNFCCC adopted the Kyoto Protocol, which introduced a hard law regulatory agreement to the broader international regime on climate. Although the Kyoto Protocol is commonly credited with stipulating legally binding limits to greenhouse gas emissions, agreement at Kyoto was made possible, at least in part, by leaving a number of provisions only vaguely defined. In the first few years after the adoption of the Kyoto Protocol, COP meetings focused mainly on the further development of several flexibility mechanisms – e.g., emissions trading, joint implementation, and the Clean Development Mechanism – that negotiators had designed into the Kyoto Protocol. By 2000, however, many parties had turned their attention to the task of developing a formal compliance system for the Kyoto Protocol (ENB, 2000).

In addition to bargaining over the development of a formal compliance system, parties were also bargaining over carbon sinks - emissions offsets to be granted to specific parties based on their land use policies, and the carbon-absorbing potential of specific regions. The issue of carbon sinks bears directly on compliance costs. For example, a lenient definition of carbon sinks reduces compliance costs for certain affected parties, but has the potential to increase relative compliance costs for other parties, which may be forced to adopt disproportionately large policy adjustments to meet their own emissions requirements (Oberthür and Ott, 1999).

To repeat, our third proposition implies the following: the probability of a hard law proposal increases (i) with the probability that North is interested, and (ii) as South's costs increase. The probability that South accepts a hard law proposal decreases as NGO reputational value for the

North increases. To illustrate this proposition, consider the European Union as the proposing party (North). At the November 2000 conference, many EU members had great interest in reaching agreement on a hard law approach to compliance. These countries have historically experienced the greatest level of domestic political demand for a strong international regime on climate (ENB, 2000). To that end, EU parties to the UNFCCC had campaigned for the participation of more reluctant parties by agreeing to the inclusion of a number of flexibility mechanisms that were attractive to the "umbrella" group, an important negotiating bloc consisting of, among others, the United States, Canada, Australia and New Zealand. In this setting, the umbrella group would be on the receiving side (South).

Though the proposed mechanisms were in many ways expedient, they were also risky. The prevailing concern among EU parties at the November 2000 meeting was that these mechanisms should not become a *de facto* justification for non-compliance with Kyoto. To prevent this, a strong enforcement system would be needed (ENB, 2000).

There is ample evidence that South faced high compliance costs. This is evident in the Kyoto Protocol itself, which exempted developing state parties from mandatory reductions in the treaty's first compliance period. The issue of carbon sinks, which played a major role in discussions at The Hague in November 2000, was directly related to compliance costs for states with large forest areas, such as Russia and the United States. There is also evidence that EU parties faced powerful reputational incentives to cater to environmental groups. Environmental NGOs have been present in international climate negotiations, campaigning for more ambitious targets and stringent enforcement (Gulbrandsen and Andresen, 2004). At The Hague, NGO presence was so strong that negotiators broke from tradition and barred NGOs from attending some meetings on key issues, such as compliance. The fact that at least one "closed door" meeting was disrupted by public protest is yet further evidence of the degree of engagement of civil society in climate negotiations (ENB, 2000).

Following the logic described in PROPOSITION 3, we would expect the combination of high interest from North, high compliance costs for South, and civil society engagement to result in a hard law offer from North that will be rejected by South with some positive probability. Indeed, the sixth COP failed to produce an amendment to the Kyoto Protocol, and the compliance question pushed for the next meeting.

3.4 Model Extensions

In the main model, we assume that South's preferences are common knowledge. Consider now an extension in which South would prefer hard law with some probability $\upsilon \in (0,1)$. If South preferred hard law, it could credibly say so in costless pre-game communication, and thus North would always propose hard law without fear of rejection by South. This extension, therefore, would not qualify any of our substantive predictions. With some probability South would approve of hard law, and also say so in pre-game communication. If South did not approve of hard law, however, North would learn this because South would not have an incentive to issue a statement that would result in hard law for sure.

Another plausible extension pertains to South's initiative. If South could move first with some probability $\kappa \in (0,1)$, it would obviously propose soft law, and North would accept. Since South has a strong preference against hard law, it would set the negotiation agenda so as to preclude hard law. As long as $\kappa \neq 1$, though, the probability that North moves first is nonzero. In this contingency, the equilibrium analysis above would apply.

4 NGO Participation and Hard Law in International Institutions

One important lesson from this model pertains to the study of treaty compliance. All else constant, the model implies that the following hypotheses should hold. First, if negotiating states do not find hard law completely unacceptable, then NGO participation in negotiations should increase the probability of hard law. However, NGO participation should also reduce actual enforcement and implementation of hard law. By PROPOSITION 1, it is NGO participation that results in hard law. Second, if negotiating states are very worried about the cost of sanctions, then the probability of hard law should be low. This is established in PROPOSITION 2.

These hypotheses can inform studies of treaty compliance. Conventional statistical approaches are based on the assumption that states incorporate treaty provisions into treaties because these provisions help solve international cooperation problems (Abbott and Snidal, 1998; Koremenos, Lipson, and Snidal, 2001). We have shown that this assumption is doubtful under NGO participation. A serious test of treaty compliance, therefore, should separately analyze treaty negotiations with and without NGO participation. In the former case, we would expect hard law to be effective.

In the latter case, it should be less effective. By accounting for NGO participation, scholars of international law could avoid a pessimistic bias in their estimates of treaty effectiveness.

In this section, we illustrate these notions. While the examples above are suggestive of our model's validity, in this section we more systematically our key contention: does NGO participation play a key role in the inclusion of hard law in international institutions? We first examine the propositions derived above in light of evidence from the International Regimes Database (IRD) by Breitmeyer, Young, and Zuern (2006). We then examine those cases in which NGO participation occurred and sanctions were implemented, so as to examine whether this hard law was moribund or not.

The IRD establishes a common data protocol for use in aggregating a wealth of information originally identified by case studies of nearly all major international environmental regimes. The IRD is especially appropriate for our purposes because it is not singularly focused on design characteristics. Crucially, for our purposes, the IRD also records information about the bargaining context within which states have negotiated and designed environmental agreements. Equally important, the regimes described in the IRD contain a broad mixture of hard and soft law. Although the database breaks down regimes into a number of functionally and temporally defined "regime elements," not all of these can be mapped to specific international agreements. Following previous work with the database (Marcoux, 2009), we have chosen here to focus on the roughly fifty regime elements that correspond to the negotiation and adoption of a specific international environmental agreement.

In examining our propositions, we focus on three variables in particular: whether or not NGOs participated in negotiations, whether or not the resulting agreement provides for compliance monitoring, and whether or not the agreement uses potentially costly enforcement measures to promote compliance, instead of a "managerial" approach emphasizing capacity building (Chayes and Chayes, 1995).

We interpret the presence of NGOs during negotiations – whether as observers or as direct participants – as a proxy for the North's reputational value of NGO support, w. International environmental law was highly state-centric when these international agreements were negotiated, and to a large extent it remains so today. The participation of NGOs in the negotiation of an international environmental agreement is not required under international law. Rather, it reflects

the fact that states have made a conscious decision to allow their participation. Excluding NGOs from participation – even as mere observers – thus indicates that their reputational value in a given setting is low. Conversely, allowing NGOs to participate, when such participation is not required, signals that states do care about the reputational value of these organizations.[7]

The other two variables – provision for compliance monitoring and the use of sanctions in enforcement – allows us to distinguish between agreements with "hard law" mechanisms from those that use a more soft law approach. If an agreement includes either of these design features, it can be plausibly described as hard law. Since there is no universally agreed definition of what, exactly, constituted hard law, we have chosen to use two measures. Compliance monitoring reflects a relatively lower threshold for identifying hard law; the inclusion of enforcement sanctions reflects a relatively more strict definition of hard law.

If our theory is correct, we should expect to see systematic differences between the use of hard law mechanisms depending on NGO reputational value. As shown in FIGURE 2 above, when South's hard law cost c is low, we would expect to observe pooling up in any event. However, if South's hard law cost is high, the reputational value to North becomes of greater importance: when reputational value is high, we expect to observe mixed strategies. As reputational value declines, pooling down to soft law becomes a possibility. *Ceteris paribus*, we should therefore expect to see more frequent use of hard law mechanisms when reputational value is high. The question is whether the evidence from the IRD bears this out or not.

We constructed two frequency tables to determine whether we can reject the null hypothesis: that reputational value and hard law are distributed independently. TABLE 1 presents the distribution of NGO participation and the use of compliance monitoring (our less strict measure of hard law). The independent variable here is NGO participation during negotiations. Among the fifty cases for which we have compliance monitoring data, there were 29 agreements negotiated with NGOs present and 21 negotiated with NGOs absent. The 29 agreements negotiated in the absence of NGOs are almost evenly split on compliance monitoring. Fifteen agreements provide for compliance monitoring; fourteen do not. In contrast, we find that a clear majority of agreements, 15 of 21, that were negotiated in the presence of NGOs do provide for compliance monitoring. Although

[7]It is important to note here that the IRD codes expert groups, such as epistemic communities, separately from NGOs. We define NGOs narrowly to exclude these groups. This allows us to focus on reputational value alone, and to eliminate false positives, such as decisions to delegate to expert groups.

the evidence is suggestive, we should note here that we cannot reject the null hypothesis: the χ^2 test value for this crosstabulation is 1.97, with $p = 0.16$.

[Table 1 about here.]

We find a much stronger result using our stricter measure of hard law, as shown in TABLE 2. With respect to the design of enforcement sanctions, we find evidence of pooling down when reputational cost is low. Of the 30 agreements negotiated with NGOs absent, 28 do not provide for sanctions. In contrast, fully one-third of agreements negotiated with NGOs present incorporate formal sanction mechanisms. While this is not direct evidence of pooling up, it does reflect that hard law is much more common under NGO participation. Here, we can reject the null hypothesis that NGO participation and enforcement sanctions are distributed independently: the χ^2 is 6.04, with $p = 0.01$.

[Table 2 about here.]

Having established that NGO participation is positively associated with sanctions, we now examine whether this hard law is moribund or not. We examined all cases of NGO participation that resulted in the inclusion of sanctions, conducting a qualitative analysis of enforcement upon the entry into force of the relevant treaty. The results of this analysis are provided in TABLE 3. It lists those cases of treaty formation that occurred under NGO participation and resulted in a sanctioning mechanism. For each such observation, we first examined whether there was non-compliance. If not, then we deemed the sanctioning mechanism successful: the deterrent of sanctions simply prevented non-compliance. If there was non-compliance, we examined whether the sanctions were used to punish the defectors. In cases of failure to use the sanctions, we describe the hard law treaty in focus as *moribund*.

[Table 3 about here.]

Consistent with our theoretical expectations, we found instances of effective and moribund hard law. Probably the most effective hard law provisions were found in the London (1990) and Copenhagen (1992) Amendments to the Montreal Protocol on Substances That Deplete the Ozone Layer (1987). These amendments created a system that proscribes trade in the controlled substances. As such,

they encourage outsiders to join the treaty and deter members from violating their commitment to reduce the use of ozone-depleting substances. As Barrett (2003, 318) puts it, "[t]he benefit that non-signatories derive from free-riding is overwhelmed by the loss they suffer in being unable to trade with the majority of other countries." The fact that these trade sanctions have not been used does not imply, however, that they were ineffective. According to Barrett (2003, 320), "because no country will free ride, trade will not actually be restricted in equilibrium." In these cases, therefore, the absence of sanctioning is not to be equated with moribund hard law. The sanctions were effective enough to induce global participation and non-compliance has been minimal.

The MARPOL regime, or the International Convention for the Prevention of Pollution From Ships (1973, 1978), improves on an older marine oil pollution regime by imposing equipment standards on ships. It solves the difficult compliance problem of preventing oil discharges at the sea by requiring that oil tankers are equipped with systems that make illegal oil discharges unprofitable. Mitchell (1994, 452) notes that while compliance with the equipment standards has generally very good, "many states inspected tankers for compliance with equipment requirements" and some have even detained foreign ships that did not comply with these requirements. Thus, this is an example of effective sanctions inducing compliance and being used if necessary. Again, hard law is moribund.

The International Commission for the Conservation of Atlantic Tunas (1966) may not a particularly successful environmental treaty, as tuna populations continue decline due to overfishing. However, it is also not an example of moribund hard law. Desombre (1995, 56) examines the role of this convention in federal fisheries legislation of the United States. She notes that the 1975 Atlantic Tunas Convention Act "[a]uthorizes the Secretary of Commerce to prohibit imports of fish from a country" if such fishing would undermine the "recommendations by the International Commission for the Conservation of Atlantic Tunas."

The remaining cases paint a very different picture, though. The 1989 Lomé IV Agreement does contain a stringent provision banning trade in hazardous waste between the European Community and its partners. As a preferential trade agreement, Lomé IV can also enforce this provision by reducing trade concessions. But the state parties have blatantly violated the hazardous waste provision in Lomé IV (Clapp, 2001), yet we found no evidence whatsoever of any attempts to enforce it. Indeed, the hazardous waste provision disappeared from the 2000 Cotonou Agreement. This case illustrates our contention that NGO participation may result in moribund hard law.

The International Tropical Timber Agreement (1994) is a perfect illustration of moribund hard law. In principle, the treaty allows members to vote on the exclusion of any member from the treaty for undermining cooperation. This provision has never been used, though, and this is so in spite of the fact "countries do not comply with it – whether the 'it' be the guidelines or the criteria parties have adopted, reporting requirements, or even required payments" (Weiss, 1998, 129).

The La Jolla Agreement for the Reduction of Dolphin Mortality in the Eastern Pacific Ocean (1992) is an ambiguous case. It is not a formal treaty, and thus all provisions contained in it are ultimately non-binding. At the same time, though, the agreement does require that governments send observers on fishing vessels and prevent these fishing vessels from fishing should they exceed their dolphin mortality limit (Joseph, 1994, 11). In actuality, governments did send observers and recorded violations. Overall, however, the governments did not have any difficulties complying with their requirements: in 1993, allowed dolphin mortality was set at $19,500$ but the actual mortality was only $3,601$; by 1996, when allowed dolphin mortality had reduced to $9,000$, the actual mortality was only $2,547$ (Parker, 1999, 52). This striking gap indicates that compliance must have been very easy: countries achieved high levels of overcompliance. Evaluating the role or effectiveness of hard law is difficult in this case, for there is no evidence indicating that the observer program induced compliance. Thus, we conservatively code it as ambiguous. It was in 1998 that a legally binding agreement replaced the 1992 La Jolla Agreement.

Overall, we see that in 2/7 cases there is clear moribund hard law. In 4/7 cases, there is clear evidence of functioning hard law, and in 1/7 cases the role of hard law is unclear. These patterns are consistent with our expectations: under NGO participation, hard law provisions are often inserted in treaties yet they sometimes, but not always, become moribund.

5 Conclusion

Why do states often insist on including hard law in international institutions, and subsequently fail to enforce such hard law? We have advanced a theory of moribund hard law. It is a frivolous but rational form of informal governance. States often face domestic interest group pressure to negotiate hard law. If other states prefer soft law, they will only accept offers of hard law when they believe that enforcement is not forthcoming. This dynamic can explain the prevalence of moribund hard

law in international institutions.

This finding has important implications. Theoretically, the analysis highlights the importance of examining the rich variety of strategic incentives that states face in the design of international institutions. In addition to the collective action problems that classical cooperation theorists emphasize, state governments negotiate under domestic and transnational pressure (Putnam, 1988; Raustiala, 1997). Our game-theoretic analysis demonstrates that such pressures have dramatic effects on what constitutes a rational treaty design (Koremenos, Lipson, and Snidal, 2001).

It also underscores the value of examining interactions between formal rules and informal expectations in international rule, as proposed by Stone (2011) in his theory of "informal governance." Yet our approach deviates from conventional treatments of informal governance. Lipson (1991) and Stone (2011), for example, highlight that informality can facilitate international cooperation. In our model, on the contrary, informal expectations allow states to mislead domestic audiences into believing that they are willing to comply with international law. Characterizing the conditions under which states use informal governance for different purposes offers interesting opportunities for future research.

Empirically, our findings present major challenges for studies of treaty compliance. Traditionally, students of treaty compliance have sought to isolate the causal effect of a given treaty on state behavior (Simmons, 2010; Von Stein, 2005). Our analysis indicates that this approach could result in biased results. Scholars of treaty compliance should distinguish between serious and frivolous treaty provisions, and account for the international context of the treaties they analyze. We have proposed that this can be done by accounting for NGO participation. A rigorous test of the enforceability of international law should analyze treaties negotiated under NGO participation separately from those negotiated without NGO participation. Otherwise the empirical estimates confuse treaty ineffectiveness with deliberate failure to use international law.

Scholars of international institutions are not wrong to emphasize the functional dimension of treaty design. International institutions are not easy to negotiate and design, so their characteristics should not be attributed to irrational fads. This does not imply, though, that frivolous rules cannot be rational. In a strategic setting with pivotal states worrying about the cost of hard law, such hard law can only survive if it becomes moribund with some probability. Consequently, frivolity is a precondition for rationality.

Mathematical Appendix: Equilibrium Analysis

This appendix provides the formal details of the equilibrium analysis and proofs for the three propositions given in the main text. After introducing an equilibrium refinement, we characterize equilibrium behavior and posterior beliefs. We then prove the claims and proofs in the main text.

Equilibrium Refinement

Consider a deviation off the equilibrium path. The following restrictions are imposed:

1. If both types of North could benefit given the expected response, South and the NGO do not update their beliefs;

2. If only one type of North could benefit given the expected response, South and the NGO believe this type has deviated.

The former equilibrium refinement implies that South and the NGO cannot infer North's type from deviations that would benefit North regardless of its type. The latter equilibrium refinement implies that if South and the NGO see a deviation that could only benefit one type of North, they believe that North is of the profiting type. As the equilibrium analysis below shows, these refinements allow us to exclude substantively implausible equilibria that result from arbitrary belief updating off the equilibrium path.

Separating Equilibrium

The game has no separating equilibrium, so that the different types of North would play different pure strategies. Consider, first, the possibility that interested North would play $T^* = S$ while disinterested North would play $T^* = H$. In such an equilibrium, the following hold:

1. South accepts $T^* = S, H$;

2. Interested North enforces the hard obligation (off the equilibrium path) while disinterested North does not enforce it (on the equilibrium path).

3. The NGO would be able to perfectly identify North's type based on T^*.

To show this equilibrium cannot exist, consider a disinterested North's deviation by offering $T^* = S$. South accepts, and the NGO beliefs that $\mu^{int} = 1$. Thus, the expected payoff to North is $1 + w$. On the equilibrium path, a disinterested North's expected payoff is 1 because the NGO believes that only the disinterested type selects hard law. Thus, the deviation is profitable.

Suppose now interested North offers $T^* = H$ and disinterested North offers $T^* = S$. To prove that such an equilibrium does not exist, consider first disinterested North's deviation to $T = H$. South rejects, so disinterested North's payoff changes from 1 to w. Consider now interested North's deviation to $T = S$. South accepts, so interested, North's payoff changes from w to 1. Unless $q = 1$, one of these two deviations must be profitable. Thus, a separating equilibrium cannot exist outside a set of measure zero.

Pooling Up

Consider now the "pooling up equilibrium," so that North plays $T^* = H$ regardless of its type. In this equilibrium, the NGO's beliefs on the equilibrium path would be $\mu^{int} = p^{int}$ should South reject and otherwise depend on the signal q as described below.

First, suppose South rejects based on prior beliefs, as $p^{int}(1 - c) + (1 - p^{int}) < 0$. Now North's payoff from $T^* = H$ is wp given that the NGO's optimal action is $S^* = p^{int}$. North's deviation to $T^* = S$ off the equilibrium path does not induce South to change beliefs given that the payoff from that deviation is the same for both. Thus, the payoff from the deviation is $1 + p^{int}$. Clearly $1 + p^{int} > p^{int}$, so the equilibrium cannot exist.

Next, suppose South accepts based on prior beliefs, as $p^{int}(1-c)+(1-p^{int}) > 0$. North's payoff from $T^* = H$ is now $1 + w(1 - G)\frac{p^{int}}{p^{int}+(1-p^{int})(1-G)}$ for the disinterested type (given that it does not enforce): with probability G, the non-enforcement is revealed and the NGO learns for sure that North is of the disinterested type; with probability $1 - G$, it is not and the NGO cannot further update beliefs. For the interested type, which enforces, the payoff is $1 + I + w\frac{p^{int}}{p^{int}+(1-p^{int})(1-G)}$.

The NGO responds to deviations by updating posterior beliefs such that $\mu^{int} = 0$ because the interested type has more to lose. This yields $S^* = 0$, so the deviation is not profitable for either type of North.

Pooling Down

Consider now the "pooling down equilibrium," so that North plays $T^* = S$ regardless of its type. As to posterior beliefs, they would be $\mu^{int} = p^{int}$ on the path of play. In this equilibrium, the payoff to North is $1 + wp^{int}$ regardless of its type. If South accepts based on prior beliefs, deviation yields a payoff $1 + I + w$ for an interested North by the equilibrium refinement because the NGO believes that the North is interested. Thus, a deviation is profitable for both and the equilibrium implodes. If South does not accept based on prior beliefs, deviation yields wp^{int} for both types. Thus, whenever $wp^{int} \leq 1$ the equilibrium exists.

Mixed Strategies

Consider now equilibria in mixed strategies. First note that an interested North always obtains a higher payoff from $T^* = H$ than a disinterested North, so only one of the two can mix. Next note that if the disinterested North does not mix, instead playing $T^* = S$, then the interested North is the only type to ever play $T^* = H$. Thus, it follows that outside set of measure zero, South cannot be indifferent between accepting and rejecting hard law. Consequently, North cannot be indifferent between$T^* = H$ and $T^* = S$ outside a set of measure zero. South is also never indifferent between accepting and rejecting soft law, so it follows that the interested North cannot mix in equilibrium.

Suppose now that the disinterested North mixes between $T^* = H$ and $T^* = S$. Let λ^H denote the probability that it plays $T^* = H$. If $T^* = S$, South must believe that $\mu^{int} = 0$. If $T^* = H$, South's beliefs depend on the signal q as described below.

Clearly, in this case an interested North must play $T^* = H$ and South must accept $T^* = S$. For the disinterested North to be indifferent, we need

$$1 = \omega^{YES}\left((1 + w(1-G)\frac{p^{int}}{p^{int} + (1-p^{int})\lambda^H(1-G)}\right) + (1-\omega^{YES})\left(\frac{p^{int}}{p^{int} + (1-p^{int})\lambda^H}\right)$$

given the NGO's posterior beliefs. For South to be indifferent, we need

$$0 = p^{int}(1-c) + (1-p^{int})\lambda^H.$$

The second equation induces a unique λ^H, and the first equation thus induces a unique ω^{YES}. The

second equation also shows that it is necessary that the South would not accept based on prior beliefs. Thus, a unique equilibrium in mixed strategies exists.

Mathematical Analysis: Proofs

Proof of Proposition 1

As long as $p^{int}(1-c)+(1-p^{int}) < 0$, so that South accepts hard law based on prior beliefs, the above equilibrium analysis shows that a unique "pooling up" equilibrium exists. When p^{int} increases, the maximum c must decrease, as differentiation with respect to p^{int} shows. The conditional probability of enforcement is given simply by p^{int}, so it increases one-to-one with p^{int}. ■

Proof of Proposition 2

As long as $p^{int}(1-c)+(1-p^{int}) > 0$, so that South rejects hard law based on prior beliefs, and $wp^{int} < 1$, so that deviation is not profitable given the expected rejection, the above equilibrium analysis shows that a "pooling down" equilibrium exists. When p^{int} increases, the range of c and w increases so the region expands. ■

Proof of Proposition 3

As the equilibrium analysis shows, an equilibrium in mixed strategies cannot exist unless $p^{int}(1-c)+(1-p^{int}) > 0$. If this condition holds, the equilibrium analysis shows that a unique combination of λ^H, ω^{ES} exists that allows play in mixed strategies. As p^{int} increases, the range of c clearly increases.

As p^{int} increases, λ^H must clearly increase for South to be indifferent. Thus, the prior probability of a hard law offer, $p^{int}+(1-p^{int})\lambda^H$, increases. As c goes up, λ^H must also increase for South to be indifferent. Thus, the prior probability of a hard law offer, $p^{int}+(1-p^{int})\lambda^H$, increases. The conditional probability of actual enforcement is $\frac{p^{int}}{p^{int}+(1-p^{int})\lambda^H}$ and it must remain unchanged for any given c. As c increases, it must decrease. The probability that South accepts the proposal must decrease with the value of reputation w, to keep North indifferent.

The probability of a hard law proposal increases with the probability that North is interested and South's cost/ whereas the conditional probability of actual enforcement decreases. The probability that South accepts the proposal decreases with the reputational value w for the disinterested North to remain indifferent. ■

References

Abbott, Kenneth W., and Duncan Snidal. 1998. "Why States Act through Formal International Organizations." *Journal of Conflict Resolution* 42 (1): 3–32.

Abbott, Kenneth W., and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54 (3): 421–456.

Abbott, Kenneth W., Robert O. Keohane, Andrew Moravcsik, Anne-Marie Slaughter, and Duncan Snidal. 2000. "The Concept of Legalization." *International Organization* 54 (3): 401–419.

Audley, John J. 1997. *Green Politics and Global Trade: NAFTA and the Future of Environmental Politics*. Washington DC: Georgetown University Press.

Auer, Matthew. 2001. "Energy and Environmental Politics in Post-Corporatist Mexico." *Policy Studies Journal* 29 (3): 437–455.

Barrett, Scott. 2003. *Environment and Statecraft: The Strategy of Environmental Treaty-Making*. Oxford: Oxford University Press.

Bechtel, Michael M., and Jale Tosun. 2009. "Changing Economic Openness for Policy Convergence: When Can Trade Agreements Induce Convergence of Environmental Regulation?" *International Studies Quarterly* 53 (4): 931–953.

Benedick, Richard E. 1998. *Ozone Diplomacy: New Directions in Safeguarding the Planet*. Cambridge: Harvard University Press.

Bob, Clifford. 2005. *The Marketing of Rebellion*. New York: Cambridge University Press.

Börzel, Tanja A. 2000. "Why There Is No 'Southern Problem:' On Environmental Leaders and Laggards in the European Union." *Journal of European Public Policy* 7 (1): 141–162.

Breitmeyer, Helmut, Oran R. Young, and Michael Zuern. 2006. *Analyzing International Environmental Regimes: From Case Study to Database*. Cambridge, MA: MIT Press.

Cameron, Maxwell A. 1997. "North American Free Trade Negotiations: Liberalization Games Between Asymmetric Players." *European Journal of Internationa Relations* 3 (1): 105–139.

Chayes, Abram, and Antonia Handler Chayes. 1995. *The New Sovereignty: Compliance with International Regulatory Agreements*. Cambridge: Harvard University Press.

Clapp, Jennifer. 1994. "The Toxic Waste Trade with Less-Industrialised Countries: Economic Linkages and Political Alliances." *Third World Quarterly* 15 (3): 505–518.

Clapp, Jennifer. 2001. *Toxic Exports: The Transfer of Hazardous Wastes from Rich to Poor Countries*. Ithaca: Cornell University Press.

Dai, Xinyuan. 2005. "Why Comply? The Domestic Constituency Mechanism." *International Organization* 59 (2): 363–398.

Desombre, Elizabeth R. 1995. "Baptists and Bootleggers for the Environment: The Origins of United States Unilateral Sanctions." *Journal of Environment and Development* 4 (1): 53–75.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48: 147–160.

ENB. 2000. "Summary of the Sixth Conference of the Parties to the Framework Convention on Climate Change: 13-25 November 2000." *Earth Negotiations Bulletin* 12 (163).

Fearon, James D. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52 (2): 269–305.

Gartzke, Erik. 1999. "War Is in the Error Term." *International Organization* 53 (3): 567–587.

Gulbrandsen, Lars H., and Steinar Andresen. 2004. "NGO Influence in the Implementation of the Kyoto Protocol: Compliance, Flexibility Mechanisms, and Sinks." *Global Environmental Politics* 4 (4).

Hafner-Burton, Emilie M. 2005. "Trading Human Rights: How Preferential Trade Agreements Influence Government Repression." *International Organization* 59 (3): 593–629.

Husted, Bryan W., and Jeanne M. Logsdon. 1997. "The Impact of NAFTA on Mexico's Environmental Policy." *Growth and Change* 28 (1): 24–48.

Joseph, James. 1994. "The Tuna-Dolphin Controversy in the Eastern Pacific Ocean: Biological, Economic, and Political Impacts." *Ocean Development and International Law* 25 (1): 1–30.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca: Cornell University Press.

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

Koremenos, Barbara. 2005. "Contracting around International Uncertainty." *American Political Science Review* 99 (4): 549–565.

Koremenos, Barbara. 2011. "Open Covenants, Clandestinely Arrived At." Unpublished Manuscript.

Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55 (4): 761–799.

Kydd, Andrew H. 2000. "Trust, Reassurance, and Cooperation." *International Organization* 54 (2): 325–357.

Lipson, Charles. 1991. "Why Are Some International Agreements Informal?" *International Organization* 45 (4): 495–538.

Marcoux, Christopher. 2009. "Institutional Flexibility in the Design of Multilateral Environmental Agreements." *Conflict Management and Peace Science* 26 (2): 209–228.

McCall Smith, James. 2000. "The Politics of Dispute Settlement Design: Explaining Legalism in Regional Trade Pacts." *International Organization* 54 (1): 137–180.

Mearsheimer, John J. 1994-1995. "The False Promise of International Institutions." *International Security* 19 (3): 5–49.

Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103 (1): 144–181.

Mitchell, Ronald B. 1994. "Regime Design Matters: Intentional Oil Pollution and Treaty Compliance." *International Organization* 48 (3): 425–458.

Mitchell, Ronald B. 1999. "International Environmental Common Pool Resources: More Common than Domestic but More Difficult to Manage." In *Anarchy and the Environment*, ed. J. Samuel Barkin, and George E. Shambaugh. Albany, NY: State University of New York Press.

Najam, Adil. 2005. "Developing Countries and Global Environmental Governance: From Contestation to Participation to Engagement." *International Environmental Agreements* 5 (3): 303–321.

Oberthür, Sebastian, and Hermann Ott. 1999. *The Kyoto Protocol: International Climate Policy for the 21st Century*. Berlin: Springer-Verlag.

Parker, Richard W. 1999. "The Use and Abuse of Trade Leverage to Protect the Global Commons: What We Can Learn from the Tuna-Dolphin Conflict." *Georgetown International Law Review* 12 (1): 1–123.

Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 70 (1): 169–203.

Price, Richard. 2003. "Transnational Civil Society and Advocacy in World Politics." *World Politics* 55 (4): 579–606.

Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 44 (3): 427–460.

Raustiala, Kal. 1997. "States, NGOs, and International Environmental Institutions." *International Studies Quarterly* 41 (4): 719–740.

Ron, James, Howard Ramos, and Kathleen Rodgers. 2005. "Transnational Information Politics: NGO Human Rights Reporting, 1986-2000." *International Studies Quarterly* 49 (3): 557–588.

Rosen, Howard. 2004. "Free Trade Agreements as Foreign Policy Tools: The US-Israel and US-Jordan FTAs." In *Free Trade Agreements: US Strategies and Priorities*, ed. Jeffrey J. Schott. Washington, DC: Institute for International Economics.

Sell, Susan K. 1995. "Intellectual Property Protection and Antitrust in the Developing World: Crisis, Coercion, and Choice." *International Organization* 49 (2): 315–349.

Sell, Susan K. 1996. "North-South Environmental Bargaining: Ozone, Climate Change, and Biodiversity." *Global Governance* 2 (1): 97–118.

Shaffer, Gregory, and Mark A. Pollack. 2010. "Hard vs. Soft Law: Alternatives, Complements and Antagonists in International Governance." *Minnesota Law Review* 94 (3): 706–799.

Simmons, Beth A. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13: 273–296.

Stasavage, David. 2004. "Open-Door or Closed-Door? Transparency in Domestic and International Bargaining." *International Organization* 58 (4): 667–703.

Steinberg, Richard H. 1997. "Trade-Environment Negotiations in the EU, NAFTA, and WTO: Regional Trajectories of Rule Development." *American Journal of International Law* 91 (2): 231–267.

Steinberg, Richard H. 2002. "In the Shadow of Law or Power? Consensus-Based Bargaining and Outcomes in the GATT/WTO." *International Organization* 56 (2): 339–374.

Stone, Randall W. 2011. *Controlling Institutions: International Organizations and the Global Economy.* New York: Cambridge University Press.

Tolba, Mostafa K. 1998. *Global Environmental Diplomacy.* Cambridge, MA: MIT Press.

Urpelainen, Johannes. 2010. "Enforcement and Capacity Building in International Cooperation." *International Theory* 2 (1): 32–49.

Urpelainen, Johannes. 2011. "The Origins of Social Institutions." *Journal of Theoretical Politics* 23 (2): 215–240.

Von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99 (4): 611–622.

Wapner, Paul. 1995. "Politics Beyond the State: Environmental Activism in World Civic Politics." *World Politics* 47 (3): 311–340.

Weiss, Edith Brown. 1998. "The Five International Treaties: A Living History." In *Engaging Countries: Strengthening Compliance with InternationaL Environmental Accords*, ed. Edit Brown Weiss, and Harold K. Jacobson. Cambridge: MIT Press.

Young, Oran R. 1989. "The Politics of International Regime Formation: Managing Natural Resources and the Environment." *International Organization* 43 (3): 349–375.

Young, Oran R., and Michael Zürn. 2006. "The International Regimes Database: Designing and Using a Sophisticated Tool for Institutional Analysis." *Global Environmental Politics* 6 (3): 121–143.

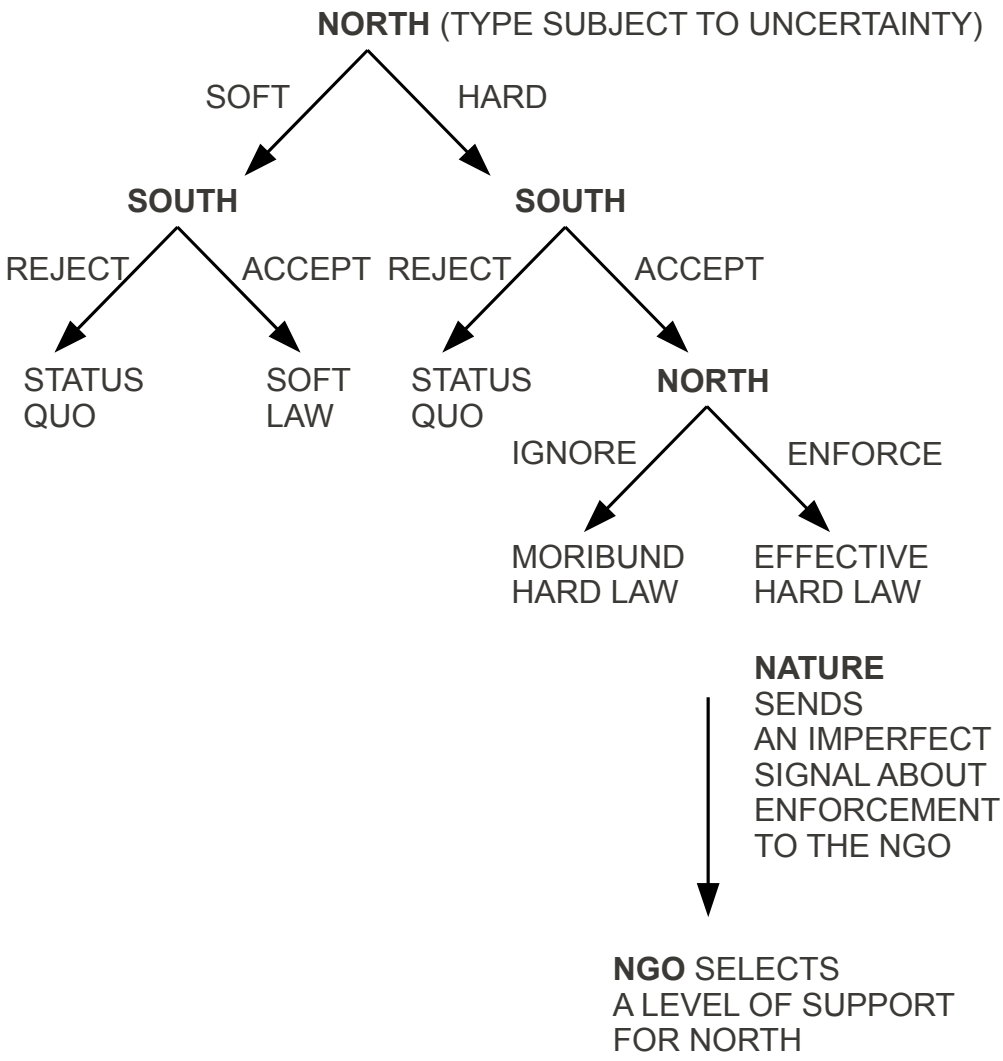

Figure 1: Sequence of moves.

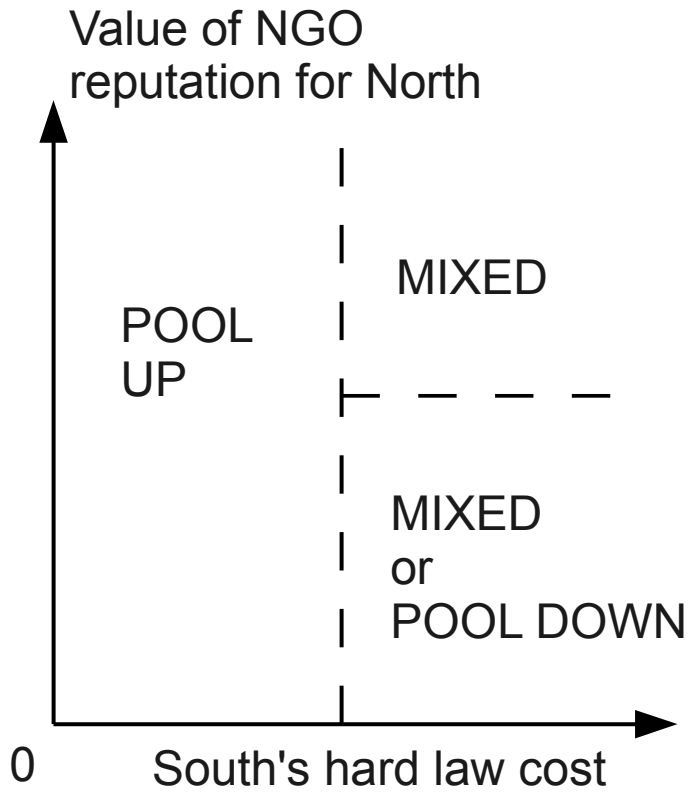

Figure 2: Graphical illustration of equilibria. If South's hard law cost c is low, "pooling up" follows and North offers hard law with certainty. If North's hard law cost c is high, either the equilibrium is in mixed strategies or "pooling down" follows with North offering soft law in all instances. The latter possibility requires that North's value of reputation is not too high.

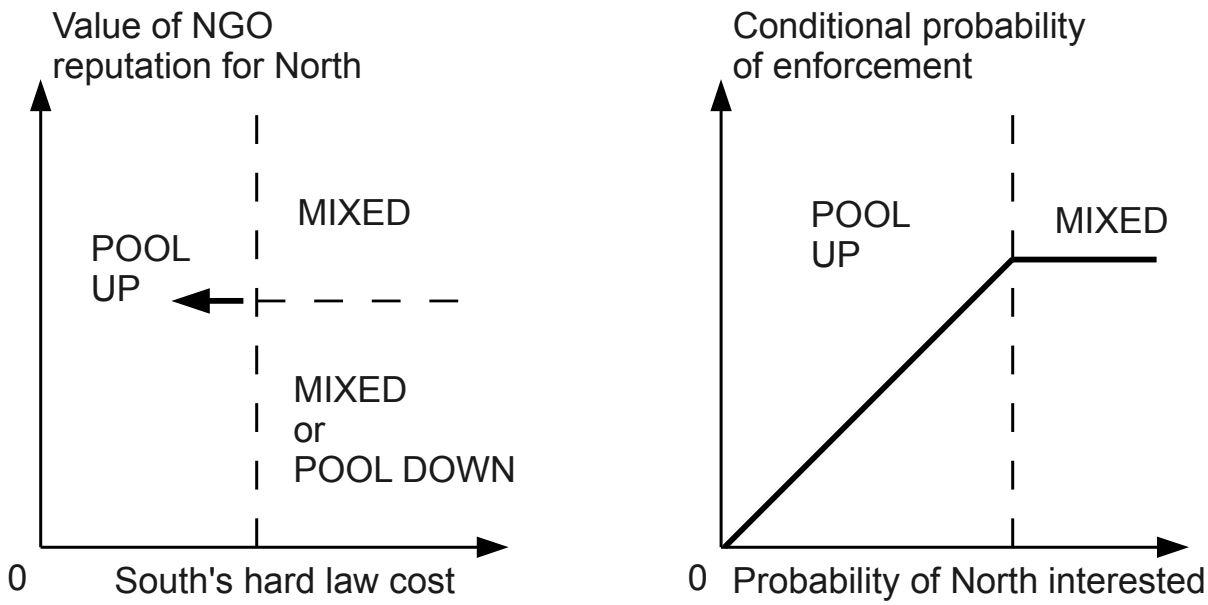

Figure 3: Graphical illustration of the effects of increasing the probability that North is interested, p^{int}. The left panel shows that an increase in p^{int} reduces the region for the "pooling up" equilibrium. The right panel shows that as p^{int} increases, the conditional probability of enforcement increases linearly until the "pooling up" equilibrium no longer exists. Beyond this point, the conditional probability remains unchanged.

		NGO Participation	
		No	Yes
Compliance	No	14	6
Monitoring	Yes	15	15

Table 1: Crosstabulation of NGO participation and monitoring measures.

		NGO Participation	
		No	Yes
Enforcement	No	28	14
Sanctions	Yes	2	7

Table 2: Crosstabulation of NGO participation and sanctioning mechanisms.

Treaty	**Issue**	**Year**	**Non-compliance**	**Moribund**
Lomé IV	Hazardous waste	1989	HIGH	*YES*
Tropical timber	Sustainable forestry	1994	HIGH	*YES*
Atlantic tuna	Fisheries	1966	HIGH	*NO*
Dolphins	Fisheries	1992	LOW	*UNCLEAR*
MARPOL	Sea oil pollution	1973, 1978	LOW	*NO*
London Amendment	Ozone depletion	1990	NONE	*NO*
Copenhagen Amendment	Ozone depletion	1992	NONE	*NO*

Table 3: Moribund hard law in treaties formed under NGO participation.