

The Customary International Law Game

George Norman^{*}
Joel P. Trachtman^{**}

Abstract

Structural realists in political science and some rationalist legal scholars argue that customary international law cannot affect state behavior: that it is “epiphenomenal.” This article develops a game theoretic model of a multilateral prisoner’s dilemma in the customary international law context that shows that it is plausible that states would comply with customary international law under certain circumstances. Our model shows that these circumstances relate to: (i) the relative value of cooperation versus defection, (ii) the number of states effectively involved, (iii) the extent to which increasing the number of states involved increases the value of cooperation or the detriments of defection, including whether the particular issue has characteristics of a commons problem, a public good, or a network good, (iv) the information available to the states involved regarding compliance and defection, (v) the relative patience of states in valuing the benefits of long-term cooperation compared to short-term defection, (vi) the expected duration of interaction, (vii) the frequency of interaction, and (viii) whether there are also bilateral relationships or other multilateral relationships between the involved states.

This model shows that customary international law is plausible in the sense that it may well affect state behavior where certain conditions are met. It shows what types of contexts, including malleable institutional features, may affect the ability of states to produce and comply with customary international law. This article identifies a number of empirical strategies that may be used to test the model.

^{*} William and Joyce Cummings Family Chair of Entrepreneurship and Business Economics, Tufts University.

^{**} Professor of International Law, The Fletcher School of Law and Diplomacy, Tufts University.

The authors express appreciation to those who made suggestions or commented on earlier drafts, including Kenneth Abbott, Jose Alvarez, Lucian Bebchuk, David Bederman, Tomer Broude, Marc Busch, Jeffrey Dunoff, Michael Glennon, Ryan Goodman, Larry Helfer, Moshe Hirsch, Robert Hockett, Ian Johnstone, Louis Kaplow, Patrick Kelly, David Kennedy, Barbara Koremenos, Lisa Martin, Sean Murphy, Philip Nichols, Eric Posner, Mark Ramseyer, Kal Raustiala, Alfred Rubin, Steven Shavell, Beth Simmons, Peter Spiro, Richard Steinberg, Paul Stephan, Edward Swaine, Detlev Vagts, six anonymous referees and participants in seminars where earlier versions of this article were presented, at Harvard Law School, Hebrew University, the Max Planck Institute for Comparative Public Law and International Law, UCLA Law School and the Wharton School. We also thank Alexander Gazis and Christine Makori for research assistance. Errors are ours.

1. Introduction

Customary international law (CIL) is under attack as behaviorally epiphenomenal and doctrinally incoherent. By contrast, the central argument of this article is that CIL, while something of a feat of levitation, is theoretically plausible in the sense that it may well affect state behavior in certain circumstances. CIL is a feat of levitation insofar as it rests not on a rock-solid natural law basis of divine principles, but on a fabric of rational acts, woven through a multiplicity of relations over time. Our argument makes one central claim: that while there are limits on and variations in the effectiveness of CIL, there are circumstances where it may independently affect the behavior of states. There is no reason in theory, or in data adduced by others, to believe CIL generally epiphenomenal. As CIL is the foundation of all international law, this article suggests the circumstances under which we would expect international law to affect state behavior.

This article refines and extends an emerging rationalist understanding of CIL.¹ Pioneering work in this field, notably that of Jack Goldsmith and Eric Posner, has begun to articulate a rationalist theory of CIL. Goldsmith and Posner have argued that CIL does not affect state behavior.² We show why this assertion is either tautological or incorrect as a matter of theory, and, to the extent that it purports to rely on factual observation, is not supported by the data presented.³ We construct a more complete model that shows that Goldsmith and Posner's assumption that the multilateral prisoner's dilemma is unlikely to allow CIL to affect state behavior is false.

The tautology of the Goldsmith-Posner argument arises from a false dichotomy between motivation by self-interest⁴ and motivation by law. In a rationalist model, behavior is *assumed* to be motivated by self-interest. If law is artificially separated from self-interest, then of course a rationalist model would assume that law has no motivating force. Yet we show how CIL rules may modify the payoffs associated with relevant behavior, and thereby affect behavior through self-interest. CIL may affect behavior,

¹ See, e.g., Mark A. Chinen, *Afterword*, 23 MICH. J. INT'L L. 201 (2001); Mark A. Chinen, *Game Theory and Customary International Law: A Response to Professors Goldsmith and Posner*, 23 MICH. J. INT'L L. 143 (2001); Vincy Fon & Francesco Parisi, *International Customary Law and Articulation Theories: An Economic Analysis*, George Mason Law and Economics Research Paper No. 02-24, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=335220; Jack L. Goldsmith & Eric A. Posner, *Further Thoughts on Customary International Law*, 23 MICH. J. INT'L L. 191 (2001); Jack L. Goldsmith & Eric A. Posner, *A Theory of Customary International Law*, 66 U. CHI. L. REV. 1113 (1999) [hereinafter Goldsmith & Posner 1999]; Andrew T. Guzman, *A Compliance-Based Theory of International Law*, 90 CAL. L. REV. 1823 (2002); Francesco Parisi, *The Formation of Customary Law*, George Mason Law and Economics Research Paper No. 01-06, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=262032; Edward T. Swaine, *Rational Custom*, 52:3 DUKE L.J. 559 (2002); Pierre-Hugues Verdier, *Cooperative States: International Relations, State Responsibility and the Problem of Custom*, 42 VA. L. REV. 839 (2002).

² The leading article here is Goldsmith & Posner 1999, *supra* note 1. For a self-described "traditionalist's" response, see Detlev F. Vagts, *International Relations Looks at Customary International Law: A Traditionalist's Defence*, 15 EUR. J. INT'L L. 1031 (2004).

³ See Vagts, *supra* note 2 (critiquing the use of examples by Goldsmith & Posner).

⁴ By "self-interest" we mean merely to refer to maximization of preferences—these preferences could be other-regarding or altruistic.

even if it only does so at the margins. While CIL is endogenous to states as a group, meaning that it is not a vertical structure produced outside or above the group of states, it is an independent (exogenous) influence on the behavior of each individual state.

Goldsmith and Posner also assert that CIL is internally incoherent, as states are not motivated by *opinio juris* to comply with CIL. However, our analysis provides a plausible basis to assign a discriminating role to *opinio juris*, and therefore we find CIL doctrine internally coherent in at least its core dimension. We understand *opinio juris* as a way of referring to the intent of states to propose or accept a rule of law that will serve as the focal point of behavior, will implicate an important set of default rules that are applicable to law and not to other types of social order, and that will implicate an important set of linkages among legal rules.

This article develops a repeated multilateral⁵ prisoner's dilemma⁶ model of CIL. Of course, game theory can never capture all real-world detail with its highly nuanced decision-making.⁷ The purpose of game-theoretic models is not to predict or prescribe behavior, but to generate testable hypotheses that, once tested, are expected to tell us something useful about the world.⁸

Our analysis focuses on the parameters of the multilateral prisoner's dilemma in the CIL context. These parameters include: (i) the relative value of cooperation versus defection, (ii) the number of states effectively involved, (iii) the extent to which increasing the number of states involved increases the value of cooperation or the detriments of defection, including whether the particular issue has characteristics of a commons problem, a public good, or a network good,⁹ (iv) the information available to the states involved regarding compliance and defection, (v) the relative patience of states in valuing the benefits of long-term cooperation compared to short-term defection, (vi) the expected duration of interaction, (vii) the frequency of interaction, and (viii) whether there are also bilateral relationships or other multilateral relationships between the involved states.

⁵ In this context, we use the term "multilateral" to refer to contexts in which more than two parties are involved. A more technical term would be "n-person."

⁶ The "prisoner's dilemma" is a form of game commonly used in social science analysis. In this article, it serves as a vehicle for providing a formalized analogy to the CIL context. The structure and assumptions that characterize the prisoner's dilemma are described in parts 2 and 3 below.

⁷ See the cautions expressed in Goldsmith & Posner 1999, *supra* note 1, and the broader treatment in Duncan Snidal, *The Game Theory of International Politics*, in Kenneth A. Oye, ed., *COOPERATION UNDER ANARCHY* (1986).

⁸ While it may be argued that the game theory that we use, initially developed in the context of analysis of individual behavior, cannot be applied to state behavior, it should be noted that game theory has been applied to behavior of firms, as well as states. For an argument regarding the adaptation of these types of models to international law, see Jeffrey L. Dunoff & Joel P. Trachtman, *The Law and Economics of International Law*, 24 *YALE J. INT'L L.* 1 (1999).

⁹ "Commons problems" are circumstances where persons share a particular resource, and one person's use of the resource may reduce the amount of the resource available to others, while conservation of the resource may increase the amount of the resource available to all. A fishery may present a commons problem in this sense. "Public goods" are goods that are available to all, where one person's use of the good does not diminish its availability to others. Networks in this context involve circumstances where wider use of the same good makes the good more valuable to all. Standards often have this characteristic.

The parameters we identify are incorporated in our model as independent variables, but from a normative standpoint it is possible for policy initiatives to select or manipulate these parameters. That is, by identifying the parameters for determining whether CIL will affect state behavior, this article opens the way to normative institutional design. States may determine to restructure certain institutions in order to facilitate the formation and operation of CIL. There may be circumstances under which it will be normatively attractive to facilitate the development of CIL, rather than to engage in more self-conscious and static treaty-making. The institutional dynamism and social immanence that make social norms attractive in the domestic context may also be attractive in at least some international contexts.

This article is organized as follows. The remainder of this introduction provides a short doctrinal review of CIL and briefly locates this article in relation to four literatures: law and economics, social norms, international organization and industrial organization. Our model is largely based on earlier work in these areas. Part 2 explains our choice of the repeated multilateral prisoner's dilemma as the basis for our model, as well as our choice of an assumed strategy for players within this model. Any game theory model depends on decisions to assume certain game structures and strategies. We explain here why the prisoner's dilemma game provides a good fit for the CIL context, and how we choose an assumed strategy for states to follow. Part 3 explains the other assumptions and parameters of our model. Part 4 sets forth four illustrative examples of CIL contexts that might be understood in terms of our model. Part 5 sets forth the general structure and implications of our model, while Appendix I sets forth the formal model. Part 6 presents some implications of this article. Part 7 briefly concludes.

In brief, this article has important implications regarding the plausibility of CIL, the possibility of institutional change to facilitate CIL in particular areas, our understanding of CIL doctrine, and future research in CIL. On the basis of a complete rationalist model, articulated in formal terms in Appendix I, we show the parameters that will determine states' compliance with CIL. These parameters include especially (i) the relative patience of individual states to wait for the benefits of reciprocal compliance by other states, (ii) the importance of the frequency of interaction and of the duration of interaction in order to increase the potential value of future reciprocal benefits, (iii) the possibility that compliance in one setting may be supported by overlapping relationships in other settings, (iv) the structure of the benefits from cooperation, including whether benefits may increase or decrease with the number of states that cooperate. Based on this understanding, states may create institutional responses in order to enhance the possibilities of compliance with particular rules of CIL.

But CIL is not merely a label for rational cooperation. Rather, CIL is a special branch of cooperation that has particular features. These features include the establishment of a focal point so that states may readily identify what will "count" as cooperation, attention to the motivation of states to offer and accept a rule of law (*opinio juris*), with certain default rules and prescribed consequences, and the linkage of the particular rule with the broader international legal system. In this last regard, we might

say that by including a particular rule in CIL, states are accepting that the rest of the legal system is hostage to compliance with the particular rule. This adds strong incentives for compliance. Yet this is only the beginning of a rationalist research program in CIL. Theory such as the one articulated in this article must be tested and refined based on empirical observation. By articulating a set of parameters for determining the likelihood of compliance, this article also suggests the contours of a broader research program in CIL.

a. Customary International Law and its Doctrine

As an introductory matter, it is useful to review the fundamental doctrine of CIL. Article 38(1)(b) of the Statute of the International Court of Justice, listing the sources of international law applicable by the Court, includes “international custom, as evidence of a general practice accepted as law.” Article 102 of the American Law Institute’s Restatement (Third) of Foreign Relations Law states that “[c]ustomary international law results from a general and consistent practice of states followed by them from a sense of legal obligation.” The sense of obligation is referred to in Latin as “*opinio juris sive necessitatis*.”

CIL has two core doctrinal problems relating to *opinio juris*. First, can CIL ever come into existence if it requires *opinio juris*—a sense of legal obligation—before it can exist? The second, related, problem relates to the motivation of states. Are states ever motivated by *opinio juris*? The rationalist assumption that states are motivated by their own preferences seems *prima facie* inconsistent with the CIL requirement of motivation by *opinio juris*, and this is the criticism that Goldsmith and Posner level at CIL doctrine.¹⁰ They argue that what appears to be CIL is not, because it is actually motivated by coincidence of interest, coercion, cooperation predicated on self-interest, or coordination predicated on self-interest. We respond to this argument in part 6.

As suggested above, for the social scientist studying law, the critical descriptive question relates to the effects of legal rules on behavior. For CIL, this descriptive question is also a doctrinal question, as CIL doctrine requires some level of generality and consistency of practice—some quantum of state behavior. And again, at least under the Restatement formulation, this behavior must be motivated by *opinio juris*. Under the Statute of the International Court of Justice, the custom itself may serve as evidence of acceptance as law, of *opinio juris*. Some commentators have suggested that *opinio juris* in a formal sense may not be necessary at all, but that the requirement should be understood in terms of state consent or acceptance.¹¹

As we address the theory of CIL, it is useful to have in mind some examples of CIL. CIL exists in a wide variety of fields, and coexists in many areas with treaty law. As already noted, CIL is its own foundation. Thus, there is a CIL of CIL. In other words, the CIL doctrine discussed above is understood as law pursuant to CIL.

¹⁰ Goldsmith & Posner 1999, *supra* note 1.

¹¹ Maurice H. Mendelson, *The Formation of Customary International Law*, 272 RECUEIL DES COURS 155, 268-293 (1998).

Furthermore, the law of treaties came into existence pursuant to CIL, although this law is today codified in treaty. But moving beyond these foundational areas, CIL addresses, for example, issues of territorial integrity, jurisdiction to apply law, diplomatic immunity, the rights of states with respect to coastal areas, human rights, cross-border pollution and the use of force to settle international disputes.

These are diverse fields, each with its own characteristics. Our theoretical approach calls for discrimination among these fields: not all CIL is created equal. Not all law is equally or peremptorily binding. Our model indicates the parameters for discrimination. Simply put, we would expect greater possibility for formation of and compliance with CIL in some fields than in others. Of course, this is an area in which empiricism would require analysis of areas in which CIL has *not* developed: the dog that did not bark. Thus, we might ask, why is there no CIL rule that requires extradition of criminals?

Our theoretical approach also accepts the possibility of linkage among diverse fields, integrating and therefore homogenizing the behavioral effect of each rule of law.

b. Four Literatures

This article draws on four semi-autonomous literatures. First, as noted above, there is an emerging rationalist, law and economics-based literature of CIL. The leading work in this area is by Goldsmith and Posner, but there are other important contributions. Second, this article draws on a burgeoning literature on social norms in the law, although social norms are studied by all manner of social scientists. Third, this article draws on the economics field of industrial organization for game theory-based insights about collusion among competitors in markets. Fourth, this article draws on the political science literature of international organization, which has addressed in detail the game theoretic analysis of cooperation among groups of states.

i. Law and Economics of CIL

Goldsmith and Posner provide a game theoretic analysis of CIL. They examine a variety of CIL circumstances, and argue that these circumstances can be categorized into four game types: (i) coincidence of interest, (ii) coercion, (iii) bilateral cooperation, and (iv) bilateral coordination. This is a useful exercise, as it invites us to consider the motivation of states, and the degree to which CIL affects behavior. In the perhaps hypothetical cases of pure coincidence of interest and pure coercion, Goldsmith and Posner are correct that there can be no *opinio juris*, and that law does little work. This is not new to CIL doctrine, but it is useful to emphasize. It is also worth noting that Goldsmith and Posner seem to assume a purity of motivation that may not exist in the real world.

Goldsmith and Posner suggest that many instances of observed CIL may be understood in terms of bilateral cooperation along the lines of a bilateral prisoner's dilemma game. They then argue that "Although game theory does not rule out the

possibility of n-state cooperation, the assumptions required for such an outcome are quite strong and usually unrealistic. For this reason, we doubt the utility of n-player prisoner's dilemmas as an explanation for multilateral or 'universal' behavioral regularities."¹² Their views with respect to coordination games are similar. The present article shows that the assumptions for multilateral or "n-state" cooperation are neither strong nor unrealistic, but are in some cases quite plausible.

In addition to developing this theoretical perspective, Goldsmith and Posner examine several examples of CIL. The areas they consider are neutrality, diplomatic immunity and maritime jurisdiction. They believe that in these areas, states were motivated by coincidence of interest, coercion or a bilateral reciprocity along the lines of the prisoner's dilemma. They conclude that if state behavior can be explained by coincidence of interest or coercion, or any other self-interested-motivation, then neither *opinio juris* nor CIL does any motivational work.

ii. Social Norms

We might ask, however, whether the CIL problem is different in structure from the social norms context, and whether if social norms can affect behavior, CIL can as well. Since the publication by Robert Ellickson of *Order Without Law*¹³ in 1991, legal scholars have examined the role of informal norms in society, and the relationship of these norms to law. Ellickson investigates how cattle farmers in Shasta County, California, manage to establish and apply their own non-legal rules, with a notable level of compliance, without direct intervention by the state. It is an insightful story about how order can arise without law, or in spite of law.¹⁴

We may draw a rough, and limited, analogy between the development of social norms in a municipal, or private setting,¹⁵ and the development of CIL in the international public setting. In the international community, CIL is substantively similar to the phenomenon Ellickson describes.¹⁶ In international political science, regime theorists such as Robert Keohane,¹⁷ Stephen Krasner, and Beth Simmons¹⁸ have told a similar story of the possible rise of order in international society. However, regime theory has

¹² Goldsmith & Posner 1999, *supra* note 1, at 1130.

¹³ ROBERT ELLICKSON, *ORDER WITHOUT LAW* (1991).

¹⁴ However, one might argue that the general legal system, including especially its rules against violence, forms an important background or infrastructure that may provide support to the farmers' social norms.

¹⁵ For a recent work synthesizing and extending some of the social norms learning, see ERIC A. POSNER, *LAW AND SOCIAL NORMS* (2000). See also Richard H. McAdams, *Book Review: Signaling Discount Rates: Law, Norms and Economic Methodology*, 110 YALE L.J. 625 (2001).

¹⁶ For an example of this type of analogical allegory, comparing domestic custom to international custom, see Mendelson, *supra* note 11, at 165-168. For an early statement that CIL is produced in an evolutionary fashion, see ANTHONY A. D'AMATO, *THE CONCEPT OF CUSTOM IN INTERNATIONAL LAW* 104 (Cornell 1971).

¹⁷ See, e.g., ROBERT O. KEOHANE, *AFTER HEGEMONY: COOPERATION AND DISCORD IN THE WORLD POLITICAL ECONOMY* (1984); ROBERT O. KEOHANE, *INTERNATIONAL INSTITUTIONS AND STATE POWER: ESSAY IN INTERNATIONAL RELATIONS THEORY* (1989).

¹⁸ See, e.g., STEPHEN D. KRASNER, *INTERNATIONAL REGIMES* (1983); Stephen D. Krasner & Beth Simmons, *Theories of International Regimes*, 41 INT'L ORG. 491 (1987).

generally avoided CIL.¹⁹ Moreover, recognition that a rule has become part of CIL may signal its support by, or linkage with, the multi-sector international legal system. This system of accepted and enforced linkage may distinguish legal rules from non-legal regimes.

The difference between law and social norms in the municipal setting is that law is the province of the state (setting aside for the moment religious law, other non-state rules, and circumstances in which non-state made rules are incorporated in the state-enforced law).²⁰ However, this distinction is inapposite to the international system, which has been characterized as a horizontal, as opposed to vertical, system, where there is no overarching state, *per se*.²¹ So, in the international system, there is more overlap, and an indistinct border, between law and social norms. This overlap is perhaps easier to see in the international context than in the domestic context, as, in the international context, a significant subset of social norms is termed "law."

Ellickson states that the social norms literature defines a "social norm" as "a rule governing an individual's behavior that third parties other than state agents diffusely enforce by means of social sanctions."²² The focus of this definition on decentralized means of enforcement shows the strong analogy between social norms in the municipal setting and CIL in the international setting. Of course, to the extent that international courts may apply, and institutions of global governance may enforce, CIL, there is a difference. But this application and enforceability is quite limited. There are few circumstances in which CIL rules benefit from mandatory adjudication in *international* tribunals. We would not consider application of CIL by *domestic* courts to amount to the action of "state agents" *at the international level*, although action by domestic courts would certainly be considered action of "state agents" *at the municipal level*. This is because in the international context, *domestic* courts are simply internal deliberative processes of national governments. The application by domestic courts of CIL may be understood as a kind of norm internalization.²³

¹⁹ Duncan Snidal, *Political Economy and International Institutions*, 16 INT'L REV. L. & ECON. 121, 124 (1996).

²⁰ While there is no state at the global level, there is an international legal and organizational order, which is quite a bit more fragmented than most nation-states.

²¹ We add this qualification, because one might argue that the CIL and conventional law framework, as it exists, is at least comparable to a municipal state, or at least that this framework is comparable to the constitution that a municipal state might have. Of course, while it responds to some of the same questions, including a rule of recognition, it is not as detailed or fertile as a constitution.

²² Robert Ellickson, *The Evolution of Social Norms: A Perspective from the Legal Academy*, in SOCIAL NORMS 35 (Michael Hechter & Karl-Dieter Opp eds., 2001). Note that Ellickson assumes multilateral, as opposed to bilateral, retaliation.

²³ See Harold Hongju Koh, *Why Do Nations Obey International Law?*, 106 YALE L.J. 2599 (1997) (reviewing Abram Chayes & Antonia Handler Chayes, *THE NEW SOVEREIGNTY: COMPLIANCE WITH INTERNATIONAL REGULATORY AGREEMENTS* (1995)). This internalization may be desirable under certain circumstances, and may be developed as a tool of enforcement of CIL. That is, states may persuade one another to harness their domestic legal system for certain of their international legal obligations, effectively holding the domestic legal system hostage. See also Philip Moremen & Joel P. Trachtman, *Whose Right is it Anyway? Private Parties in EC-U.S. Dispute Settlement at the WTO*, 44 HARV. INT'L L. J. 221 (2003).

One important set of explanations of social norms understands norms as preferences that individuals (or in our case states) acquire through education, acculturation or other processes, such as an expressive or articulation function.²⁴ It may well be possible to modify preferences of states through social norms. It seems obvious that the way to modify the preferences of states is to modify the preferences of individual government officials, or voters. Indeed, it may be appropriate to consider epistemic communities and networks among government officials as channels of preference-modification. While this approach may have traction in the CIL setting, and may re-emphasize the role of arguments predicated on legitimacy, justice and morality as bases for preference-modification, this article will bypass that discussion, and focus on exogenous explanations of social norms, that is, explanations that look at external influences on actors, rather than internal changes in actors.²⁵ Modeling always involves simplification, and our goal in this article is to elaborate a rationalist model for future testing. In fact, testing an exogenous model would help to advance the debate between exogenous and endogenous causes of compliance with CIL.

In connection with exogenous explanations, the law-based social norms literature has not embraced the repeated multilateral prisoner's dilemma.²⁶ This is due to two types of concerns. First, there are concerns that game theory does not reflect the nuance of social interaction. We hope that the incorporation in our model of multi-sector contact helps to address this concern. Second, there are concerns regarding the credibility of third-party enforcement.²⁷ Will third parties be properly motivated to join in retaliation against violators? If the threat of retaliation is not credible, there will be strong incentives for violation. We address these concerns below.

iii. Industrial Organization

Much of our understanding of the utility, structure and dynamics of multilateral prisoner's dilemma games comes from the economics literature of industrial organization. This literature considers the possibility that firms may enter into cartels or other restrictions of competition that violate antitrust laws. While firms may find opportunities to communicate, their communications and agreements must be kept secret from the regulatory authorities and are not enforceable at law. This legal restraint on the enforcement of agreements is analogous to the limitation in the international law setting on enforcement of agreements. An important concern for industrial organization

²⁴ See Robert D. Cooter, *Expressive Law and Economics*, 27 J. LEG. STUDS. 585 (1998).

²⁵ Robert D. Cooter, *Do Good Laws Make Good Citizens? An Economic Analysis of Internalized Norms*, 86 VA. L. REV. 1577 (2000); Robert C. Cooter, *Models of Morality in Law and Economics: Self-Control and Self-Improvement for the "Bad Man" of Holmes*, 78 B.U. L. REV. 903, 911-912 (1998); JON ELSTER, *THE CEMENT OF SOCIETY: A STUDY OF SOCIAL ORDER* (1989). Basu refers to these as "preference-changing norms." See Kaushik Basu, *Social Norms and the Law*, 3 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 476 (1998).

²⁶ See Paul G. Mahoney & Chris William Sanchirico, *Norms, Repeated Games, and the Role of Law*, 91 CAL. L. REV. 1281, 1284 (2003). Mahoney and Sanchirico explain the state of the social norms literature with respect to the multilateral prisoner's dilemma. They explain that the objection to these models is that third party enforcement is not individually rational: that the players lack incentives to retaliate.

²⁷ *Id.* at note 12, citing work by Ellickson, Katz, McAdams and Posner.

economists is to identify circumstances under which agreements can be made self-sustaining through the self-interest of the parties to the agreement.

While the analogy is apparent, we must recognize, of course, that in the CIL setting, public communication is possible, as are agreements that at least purport to be binding: treaties. The degree to which agreements may actually constrain behavior is a question for another article, but we point out that the international legal rule that treaties must be observed (*pacta sunt servanda*) is itself part of CIL.

Another, perhaps more important, distinction is that a cartel has certain characteristics that may differ from any particular CIL setting. That is, in a cartel, the more others adhere to the cartel, the greater the monetary incentives for any particular member to defect. This context is more like a commons problem than like a public good or network problem. We will discuss some of these distinctions based on variations in payoff structures below.

iv. International Organization

Political scientists and economists working in the field of international organization have made a good deal of progress in analyzing the problem of international cooperation more generally. In various works, they have examined most of the parameters that we utilize here.²⁸ This literature has not examined CIL. The model we present here formalizes certain considerations that remain informal in much of the political science literature, and has other distinctive features that we describe below.

We have structured our model to match most closely the CIL context, rather than attempt to structure a model that would address other international cooperation devices. However, we recognize that the question of which device—treaty, CIL or softer law—itself depends on a set of variables. We also recognize that CIL may be understood as a phase in the formalization of law, or in “legalization.”²⁹ The International Law Commission of the United Nations often codifies CIL, and CIL often forms the basis for treaties. This article does not present an explanation of choice between custom and treaty, or of the relationship among these instruments.³⁰

²⁸ There is disagreement between institutionalists and “realists,” who claim that states’ interests in international relations are characterized by a search for relative gains, rather than absolute gains. These realists reject the possibility of cooperation where it results in relative gains to a competitor. See Marc Busch & Eric Reinhardt, *Nice Strategies in a World of Relative Gains: The Problem of Cooperation under Anarchy* 37 J. CONFL. RES. 427 (1993); Robert Powell, *Absolute and Relative Gains in International Relations Theory*, 85 AM. POL. SCI. REV. 1303 (1991); Duncan Snidal, *Relative Gains and the Pattern of International Cooperation*, 85 AM. POL. SCI. REV. 701 (1991).

²⁹ See the special issue of International Organization devoted to the phenomenon of “legalization.” 54:3 INT’L ORG. (2000).

³⁰ See Kenneth Abbott & Duncan Snidal, *Hard and Soft Law in International Governance*, 54 INT’L ORG. (2000); John K. Setear, *Treaty, Custom, Iteration and Public Choice*, February 2004, available at <http://ssrn.com/abstract=492604> (arguing that custom is more attractive to executive branches). To the extent that CIL is less detailed—less specific—than treaty norms, it is amenable to a rules versus standards type of analysis. See Louis Kaplow, *Rules versus Standards: An Economic Analysis*, 42 DUKE L.J. 557,

2. The Multilateral Prisoner's Dilemma

Game theory develops models to depict, to abstract, and to formalize, various social settings. Given the diversity of social contexts, there is a wide choice of models from which to choose.³¹ In this section we explain our choice of the multilateral prisoner's dilemma as the basic model for the CIL problem, examine the potential strategies that might be played within this game, and explain some of our assumptions. The multilateral repeated prisoner's dilemma is itself really a group of varying models, with a number of varying features, including the number of players and their preferences.

a. Choice of Game

The basic payoff structure assumed in the prisoner's dilemma game captures the essential problem of cooperation in a horizontal social setting with externalities,³² in which parties have a choice between compliance and defection, and in which they can enrich themselves individually through defection while they could enrich society, and in light of the anticipated actions of others, themselves, through compliance.³³ Of course, some CIL contexts might be better modeled using other methods, but by using the prisoner's dilemma we hope to capture the essence of informal contracting under opposed interests.³⁴

One of the reasons that we choose this game is because it allows us to contextualize a number of insights and concerns that cannot easily be included in other analytical models. For example, we believe that the multilateral prisoner's dilemma can

577 (1992); Joel P. Trachtman, *The Domain of WTO Dispute Resolution*, 40 HARV. INT'L L.J. 333, 346-47 (1999).

³¹ For a useful analysis of the "fit" of other games, including "battle of the sexes" and "stag hunt," see Swaine, *supra* note 1. See also Fiona McGillivray & Alastair Smith, *Trust and Cooperation Through Agent-Specific Punishments*, 54 INT'L ORG. 809, 810 (2000) (noting that the prisoner's dilemma is often used to model international cooperation).

³² By use of the term "externalities," we mean to be quite inclusive, including both pecuniary and non-pecuniary externalities: any circumstance in which an action by one state has adverse or beneficial effects on another state.

³³ At another level of complexity, it would be possible to model the game of formation of a CIL rule separately from enforcement. See James D. Fearon, *Bargaining, Enforcement, and International Cooperation*, 52:2 INT'L ORG. 269 (1998) (separating the bargaining problem, modeled as a coordination game, from the enforcement problem, modeled as a prisoner's dilemma); Stephen D. Krasner, *Global Communications and National Power: Life on the Pareto Frontier*, 43 WORLD POL. 336 (1991) (arguing that many international issues are better modeled as coordination games). Fearon's two stage approach may be more appropriate to the treaty context than to the custom context. Fearon points out that relatively large "shadows of the future" might inhibit bargaining to achieve an initial coordination game agreement, while making the enforcement game more tractable. In the CIL context, there is less natural separation, and there may even be first mover advantages that would counteract the effect Fearon suggests. Finally, our interest in this article is not so much in establishing CIL rules, as in enforcing them.

³⁴ For a discussion of the use of coordination games to model certain types of international contexts, see Barbara Koremenos, Charles Lipson, & Duncan Snidal, *The Rational Design of International Institutions*, 55:4 INT'L ORG. 761, 774 (2001); Duncan Snidal, *Coordination versus Prisoner's Dilemma: Implications for International Cooperation and Regimes*, 79:4 AM. POL. SCI. REV. 923 (1985).

take account of a number of the diverse considerations often referred to together as “reputation” or “reputational sanctions.”³⁵ We also believe that the multilateral prisoner’s dilemma must be at the core of a rationalist explanation of the effectiveness of social norms. Finally, we believe that the multilateral prisoner’s dilemma offers parsimony: in the CIL context, the factors that it takes into account seem necessary, and there are no factors that seem superfluous.

The prisoner’s dilemma is a non-cooperative game, meaning that the players are unable to enter into binding agreements with one another. Although treaties are binding in a formal legal sense, as the source of their binding force is CIL itself, it seems appropriate at the outset to model CIL as a non-cooperative game. However, it may be appropriate to relax the assumption that the general international legal context is a non-cooperative game once we determine that the CIL of *pacta sunt servanda* has binding force, lending binding force to treaties. While this is critical to *treaty* law, it should be emphasized that it does not affect our analysis of CIL.

Thus, in a non-cooperative, single-play circumstance, with a standard prisoner’s dilemma payoff structure, we would expect non-compliance.³⁶ This is each player’s dominant strategy, meaning that each player’s payoff from defection is superior to its payoff from cooperation, no matter what the other player does. This dominant strategy is illustrated in the table below, in which no matter what State B does, State A obtains a better payoff by defecting.³⁷

Prisoner’s Dilemma Game		State B	
		Cooperate	Defect
A State	Cooperate	3;3 [sum=6]	1;4
	Defect	4;1	2;2 [sum=4]

³⁵ See Guzman, *supra* note 1.

³⁶ STEPHEN MARTIN, *ADVANCED INDUSTRIAL ECONOMICS* 98 (1993).

³⁷ Bilateral defection is a Nash equilibrium because under the payoffs assumed in the prisoner’s dilemma, each party is better off defecting, no matter what the other party does. A “Nash equilibrium” is a set of “strategies such that each player’s strategy is an optimal response to the other players’ strategies.” DREW FUDENBERG & JEAN TIROLE, *GAME THEORY* 11 (1991).

Therefore, under the rather restrictive assumptions of the true prisoner's dilemma, the parties each invariably choose the strategy that results in reduced individual welfare, and reduced aggregate welfare, compared to the non-defecting strategy. This is an inefficient outcome. By analogy, states playing the CIL game (assuming prisoner's dilemma-type payoffs) in a bilateral single-play setting would fail to form or comply with a CIL rule that increased individual and aggregate welfare. Cooperation is strongly dominated, and the resulting equilibrium is for both states to defect.³⁸ The same is true of a prisoner's dilemma game repeated a finite number of times known in advance to the players. Again, the unique resulting equilibrium is for each player to defect in each period.³⁹

This conclusion is inescapable in theory, given the constraints of the game: by definition, the outcome of the prisoner's dilemma game is an inefficient strategic equilibrium. This conclusion presents a normative goal: to modify the real world circumstances so as to produce stable equilibria that are efficient. This is the major role of CIL, and of international law generally.

Of course, in a world of effective third-party enforcement of agreements, the response to the prisoner's dilemma is clear: the parties would enter into a binding agreement to cooperate, thereby modifying the payoff structure and escaping the prisoner's dilemma. The prisoner's dilemma assumes, however, that its prisoners are held separately, and cannot negotiate, reach, or enforce a binding agreement.

In the CIL setting, there is no court of general mandatory jurisdiction nor any publicly appointed "policeman." While we may draw analogies to the World Court and the U.N., these institutions have substantial differences compared to domestic courts and police. Therefore, we begin by assuming that there is no capacity to make agreements that are binding. This is obviously a simplifying assumption. In fact, our model is meant to show that there are substitutes for formal binding agreements, and that these informal substitutes may have binding force. Once the capacity to bind is established, the players are no longer in a prisoner's dilemma.

In the CIL game, there are four additional important distinctions from the assumptions of non-cooperative game theory in general, and the prisoner's dilemma in particular. First, the players can communicate with one another, and can do so more readily today than during the classical period of formation of CIL. Second, states play a repeated game with one another with no defined end date, and so can respond at a later time to something done at an earlier time. Updating of information and punishment are possible. Third, not only is the narrow game characterized by a particular CIL rule, like the three-mile territorial sea, but it is embedded in a dense fabric of relationships. Fourth, information regarding compliance is often readily accessible; more so today, it would appear, than during the 19th century. Each of these distinctions alone may be sufficient to

³⁸ This is a "Nash equilibrium." *Id.* at 111.

³⁹ This is a subgame perfect equilibrium. "A subgame perfect equilibrium is a strategy profile that induces a Nash equilibrium in every subgame." M.J. OSBORNE, AN INTRODUCTION TO GAME THEORY (Oxford University Press, 2004).

transform the game into something quite different from the prisoner's dilemma—while nothing resolves the true prisoner's dilemma, modifications may result in stable and efficient equilibria. Indeed, it may be useful to use cooperative game theory to analyze some CIL circumstances.⁴⁰

As will be illustrated below, one of the more difficult types of multilateral cooperation problem is a commons problem in which, as in the cartel context, incentives to defect increase with the number of other states that comply. The more that the common is protected, the greater the benefits of defecting. Where incentives to defect increase with the number of players, we would expect the most severe challenge to cooperation. Not all CIL contexts exhibit this characteristic.

Observation suggests that even certain commons problems, or other prisoner's dilemma-type circumstances, may achieve resolution despite theory. Elinor Ostrom states that “[a] substantial gap exists between the theoretical prediction that self-interested individuals will have extreme difficulty in coordinating collective action and the reality that such cooperative behavior is widespread, although far from inevitable.”⁴¹ Ostrom cites considerable evidence regarding the amount and circumstances of cooperation by individuals in multilateral collective action problems. The evidence shows that individuals contribute to the resolution of these problems in substantially greater amounts than the standard prisoner's dilemma model would suggest. Of course, much of the evidence is obtained in circumstances where the assumptions of the prisoner's dilemma are violated—by allowing individuals to communicate, enter into agreements and repeat play.

We do not deploy any statistically significant data in this article. However, we advance a plausible theory of potential efficient equilibria in the multilateral prisoner's dilemma, and suggest some possible approaches to empirical testing.

b. Choice of Strategy

The next step in constructing our model is to postulate a plausible strategy that states might play within the prisoner's dilemma. There are many choices. We mention three strategies of retaliation derived from the literature of game theory: tit-for-tat, grim trigger, and penance. These strategies are stylized assumptions about state behavior, and they are necessary in order to complete the model. However, the reader will see that these strategies seem like plausible state responses to defection by others in particular circumstances. There are two ways in which retaliation might occur: the harmed state alone retaliates (bilateral retaliation) or the broader community of states retaliate (multilateral retaliation). In section 4 below, we discuss the possibility of bilateral or multilateral retaliation, and how this choice would affect our model.

⁴⁰ For an application of cooperative game theory in the international context, see Daniel G. Arce M. & Todd Sandler, *A Cooperative Game Theory of Non-Contiguous Allies*, 3 J. PUB. ECON. THEORY 39 (2000) (applying cooperative game theory to international security arrangements).

⁴¹ Elinor Ostrom, *Collective Action and the Evolution of Social Norms*, 14:3 J. ECON. PERSP. 137, 138 (2000).

i. Tit-for-Tat

Under “tit-for-tat,” states may respond to defection with a single defection. Tit-for-tat is one of the most frequently-discussed strategies in connection with repeated prisoner’s dilemmas. While tit-for-tat may win evolutionary games,⁴² it is not credible: after a defection, the wronged state will have incentives to accept an undertaking from the defecting state that it will cooperate in future.⁴³ Even more devastating to tit-for-tat is the fact that once one state defects the game cycles endlessly between defection and compliance.⁴⁴ Given the implausibility of tit-for-tat, we reject it and do not evaluate the implications of its use.

ii. Grim Trigger

Second, states may respond to a single defection with defection forever: a “grim trigger” strategy. There are two basic approaches that have been developed in the theory of repeated games. The first assumes that any deviation is met with a response that maximizes the loss that the deviator suffers – a “minmax” strategy – even if this imposes costs on the punishers. The second approach assumes that deviation results in reversion to the Nash equilibrium of the prisoner’s dilemma game: defection. We adopt the latter approach since it appears to be more appealing to players. Essentially, we assume that in the event of deviation the states revert to the strategies that they would have adopted if no CIL rule had developed in the first place. The grim trigger strategy is credible, as it calls for a reversion to the dominant strategy of defection in response to an initial defection.

However, the grim trigger strategy is unappealing for the same reason that failure to resolve the prisoner’s dilemma of cooperation is unappealing: it results in inefficiency.⁴⁵ Thus, in his work on treaties relating to environmental commons problems,⁴⁶ Barrett rejects the grim trigger strategy because it fails to satisfy the criterion of collective rationality. The collective rationality consideration is a formal articulation of the intuitive concern that it would be extraordinarily wasteful to abandon an efficient multilateral agreement because of a single defection. While it would be *individually* rational to respond with defection forever—it simply calls for reversion to the Nash equilibrium—it is collectively irrational insofar as rational negotiators will have incentives to renegotiate a cooperative arrangement after defection. Therefore, it is not credible.

⁴² See ROBERT AXELROD, *THE EVOLUTION OF COOPERATION* (1984).

⁴³ In the language of game theory, it is not “subgame perfect.” Subgame perfection means that at every stage of a repeated game, no player will have an incentive to deviate from the equilibrium strategy, even when others do. See note 39, *supra*.

⁴⁴ *Id.*, at 138.

⁴⁵ Goldsmith and Posner appear to recognize the possibility for stable and efficient equilibria under certain circumstances where states play the grim trigger strategy. However, using an overfishing of commons context as their example, they suggest that the grim trigger is not used and would be collectively irrational. Goldsmith & Posner 1999, *supra* note 1, at 1129-1130.

⁴⁶ Scott Barrett, *A Theory of Full International Cooperation*, 11 J. THEORETICAL POL. 519 (1999); SCOTT BARRETT, *ENVIRONMENT AND STATECRAFT: STRATEGIES OF ENVIRONMENTAL TREATY-MAKING* (2003).

With respect to the more empirical question of whether a grim trigger strategy is actually used, if we think not about the CIL that exists, but about the CIL that does not exist, it is clear that states do play the grim trigger strategy at least in some contexts. In fact, one might argue that the grim trigger is the existing default strategy in CIL. That is, where a CIL rule exists or is proposed for formation, and one state deviates, that may be sufficient provocation to cause others to deviate forever in response: to kill the rule multilaterally.

The grim trigger strategy is not credible because it is not “renegotiation-proof.”⁴⁷ That is, after a defection, non-defecting states will have incentives to come together and cooperate with the defector, depriving the grim trigger of credibility and therefore effectiveness. The defector can make an appealing “let bygones be bygones” argument.⁴⁸ An obvious counterargument to this, of course, is that renegotiation unravels if states perceive that this argument may be made repeatedly. That is, once the players understand that defection and promises of future compliance will go on indefinitely, would they not decline to renegotiate the first time? Moreover, in the CIL context, we are operating on the assumption that states do not have the possibility of forming binding agreements through renegotiation, preventing this indefinite defection.⁴⁹

Nevertheless, there may be circumstances in which such renegotiation is possible, and the offer to renegotiate *and abide by the results* is credible. If so, states will have incentives to defect and an alternative, “renegotiation-proof,” strategy is needed. We offer the penance strategy described below as a “weakly renegotiation-proof” alternative.

iii. Penance

If renegotiation is possible and credible, the states will prefer to renegotiate after a defection, making defection an attractive option.⁵⁰ The strategy known as “penance”⁵¹, is both individually credible in that states would individually find it attractive to play it, and collectively credible in the sense that it is likely to be more attractive than

⁴⁷ See, e.g., Joseph Farrell & Eric Maskin, *Renegotiation in Repeated Games*, 1 GAMES & ECON. BEHAVIOR (1989); Fudenberg & Tirole, *supra* note 38 at 174.

⁴⁸ Indeed, this is not uncommon in international law discourse. See, e.g., Scott M. Sullivan, *Changing The Premise Of International Legal Remedies: The Unfounded Adoption Of Assurances And Guarantees Of Non-Repetition*, 7 U.C.L.A. J. INT’L L. & FOR. AFF. 265 (2002-2003).

⁴⁹ Barrett notes that collective rationality is less of an issue in the field of domestic antitrust law, because renegotiation of an agreement in restraint of trade is illegal. Barrett, *supra* note 46 at 11.

⁵⁰ While a precise definition of renegotiation-proofness has not yet been agreed in the game theory literature, the treatment by Farrell and Maskin is worth considering. They define a “weakly renegotiation-proof” (WRP) equilibrium for an infinitely repeated game to be a subgame perfect equilibrium strategy profile that is not Pareto-dominated by any other subgame perfect strategy profile. Using this definition, the grim trigger strategy profile described above is not WRP, since after defection the payoffs to cooperation Pareto-dominate those of punishment. Farrell & Maskin, *supra* note 47.

⁵¹ See Fudenberg & Tirole, *supra* note 38, at 179-182. “Penance” is sometimes referred to as “getting even.” See ROGER B. MYERSON, GAME THEORY: ANALYSIS OF CONFLICT 326-27 (1991).

renegotiation.⁵² Fudenberg and Tirole show that the following “penance” strategy profile is “weakly renegotiation proof”: “Begin in the cooperative phase where both states play Cooperate. If a single state *A* switches to Defect, switch to the punishment phase for *A*. In this phase, state *A* plays Cooperate and the other state plays Defect. Play remains in this phase until the first time state *A* plays Cooperate, at which point play returns to the cooperative phase.”⁵³

The logic of this strategy in the CIL context is that a state having defected from a rule can have the rule reinstated only by accepting a period of punishment, in which it cooperates while the other state defects against it. The period of punishment is equal to the period of defection, ensuring that there is no net gain from defection.

A form of penance seems to be endorsed by the International Law Commission of the United Nations as the CIL rule for application in international law generally. Articles 49 to 54 of the Articles on State Responsibility provide that countermeasures may be used only to induce a state to cease a wrongful act *and to make reparations*;⁵⁴ they must be commensurate with the injury.⁵⁵

iv. Equilibrium Selection, Coordination and the Role of CIL

One of the problems in a multilateral prisoner’s dilemma is identifying the strategy that other players are playing and coordinating on a single strategy. Fudenberg and Tirole conclude: “Thus, repeated play with patient players not only makes ‘cooperation’—meaning efficient payoffs—possible, it also leads to a large set of other equilibrium outcomes. Several methods have been proposed to reduce this multiplicity of equilibria; however, none of them has yet been widely accepted, and the problem remains a topic of research.”⁵⁶ Under circumstances of multiple equilibria, “anything that tends to focus the players’ attention on one particular equilibrium, in a way that is commonly recognized, tends to make this the equilibrium that the players will expect and thus actually implement.”⁵⁷ While there is no formal solution to this problem in the game theory literature, states may coordinate through diplomacy, through other communication, or through their actions advancing particular customary rules. The selection among multiple equilibria may be understood as a separate, coordination game. Here, CIL (as reflected for example in the Rules of State Responsibility) and institutions like the International Law Commission may also play a role.

3. Assumptions Within the Multilateral Prisoner’s Dilemma

⁵² That is, it is both subgame perfect and is “weakly renegotiation-proof.” Fudenberg & Tirole, *supra* note 38, at 180, *citing* Farrell & Maskin, *supra* note 47, at 327; Eric van Damme, *Renegotiation-Proof Equilibria in Repeated Prisoner’s Dilemma*, 47 J. ECON. THEORY 206 (1989).

⁵³ Fudenberg & Tirole, *supra* note 38, at 180. *See also* the “defect for deviate” strategy proposed by Mahoney and Sanchirico, *supra* note 26.

⁵⁴ *See* The Chorzow Factory Case, 1928 P.C.I.J., Ser. A, Nos. 7, 9, 17, 19 (obligation to make reparations).

⁵⁵ *See* David J. Bederman, *Counterintuiting Countermeasures*, 96 AM. J. INT’L L. 817 (2002).

⁵⁶ Fudenberg & Tirole, *supra* note 37, at 112.

⁵⁷ Myerson, *supra* note 51, at 371.

Having selected the multilateral repeated prisoner's dilemma game, and the penance or grim trigger strategy, in this section, we develop the more specific parameters and assumptions of our model. We introduce the concepts behind these parameters, explain their salience, and justify the assumptions made in our model.

Recall that the basic model is a repeated prisoner's dilemma, in which repetition of the game in the future, with future payoffs from cooperation, increases the incentives to comply. Under these circumstances, much depends on the frequency of future interaction, the patience of states to receive these future payoffs, and the ability of states to identify defection by others.

a. Payoffs, Efficiency and Symmetry

By postulating the payoff structure of the prisoner's dilemma, we implicitly assume that failure to reach a cooperative equilibrium—failure to reach an implicit agreement—is inefficient. That is, the payoffs from cooperation are greater than the payoffs from mutual defection. Of course, there are many circumstances in which no implicit agreement is needed, and reaching one would be inefficient. However, our goal is to examine strategic barriers to implicit agreement; reducing these barriers would generally increase efficiency, just as reducing the general barriers to contract between private parties would generally increase efficiency without requiring that parties contract in every circumstance. This perspective is consistent with the first theorem of welfare economics, the Coase Theorem, and the “efficiency principle”: “If people are able to bargain together effectively, and can effectively implement and enforce their decisions, then the outcomes of economic activity will tend to be efficient (at least for the parties to the bargain).”⁵⁸ If the barriers to bargaining are eliminated, and parties reach no bargain, we may assume that there was no Pareto improving bargain available.

As stated above, we assume payoffs along the lines of the classic prisoner's dilemma. Even within this category, however, there is variation. Some circumstances will be more like a commons problem, or a cartel, in which the greater the number of players that comply, the greater the incentives to defect. Others will be the opposite, based on network effects, public goods or economies of scale: the more players that comply, the greater the incentives to comply.⁵⁹ In other cases, the payoffs from defection may not be substantially greater than the payoffs to compliance.

Different players may be affected differently by defection or compliance. Certainly in the CIL field there are notable cases of asymmetry. For example, a state with an extensive diplomatic service will have more at stake in connection with a rule of diplomatic immunity. A land-locked state may have a different perspective on the territorial sea than a state with extensive coastlines. Asymmetry affects each state's

⁵⁸ PAUL MILGROM & JOHN ROBERTS, *ECONOMICS, ORGANIZATION AND MANAGEMENT* 24 (1992).

⁵⁹ See note 9, *supra*.

incentives to comply. Linkage among varying issues may either increase or decrease asymmetry.⁶⁰

b. Identity and Number of Players

We disagree with the skepticism of prior work regarding the possibility for cooperative multilateral outcomes even in contexts that do not provide the possibility for formally binding agreements. The industrial organization literature recognizes important possibilities for cooperative outcomes in even large-numbers multilateral settings. The same seems to apply to the CIL game.

The number of players in any particular instance of the CIL game will vary. The maximum number of players is the total number of states in the world, although even this may ignore some detail. There are other players besides states, including sub-state governmental entities, international government entities, non-government organizations and individuals, so we have a potentially unlimited universe. In this connection, as a simplifying assumption, we assume “billiard ball” states. While we recognize that we lose nuance by doing this, we are trying to work with a simple model that assumes that states have unitary preferences. As we discuss the implications of our model, and ways to increase the incentives for compliance, we suggest that the structure of domestic government is worth considering in order to determine the degree of patience of states and the degree to which state decision-makers are concerned with the possibility of retaliation in areas outside their functional responsibility.

Even limiting our universe to states, with approximately 200 in the world, there would seem to be a significant problem of obtaining information about positions and practices, as well as coordination. However, while states may possess formal sovereign equality, they are not substantively equal, and their participation in the CIL formation process is not homogeneous.⁶¹

Oscar Schachter writes that “As a historical fact, the great body of CIL was made by remarkably few States. Only the States with navies—perhaps 3 or 4—made most of the law of the sea. Military power, exercised on land and sea, shaped the customary law of war and, to a large degree, the customary rules on territorial rights and principles of State responsibility.”⁶² While of course historical circumstances have changed, this

⁶⁰ While our model deals with games in which information is common knowledge, Parisi develops the Harsanyi concept of stochastic symmetry and role reversibility: the longer the shadow of the future, the less any one state can be certain of the way in which it will be affected by a particular rule. See Parisi, *supra* note 1. See also, Robert O. Keohane, *The Demand for International Regimes*, 36:2 INT’L ORG. 325 (1982).

⁶¹ See Lisa L. Martin, *The Rational State Choice of Multilateralism*, in MULTILATERALISM MATTERS 91, 99 (John G. Ruggie, ed. 1993) (suggesting techniques, including delegation as in the Security Council, of reducing the effective number of players).

⁶² Oscar Schachter, *New Custom: Power, Opinio Juris, and Contrary Practice*, in Jerzy Makarczyk, ed., *THEORY OF INTERNATIONAL LAW AT THE THRESHOLD OF THE 21ST CENTURY: ESSAYS IN HONOUR OF KRZYSZTOF SKUBISZEWSK*, 531, 536 (Kluwer 1996). See also Mendelson, *supra* note 11, at 194, 215, 225 (in the past, “civilized” states were considered sufficient, and the applicable group of states need not be geographically constrained).

suggests a game in which the number of players varies, depending on the degree of implication of their interests, and depending on their ability to affect outcomes. In this game, players are heterogeneous across a number of parameters, including interest and power, and as will be seen below, degree of patience. In this sense, we may think of powerful states engaging in the CIL formation and maintenance game as exerting power through the articulation, formation and maintenance of CIL rules. Schachter wrote of general CIL, but it is also possible to have regional or other plurilateral CIL.⁶³

Some of the leading authors in this field are skeptical of the possibility for multilateral customary processes to result in stable and efficient strategic equilibria under circumstances other than pure self-interest or coercion. For example, Goldsmith and Posner see little possibility for efficient resolution of prisoner's dilemma games in multilateral settings.⁶⁴ While they see the possibility for efficient equilibria in certain bilateral settings, they assume that "the bilateral prisoner's dilemma cannot in any event be generalized to the situation of multilateral cooperation, which is such an important part of the traditional account."⁶⁵ In this connection, they follow an established tradition, led by Mancur Olson in 1965:

[U]nless the number of individuals in a group is quite small, or unless there is coercion or some other special device to make individuals act in their common interest, rational, self-interested individuals will not act to achieve their common or group interests.⁶⁶

Olson based his perspective on the assumptions that the benefit of cooperation declines with the number of players, that the costs of monitoring increase with the number of players, and that the costs of organizing retaliation increase with the number of players.⁶⁷ However, it can readily be seen that these are conjectures about the world,⁶⁸ and are not necessarily true of any particular circumstance. Moreover, these are only a subset of the parameters worth considering. Finally, technological and social change has made it easier in some circumstance to monitor and to organize retaliation.⁶⁹ Our model provides a broader context in which to consider these, and other, parameters.

Oye identifies three slightly different ways in which increasing the number of players reduces the likelihood of cooperation: (i) increasing transaction costs, (ii) the

⁶³ See *Asylum Case (Haya de la Torre)*, 1950 I.C.J. 266; *Case Concerning the Right of Passage Over Indian Territory*, 1960 I.C.J. 6. Cf. MALCOLM N. SHAW, *INTERNATIONAL LAW* 72-73 (4th ed. 1997); Edward T. Swaine, *The Local Law of Global Antitrust*, 43 WM. & MARY L. REV. 627, 706-25 (2001).

⁶⁴ Goldsmith & Posner 1999, *supra* note 1, at 1132.

⁶⁵ *Id.*

⁶⁶ MANCUR OLSON, *THE LOGIC OF COLLECTIVE ACTION* 2 (1965). See also Kenneth A. Oye, *Explaining Cooperation Under Anarchy: Hypotheses and Strategies*, in *COOPERATION UNDER ANARCHY* (Kenneth A. Oye, ed. 1986).

⁶⁷ Olson, *supra* note 66, at 48.

⁶⁸ RUSSELL HARDIN, *COLLECTIVE ACTION* 43 (1982).

⁶⁹ See Arthur Lupia & Gisela Sin, *Which Public Goods are Endangered?: How Evolving Communication Technologies Affect the Logic of Collective Action*, 117 PUBLIC CHOICE 315 (2003); Ronald B. Mitchell, *Sources of Transparency: Information Systems in International Regimes*, 42 INT'L STUDS. Q. 109 (1998)..

lack of credibility of retaliation by third parties, suggesting that players would not retaliate against a defector, and (iii) increased heterogeneity of relative patience among larger numbers of players.⁷⁰ The transaction cost problem (compared to transaction benefits) with increasing numbers of players is merely a conjecture and could readily be countervailed by the possibility of economies of scale and scope. As discussed above, the credibility issue may vary depending on the strategy assumed. Furthermore, as discussed below, it is entirely possible that merely bilateral retaliation could support an efficient equilibrium. With respect to heterogeneous degrees of patience, as set forth in more detail below, we would agree with Oye that the ability to achieve cooperation would depend on the degree of patience of the least patient state. The implication of this constraint is that patient states may find it useful to exclude impatient states from certain cooperative arrangements. They may also be interested in increasing the patience of impatient states.

c. Information and Bilateral versus Multilateral Retaliation

The relative scale of information in the international system is somewhat different from that in a municipal setting. That is, the cost of producing and distributing information regarding state behavior may be a much smaller fraction of the utility of cooperation in the international setting than it may be in municipal inter-firm contexts. Furthermore, there are significant asymmetries among states in terms of the relative cost and value of producing information. Epistemic communities among government officials may play an important role in information transmission.

There are significant differences between a bilateral game and a multilateral game, and between a multilateral game with bilateral retaliation and one with multilateral retaliation. Under bilateral retaliation, information problems are significantly reduced, albeit not eliminated. Our model assumes bilateral retaliation, but multilateral retaliation is possible under some information conditions and would sustain cooperation in a wider range of circumstances.

An important aspect of the structure of the game pertains to the ability to retaliate in a discriminatory manner. States may have trouble discriminating in the application of sanctions for several reasons. First, they may not be able to obtain information regarding the author of the violation. This might occur, for example, with respect to pollution at sea, or terrorist attack. Second, it may be costly for states to respond in a discriminating way. For example, if the sanction involves trade barriers, the sanctioning state must instruct its customs officers to discriminate among goods by origin. Third, and most important, the relevant good being produced by cooperation may be non-excludable. This would occur with the provision of public goods or the protection of international commons. To the extent that states are unable to discriminate, their retaliation, if any, must be multilateral, instead of bilateral. This obviously limits the strategies that they are able to play and the relationships that they may enter into. Thus, given that the strategies available to a state are “cooperate” or “defect,” there are at least two possibilities that we need to consider in connection with a multilateral game:

⁷⁰ Oye, *supra* note 66, at 18-19.

- (i) **Bilateral Retaliation:** Defection by state *a* against state *b* leads to punishment of *a* only by state *b*: bilateral defection leads to bilateral punishment; or
- (ii) **Multilateral Retaliation:** Defection by state *a* against state *b* leads to punishment of *a* by all states: bilateral defection leads to multilateral punishment.

We focus our analysis on bilateral retaliation for three reasons. First, it makes the analysis simpler without changing any of the qualitative conclusions. Second, multilateral retaliation simply increases the incentives to comply over bilateral retaliation. Since bilateral retaliation involves milder punishment of defection than multilateral retaliation, the conditions that support cooperation with bilateral retaliation under bilateral retaliation will certainly support cooperation with multilateral retaliation under multilateral retaliation. In other words, where multilateral retaliation is possible, the conditions that we identify below are sufficient but not necessary to support a multilateral rule: any degree of patience that supports a multilateral CIL rule with bilateral retaliation will also support such a rule with multilateral retaliation. Third, there is some force to the argument that bilateral retaliation is a more plausible scenario than multilateral retaliation in most of the situations in which the formulation of CIL is likely to be considered. Articles 42, 48 and 54 of the International Law Commission's Rules on State Responsibility⁷¹ generally prevent retaliation by third states for truly bilateral injury, while there seems no basis for retaliation by an injured state against non-injuring states. These rules seem to limit the formal possibility for multilateral retaliation against truly bilateral defection, at least *within* a particular CIL rule. On the other hand, where a CIL rule is not formed, or falls into desuetude, we might understand that there is multilateral retaliation, as well as retaliation against non-injuring states.

In any event, in our model, we assume that retaliation is applied bilaterally—that if state *a* defects vis-à-vis state *b*, only state *b* will respond, and only against state *a*. If multiple states responded against state *a*, it would simply make cooperation more likely by increasing the punishment for defection.

Thus, assuming bilateral retaliation, we can represent a multilateral prisoner's dilemma game as a set of bilateral games. As will be seen below, this is not the same as assuming a bilateral game. Rather, it is a multilateral game with bilateral retaliation.

Although we assume bilateral retaliation, it is worthwhile to discuss the possibility and implications of multilateral retaliation. Under multilateral retaliation, we would be concerned about the ability of players to find out about the compliance or defection history or characteristics of other players. On the one hand, it may be costly for

⁷¹ Articles on Responsibility of States for Internationally Wrongful Acts, in Report of the International Law Commission on the Work of Its Fifty-third Session, UN GAOR, 56th Sess., Supp. No. 10, at 43, UN Doc. A/56/10 (2001), available at <www.un.org/law/ilc>, subsequently noted by the General Assembly in A/RES/56/83, Responsibility of States for Internationally Wrongful Acts, adopted 12 Dec. 2001. These articles do not themselves represent international law, but are an attempt to codify existing custom.

an individual to find out for himself the history of many other players. On the other hand, the potential responses of many other players, if engaged, may add to the disincentives for defection. There are economies of scale and scope in this type of system, which may countervail increased information costs that exist in an n-player setting.⁷²

Along these lines, Kandori explains that informal enforcement mechanisms fall into two categories: those that use personal enforcement and those that use community enforcement. These correspond to what we have been referring to as bilateral retaliation and multilateral retaliation. Kandori examines circumstances in which social norms work to support efficient outcomes in infrequent transactions—absent repetition that can allow personal enforcement, but under circumstances where community enforcement may occur.⁷³ Thus, but there may be circumstances in which multilateral retaliation could substitute for repetition. Alternatively, as noted above, multilateral retaliation could result in cooperation in circumstances in which bilateral retaliation would be insufficient. Where members of a community can observe each other's behavior, community enforcement works in much the same way as personal enforcement.⁷⁴ Kandori assumes this is the case in small communities. We might suggest that the global community can be understood as a small community in this sense. CIL rules often address matters that are public knowledge and are reported in the press. Imagine a municipal community where each individual's behavior is subject to journalistic and intelligence investigation.

Kandori assumes private information that is not shared among community members. When observability is not perfect, private information regarding compliance with a norm, as well as the distribution of the private information, will be more complicated, and costly. Cooperation may be difficult to sustain because the community may not have defined adequately the social norm or the determination and labeling of defectors. Kandori shows, in theoretical terms, that even where an agent does not have any direct information of other agents' behavior vis-à-vis other members of the community, cooperative behavior can sometimes be sustained. In the CIL game, with seemingly greater relative ability to observe treatment of third parties compared to private society, we would expect a greater basis for cooperation.

Of course, multilateral sanctions are dependent on information regarding defection, and a judgment that the subject has violated the relevant norm. Information may be a trivial problem in certain areas of CIL, but a difficult problem in others. Various institutional responses are possible to provide greater certainty in judging violations. Judgments can depend on individual state determinations, or on community views, or consensus, which may be led by a "reputation entrepreneur" or other opinion leader. On the other hand, it can be achieved through an independent institution such as a court. And it must be recalled that in this type of setting, even a court's views can be criticized and challenged.

⁷² See Samuel Bowles & Herbert Gintis, *The Moral Economy of Communities: Structured Populations and the Evolution of Pro-Social Norms*, 19 *EVOLUTION & HUMAN BEHAVIOR* 3, 11-14 (1997).

⁷³ Michihiro Kandori, *Social Norms and Community Enforcement*, 59 *REV. ECON. STUDS.* 63 (1992).

⁷⁴ *Id.*

As Milgrom, North and Weingast argue with respect to the non-state institutions that enforced compliance among early medieval merchants, “It is the costliness of generating and communicating information—rather than the infrequency of trade in any particular bilateral relationship—that, we argue, is the problem that the system of private enforcement was designed to overcome.”⁷⁵

Unlike in the municipal setting, the international setting provides no system of courts with broad mandatory jurisdiction. Milgrom, North and Weingast argue that third-party dispute settlement can assist in developing cooperation. Third-party dispute settlement can solve the following information problem. If two parties have a dispute, in which one accuses the other of defection, how can other members of the community determine whether the accusation is true?⁷⁶ Third-party dispute settlement, along the lines of the “law merchant,” may be more valuable to resolve information problems in the municipal context among traders than in the international context among states. First, there may be fewer states than there are potential traders in the medieval traders’ setting. Second, again, the cost of information about state compliance may be a smaller proportion of the value of CIL “transactions.” Milgrom, North and Weingast conclude within the municipal context that, given the lack of empirical evidence about the costs of running different kinds of institutions, it is not possible to develop a formal model to show that their third-party dispute settlement minimizes information costs. However, they opine that the system seems to incur only the kind of costs that are inevitable and seems well-designed to minimize those costs.⁷⁷

The Milgrom, North and Weingast “law merchant” is a private purveyor of information and evaluation. The players accept its use in order to develop an efficient equilibrium. We might consider the extent to which formal international institutions such as the International Court of Justice, the WTO’s dispute settlement process or its Trade Policy Review Mechanism fill a similar role in connection with states, and whether NGOs playing an informal role, such as Amnesty International or the World Wildlife Fund, or informal institutions such as the Basle Committee (bank regulation) or the Waasenaar Arrangement (export controls on dual use commodities), can do so in particular niches.

We assume perfect information in our model. This assumption seems reasonable given our assumption of bilateral retaliation. However, in order for multilateral retaliation to operate, information problems must be overcome.

d. Patience/Discount Factors, Horizon and the Shadow of the Future

When international lawyers discuss incentives for compliance, they often refer to the possibility of role-reversal in the future: of reciprocity and retaliation. However,

⁷⁵ Paul R. Milgrom, Douglass C. North and Barry R. Weingast, *The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges and the Champagne Fairs*, 2 *ECON. & POL.* 1,3 (1990). Note that these authors are suggesting that information may compensate for infrequency of bilateral interaction.

⁷⁶ Milgrom, North & Weingast, *supra* note 75 [Error! Bookmark not defined.](#), at 8.

⁷⁷ *Id.* at 15.

from a rationalist standpoint, a critical question is how powerful is the shadow of the future? In repeated prisoner's dilemmas, theorists have shown that the degree to which players value future payoffs will have an effect on players' incentives to comply with a norm. All things being equal, the extent to which the player values future payoffs will determine the extent to which these future payoffs affect the player's behavior.⁷⁸

We represent the extent of valuation of future payoffs as a discount factor⁷⁹—a factor used to reflect the present value to a particular player of future payoffs. The discount factor reflects the preference of the player for payoffs now versus payoffs in the future. We may understand the discount factor as a central variable, which interacts with other variables such as the per-period magnitude of future payoffs, the relative payoffs of defection versus cooperation, the horizon or number of periods predicted, frequency of repetition, the number of other players under multilateral retaliation, and the degree of linkage to other relationships.

While it might be argued that the CIL game will continue infinitely, in order to emphasize the role of patience in our model, it is useful to assume that the game will be finite, but that at any given time it is unknown when it will end. Thus, we assume a low probability that the game will be short. At any given moment, there will be a long, but finite, horizon.

Public choice considerations would counsel that horizons vary. After all, if the real interaction is not between states but between governments, we must recognize that governments have varying effective horizons. Governments come and go. The relationship between state and government horizons is to some extent determined by the degree of accountability of the government—the degree to which it represents the interests of the state. However, a democracy may have a shorter horizon than a dictatorship. Some states may have more frequent or more imminent elections at particular moments. Separately from the frequency or imminence of elections, we would want to model the relative stability of the ruling party or coalition. Much depends on the prospects for reelection, the stability of the dictatorship, and the stability of the dictatorship's policies, including their susceptibility to variation due to corruption. Furthermore, it may be useful to examine whether the real actor is neither the state nor the government, but a more entrenched bureaucracy. Transnational networks may have

⁷⁸ Of course, to the extent that players may exchange future payoffs for present payoffs, the patience variable becomes less important, or less diverse. Thus, where efficient capital markets allow for the exchange of future cash or non-cash payoffs for cash payoffs, we would expect the patience variable to be constant across players. *See*

⁷⁹ A "discount factor" is a mathematical factor structured to reflect the degree of patience of a player. It represents the present value today of a payoff in a future period. A discount factor of 1 means that future payoffs are valued equally to present payoffs. A discount factor of .75 means that future payoffs are valued at 75% of present payoffs. A high discount factor indicates patience, while a low discount factor indicates impatience. Discount factors should be contrasted with "discount rates" applied to future payoffs in order to reduce them to a present value, where the discount factor equals $1/(1+r)$, with r representing the discount rate. Therefore, discount factors are less than 1 (assuming a positive discount rate). A low discount rate corresponds to a high discount factor. For example, a discount rate of 10% would result in a discount factor of approximately 91%. We discuss below the circumstances under which a state might be thought to have a high or low discount factor.

greater durability than international networks. We combine this question of an individual state's or government's time horizon with "patience."

Included in the set of assumptions underlying the prisoner's dilemma is the assumption of the isolation of the game under analysis. However, casual observation of international society suggests that there are many linkages,⁸⁰ with the result that few issues can be isolated.⁸¹ Players can bind one another in a variety of ways, including by linking the present game to other games in a "supergame."

It is generally understood that a mutually beneficial outcome can exist as a credible equilibrium of the prisoner's dilemma where the game is repeated, subject to conditions relating to the players' discount rates and the time horizon.⁸² The reason is simple to see. With repeated play current actions can be conditioned on past actions, introducing the possibility of rewarding cooperation and punishing defection. Repetition of itself is not, however, sufficient to secure continued cooperation. If the game is repeated a known finite number of times, both players will have an incentive to defect in the final period, and the game unravels from there to immediate defection.⁸³ By contrast, if the game is repeated indefinitely, then "all players defect every period" will remain a credible equilibrium, but there may be additional credible equilibria, depending on the parties' discount factors.⁸⁴

"The 'folk theorems' for repeated games assert that if the players are sufficiently patient, then any feasible, individually rational payoffs can be enforced by an equilibrium. Thus, in the limit of extreme patience, repeated play allows virtually any payoff to be an equilibrium outcome."⁸⁵ Under circumstances of high discount factors, when players are "patient," the short-term gain from defection in one play is outweighed by the aggregation of even small losses in all future periods. Fudenberg and Maskin show that frequent transactions with the same partner, regardless of the number of players, the number of strategies available, or the size of the payoffs, make it possible to reach an equilibrium with efficient trading.⁸⁶

e. Multi-Sector Contact

⁸⁰ See ROBERT KEOHANE, *AFTER HEGEMONY* 91 (1984); Ernst Haas, *Why Collaborate? Issue Linkage and International Regimes* 32 *WORLD POLITICS* 357 (1980); Michael D. McGinnis, *Issue Linkage and the Evolution of International Cooperation*, 30:1 *J. CONFL. RES.* 141 (1986); Robert E. Tollison & Andrew D. Willett, *An Economic Theory of Mutually Advantageous Issue Linkage in International Negotiations*, 33 *INT'L ORG.* 425 (1979). McGinnis shows formally that in a prisoner's dilemma "multisupergame" players may adopt strategies that create linkages across time and games, providing opportunities for cooperation where cooperation would not be possible for isolated games.

⁸¹ See GIANCARLO SPAGNOLO, *ISSUE LINKAGE, CREDIBLE DELEGATION, AND POLICY COOPERATION* (Ctr. for Econ. Policy Research, Discussion Paper No. 2778, May 2001).

⁸² See ROGER B. MYERSON, *GAME THEORY: ANALYSIS OF CONFLICT* 308-369 (1991).

⁸³ Fudenberg & Tirole, *supra* note 38, at 111.

⁸⁴ *Id.*

⁸⁵ *Id.* at 150.

⁸⁶ Drew Fudenberg & Eric Maskin, *The Folk Theorem in Repeated Games with Discounting or with Incomplete Information*, 54 *ECONOMETRICA* 533 (1986).

Firms, and states, operate in multiple markets and encounter other firms, or states, in multiple contexts: as competitor here, as supplier there, as co-conspirator elsewhere. Industrial organization economists studying the effect of multi-market contact have found that this cross-sectoral activity may support cooperation.⁸⁷ “With multimarket contact collusion can be viable in a set of markets even when in the absence of multimarket contact it could not be supported in *any* of these markets.”⁸⁸

“Multimarket contact allows firms to use the threat of a *simultaneous* punishment in more markets, which is stronger than the sum of the independent punishments because a firm being punished in one market has a higher marginal valuation of profits, therefore it values more the losses from punishments in other markets.”⁸⁹

One important difference between the commercial context and the international relations context is that state relations in the international context almost always cross a number of sectors.⁹⁰ States relate to one another in a variety of contexts, with varying roles in each context. Thus, in one context, State A may be concerned about the scope of its prescriptive jurisdiction, while in another context it may be concerned about the scope of its responsibilities to protect foreign diplomats. As a result, while there may be a “prescriptive jurisdiction game” that is separate from the “diplomatic immunity game,” these games may be linked. In fact, states regularly link issues in international relations,⁹¹ with the result that it is not possible to establish precise boundaries for any particular game.

Defection in one area may have consequences in another, with the possibility of cross-sectoral punishment. Thus, it is not enough to examine whether states have sufficient incentives for compliance within a particular sector or arrangement; one must also analyze the effect of activity in other sectors.

⁸⁷ See, e.g., D.B. Bernheim & M.D. Whinston, *Multimarket Contact and Collusive Behavior*, 21 RAND J. ECON. 115 (1990); Corwin Edwards, *Conglomerate Bigness as a Source of Power*, in “Business Concentration and Price Policy,” NBER Conference Report (Princeton University Press 1955); Hitoshi Matsushima, *Multimarket Contact, Imperfect Monitoring and Implicit Collusion*, 98 J. ECON. THEORY 158 (2001); Giancarlo Spagnolo, *On Interdependent Supergames: Multimarket Contact, Concavity and Collusion*, 89 J. ECON. THEORY 127 (1999).

⁸⁸ Spagnolo, *supra* note 87, at 128 (emphasis in original). Spagnolo shows that when players have a concave objective function in respect of profits, inducing wealth effects and therefore generating scale economies in cooperation, multi-market contact can enhance cooperation. A concave objective function in this context arises from an interest in smoother payoffs, and less interest in unusually high payoffs in any given period. There are reasons why a political “manager” might also have a concave objective function. See Edward D. Mansfield, Helen V. Milner, & B. Peter Rosendorff, *Why Democracies Cooperate More: Electoral Control and International Trade Agreements* 56:3 INT’L ORG. 477 (2002) (modeling trade agreements as signaling and commitment devices in the domestic context of democracies). In the international relations context, in which executives play a two-level game, a concave objective function may arise from accountability via elections, or in the trade context, from a desire to avoid disruption to particular constituencies. This is Corden’s “conservative social welfare function.” W.M. CORDEN, TRADE POLICY AND ECONOMIC WELFARE 107 (1974).

⁸⁹ Spagnolo, *supra* note 87, at 133.

⁹⁰ See Guzman, *supra* note 1, at 1869-1870; Snidal, *supra* note 34, at 939.

⁹¹ See note 80 *supra*.

Matsushima argues that multimarket contact can take the place of perfect information as a basis for a stable equilibrium of implicit collusion. Matsushima shows that with multimarket contact, cooperation can take place even under circumstances of relatively low discount factors.⁹²

This conclusion suggests that international cooperation in different sectors may be mutually supportive, and that there may be a kind of network effect that makes each additional instance of cooperation more attractive than it would be absent existing instances.⁹³ This game theoretic perspective provides support for the early neo-functionalism hypotheses regarding international economic integration.⁹⁴

However, the possibility that the “real” decision-maker is a subnational actor,⁹⁵ such as a bureaucracy with a limited sectoral mandate, may affect the possibility of effective multi-sector contact. This may be a reason why assignment of broad international relations authority to centralized ministries of foreign affairs may be useful. While as argued by Downs and Jones,⁹⁶ a particular state may have multiple reputations, within multiple contexts, segmentation need not be complete, so that it may be that defection in one context may have consequences within another context. Without further empirical study, it is not possible to know how much segmentation exists in states’ reputations. Such study would examine the extent to which different ministries within states, and different “epistemic communities” on a transnational basis, communicate across sectors. It is important to note, however, that particular segments, such as trade, environment or arms control, have within them multiple relationships. Finally, it may be that segmentation of reputation is efficient from the standpoint of domestic accountability, allowing different parts of a national government to take responsibility for their own relations, and to develop the kind of reputation that maximizes returns within the relevant segment.

Furthermore, Downs’ and Jones’ argument is explicitly about treaty, not custom. In a treaty setting, despite the broad scope of remedies available under the Rules of State Responsibility, states might be understood as implicitly accepting only intra-treaty remedies for breach. However, in the custom context, there is no explicit or implicit limitation on responsive or remedial action. Therefore, it may well be that in this more delicate and nuanced context, where there is no implicit consent to limitation, states would consider themselves less constrained in their responsive or remedial action. In addition, much responsive action in this informal setting is likely to take the form of

⁹² Matsushima, *supra* note 87, at 164-65.

⁹³ See Barbara Koremenos, Charles Lipson, and Duncan Snidal, *The Rational Design of International Institutions*, 55:4 INT’L ORG. 761, 764-65 (2001).

⁹⁴ See, e.g., ERNST B. HAAS, BEYOND THE NATION-STATE: FUNCTIONALISM AND INTERNATIONAL ORGANIZATION (1968); Jeppe Tranholm-Mikkelsen, *Neofunctionalism: Obstinate or Obsolete?* 20:1 MILLENNIUM: JOURNAL OF INTERNATIONAL STUDIES 1 (1991).

⁹⁵ See Anne-Marie Slaughter, *The Real New World Order*, 76 FOREIGN AFFAIRS 183 (1997).

⁹⁶ George W. Downs & Michael A. Jones, *Reputation, Compliance and International Law*, 31 J. LEG. STUDIES 95, 101 (2002) (suggesting that reputation varies by field of activity, and that this segmentation reduces the effects of reputation).

abstention from future transactions, rather than some form of punishment within the context of the present transaction.

Milgrom, North and Weingast, examining the behavior of medieval merchants, explain that “if the relationship itself is a valuable asset that a party could lose by dishonest behavior, then the relationship serves as a *bond*. . . .”⁹⁷ Thus the shadow of the future effect is intensified by multi-market contact and perfect information. The broader this effect, the greater the likelihood that individual states will respect individual rules.

4. Plausible Examples of the CIL Game

Before we go on to articulate the model and its implications, it is worthwhile to stop and suggest how certain actual CIL rules might fit into this framework.

Our purpose in this article is merely to elaborate a theory that shows the plausibility of efficient equilibria in a multilateral prisoner’s dilemma model of CIL. We did not set out to prove that such efficient equilibria exist in nature, or to prove any of the hypotheses that flow from our model. However, it is useful, by way of illustration rather than proof, to set out some examples of circumstances that can plausibly be characterized as international multilateral prisoner’s dilemmas that seem to have reached cooperative equilibria, as opposed to circumstances of intrinsic self-interest, narrow coercion, a bilateral prisoner’s dilemma, or a bilateral coordination game. We hasten to note that our characterization of these examples, like those advanced by others, is dependent upon subjective judgment as to the payoff structure that these circumstances may entail. However, without empirical proxies for payoffs, it is not possible to do better.

Much of international law relates to the allocation of authority, or responsibility for harm. These types of rules may be analogized to rules of property and tort.⁹⁸ So, for example, iconic CIL rules such as diplomatic immunity, sovereign immunity, territorial sovereignty, and the territorial sea, may be understood in terms of allocation of authority, which may be assimilated to property. Rules such as responsibility for harm to aliens, or responsibility for environmental harms to other states, may be assimilated to tort.

The strategic context for formation of property or tort rules depends on a number of parameters. These parameters include the costs and benefits of the resource or the potential harm, the degree of natural excludability and the cost of artificial exclusion, the degree of rivalry of consumption, the degree of bilateral monopoly, and the transaction costs of negotiation. We outline below four examples that seem to follow the prisoner’s dilemma structure.

a. Restrictive Theory of Sovereign Immunity

⁹⁷ Milgrom, North & Weingast, *supra* note 75, at 1.

⁹⁸ See Joel P. Trachtman, *Economic Analysis of Prescriptive Jurisdiction and Choice of Law*, 42 VA. J. INT’L L. 1 (2001).

Under the traditional “absolute” theory of sovereign immunity, which was applied by the U.S. until the 1952 “Tate Letter,” states may engage in “sovereign” acts or in “commercial” acts, and still enjoy jurisdictional immunity from foreign courts. With the rise of state trading and international commerce, this resulted in greater frequency of disappointed private persons who had dealt with foreign states only to find that they had no legal recourse.

This state of affairs may be characterized as a multilateral prisoner’s dilemma as follows. Each state individually would be better off, in terms of its potential responsibility to private persons, maintaining the absolute theory of sovereign immunity. However, if all states take this position, then commercial enterprises may be harmed in a greater amount, trade may be diminished, and global welfare may be reduced.⁹⁹ There may be a degree of asymmetry due to differences in utilization of state trading. The Soviet Union and its satellites favored absolute immunity.

During the 20th century, states like the U.S. and the United Kingdom, and most trading states, adopted the “restrictive” theory of sovereign immunity, which holds that states lack immunity for acts of a commercial nature. When states adopted the restrictive theory, they accepted the possibility of lawsuits against them based on the restrictive theory. If a state were to decline to adopt the restrictive theory, or renege on adoption of the restrictive theory, and claim absolute immunity, they would be required to grant immunity to foreign states, reverting to the less efficient status quo (grim trigger). States therefore had sufficient incentives to “cooperate” in enhancing global welfare by adhering to and continuing to adhere to the restrictive theory.

When the first state shifted from the absolute to the restrictive theory, it was possible that other states could have claimed a violation of the traditional rule of CIL: absolute immunity. Instead, they gradually accepted the restrictive theory. Our model suggests that they may have done so in light of the possible future gains from acceptance of the restrictive theory. This case is not easy to characterize as a series of bilateral games, but is better characterized as a multilateral game, with the possibility of bilateral retaliation while adhering to the restrictive theory, or multilateral retaliation by rejecting the restrictive theory completely. It is a multilateral game because the uniformity acceptance of the restrictive theory would create both public goods effects in terms of increased world trade and network externality effects in terms of reducing the costs of contracting by reducing the need to check which states have adopted the restrictive theory.

b. Cross-Border Environmental Harm

⁹⁹ While it is possible that contractual waivers of immunity might be used, they would not cover extra-contractual responsibility, and might entail costly negotiations. In any event, these examples are intended as sketches of possible analyses, rather than as definitive analyses.

The rule of *sic utere tuo*, adopted in the Trail Smelter Arbitration,¹⁰⁰ is derived from, and remains akin to, the early law of nuisance.¹⁰¹ In allocating responsibility to the source state with respect to transboundary pollution, it requires the internalization of externalities. We may understand this rule, like the coordinate domestic law of nuisance, to respond to a collective action problem. Each individual state has incentives to externalize pollution if it fails to take into account harm to downwind or downstream states. The Coase theorem, criticizing Pigou, teaches that every externality need not be internalized. However, there will be transaction cost contexts in which internalization is useful.

The case of transboundary pollution may be understood using the prisoner's dilemma, assuming that global welfare may be increased if each state internalizes externalities in making decisions about their pollution. Each state individually has incentives to violate this rule. The CIL rule of *sic utere tuo* arose in order to establish a rule of internalization. It is enforced by the threat of grim trigger, in the form of a "pollute thy neighbor" rule, or by the implicit threat of smaller retaliation along the lines of penance. The *sic utere tuo* rule applies where pollution crosses a single border, as well as where it crosses multiple borders. Furthermore, the transboundary pollution problem, like the nuisance problem, is one where there may be economies of scale in uniformity of arrangements. Therefore, this is not simply a bilateral game.

The emergence of the *sic utere tuo* rule in international law could plausibly follow a similar path to the emergence of the coordinate rule, or other property rights, in domestic law.¹⁰²

To the extent that a specific type of environmental protection is inexhaustible (whether or not it satisfies the other criterion to be a public good and is non-excludible), such as the ozone layer, it may be that the benefits of participation rise with the number of states that participate. Not all environmental goods will have the same payoff structure, and so the *sic utere tuo* rule would require extensive analysis.

¹⁰⁰ Trail Smelter Case (U.S. v. Can.), 3 R.I.A.A. 1905 (1938 & 1941). The principle of *sic utere tuo* was included as Principle 21 of the Stockholm Declaration, UN Conference on the Human Environment, Stockholm Declaration, June 16, 1972, UN Doc. A/CONF.48/14, princ. 21, 11 ILM 1416 (1972). See also Principle 2 of the 1992 Rio Declaration. See UN Conference on Environment and Development, Rio Declaration on Environment and Development, June 14, 1992, UN Doc. A/CONF.151/5/Rev.1, princ. 2, 31 ILM 874 (1992).

¹⁰¹ See also International Liability for Injurious Consequences Arising out of Acts Not Prohibited by International Law, Chapter VI of the Report of the 55th Session of the International Law Commission, General Assembly 58th Session, Supplement 10, A/58/10 (2003).

¹⁰² See Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. PAPERS AND PROCEEDINGS 347, at 350 (1967). See also B.C. Field, *The Evolution of Property Rights*, 42 KYKLOS 319 (1989); Thomas W. Merrill, *Trespass, Nuisance, and the Costs of Determining Property Rights*, 14 J. LEG. STUD. 13 (1985); R.S. Hartman, *A Note on Externalities and the Placement of Property Rights: An Alternative Formulation to the Standard Pigouvian Results*, 2 INT'L REV. L. & ECON. 111 (1982); John Umbeck, *Might Makes Rights: A Theory of the Formation and Initial Distribution of Property Rights*, 19 ECON. INQ. 38 (1981); David Ault & Gilbert Rutman, *The Development of Independent Rights to Property in Tribal Africa*, 22 J. LAW & ECON. 183 (1979); John Umbeck, *A Theory of Contract Choice and the California Gold Rush*, 20 J. LAW & ECON. 163 (1977); Anderson & Hill, *The Evolution of Property Rights: A Study of the American West*, 18 J. LAW & ECON. 163 (1975).

c. Territorial Sovereignty and Non-Acquisition of Territory by Use of Force

The general and specific rules associated with territorial sovereignty may also be modeled, like property rights, as a multilateral prisoner's dilemma. In particular, during the last century, the customary international law rule against formal acquisition of territory by force seems to have met with stronger compliance.¹⁰³ We may understand this principle as a disincentive for the use of force. If states are restrained by a principle of non-acquisition, they will have weaker incentives to use force. How can we understand the emergence of this principle?

Each state individually has incentives to retain flexibility to acquire territory through the use of force. However, each state also wishes, defensively, to remove the flexibility of other states to acquire its territory, or perhaps that of others, through the use of force. Each state has incentives to defect, but refrains from doing so due to fear of retaliation in other areas, or in the narrower area of acquisition of territory. Again, we can see that either grim trigger or penance may serve to maintain the equilibrium of non-acquisition by force.

This problem does not seem to have the characteristics of a commons or a cartel. That is, it does not appear that the incentives to violate increase with the number of states that comply. If many states had designs on the same territory, as in colonial times, this might be the case, and perhaps is the reason that a rule against acquisition by force only arose more recently. Today it appears that where a state covets other territory, it is only a neighbor's adjacent territory. This rule may have security benefits that rise with the number of states that adhere. That is, the greater the number of states that adhere, the lower the incentives preemptively to attack and therefore the lower each state's defense budget must be. Thus, it is plausible that this rule would result in increasing incentives to comply with rising numbers of adherents.

d. Territorial Jurisdiction

Customary international law includes the norm of territorial jurisdiction. Each state generally has jurisdiction over conduct within its territory. There is some dispute regarding the scope for "extraterritorial" jurisdiction, usually referring to conduct abroad that has adverse effects within the territory of the state seeking to assert jurisdiction. According to sections 402-403 of the American Law Institute's Restatement (Third) of Foreign Relations Law, and according to a number of states and scholars, there are substantial limits on the right of a state to assert extraterritorial jurisdiction. According to the Restatement, states may not exercise jurisdiction where it would be "unreasonable" to do so in light of the various connections and interests involved. For our purposes, the only important fact is that there is some arguable limit. We can understand this limit within the prisoner's dilemma model. It is noteworthy that, while the U.S. has on

¹⁰³ Mark Zacher, *The Territorial Integrity Norm: International Boundaries and the Use of Force*, 55:2 INT'L ORG. 215 (2001). This norm does not prevent states from exercising influence over policy within other states.

occasion asserted the right to apply its law “extraterritorially,” it has often been at significant diplomatic cost, and has often been subject to diplomatic adjustment.

We may assume that each state would prefer to exercise jurisdiction without limit, in order to affect circumstances that may impose negative externalities on its citizens. This is in some sense the reverse Coasian position of the *sic utere tuo* story. Each state is restrained by the fear of reciprocation or retaliation, resulting in a stable and efficient equilibrium.

It might be argued that the territorial jurisdiction problem raises commons issues. The more states that adhere to a rule of territoriality, the greater the benefits, in terms of regulatory geographic scope, to a violating state. On the other hand, compliance with a rule of territoriality would generally be aligned with rejection of foreign “extraterritoriality.” So, it may be that the benefits of violation do not rise with the number of states that comply. It is notable that there is little consensus on the existence or content of a rule of customary international law in this field.

5. The CIL Game

Our discussion in parts 2 and 3 indicates that there are many possible assumptions that we might make in developing a CIL game. As noted above, we choose to focus on one such game – the repeated prisoner’s dilemma game – for two principal reasons. First, this is a setting in which all of the relevant states prefer formation of and compliance with a CIL rule to the other possible outcomes. Second, it is a setting in which each state can gain from deviating from the cooperative outcome and in which there are no centralized means to enforce any agreement not to deviate.

The games we develop allow us to identify plausible circumstances under which the repeated prisoner’s dilemma can result in efficient equilibria both bilaterally and multilaterally. Moreover, they show what types of contexts, including malleable institutional features, may affect the ability of states to reach such equilibria.

In Appendix I, we set forth in formal mathematical terms the structure of our model. In this section, we describe the model in textual terms. The formal mathematical depiction of the model allows us to show the relationship among the various parameters of the model in a way that, if the data it calls for were inserted, would tell us something about the possibilities for cooperation.

Under the prisoner’s dilemma’s payoff assumptions, each state prefers unilateral defection to bilateral cooperation and prefers bilateral cooperation to bilateral defection. The multilateral context of this game is captured by the assumption that some or all of the payoffs in each of the component bilateral games are functions of the number of states in the multilateral context, and the number of states cooperating. Given the assumptions of the prisoner’s dilemma, we have the standard result that all states defect. In other words, no CIL rule will be formed.

Assume instead that the game is repeated indefinitely. To make this more concrete, assume that in any period each state believes that this game will be played again with some probability less than certainty. Further suppose that each state has a particular level of patience represented by a discount factor. Now consider whether cooperation can be sustained as an equilibrium when the game is repeated indefinitely. For this to be the case, cooperation must be a “subgame perfect equilibrium,” meaning in this case that cooperation induces a set of strategies such that each state’s strategy is an optimal response to the other states’ strategies at every repetition of play. One such strategy profile that has the potential to support such an equilibrium is the grim trigger strategy:

1. Cooperate if both states have cooperated in all prior periods;
2. Defect in this and all subsequent periods if either state has defected in any prior period.

As discussed above, an alternative strategy profile, addressing the problem of credibility (collective rationality), is “penance:”

1. Cooperate if both states have cooperated in all prior periods;
2. If one state defects in one period, the other state defects in all subsequent periods until the initially defecting state cooperates.
3. After the initially defecting state cooperates for one period while the other state defects, the latter state returns to cooperation.

Under either of these strategies, if both states are sufficiently patient—sufficiently value future payoffs—both states will cooperate in all periods. The question of the sufficiency of their patience is termed the “Patience Condition.” Stated differently, if both states’ actual discount factors exceed a calculated “critical discount factor,” cooperation will ensue.¹⁰⁴

The Patience Condition can be interpreted in other ways that are directly relevant to our analysis. First, CIL is more likely to emerge and be sustainable when the returns to cooperation are high relative to non-cooperation and when the returns to unilateral defection are low. Second, CIL is more likely to be formed between relatively patient states: those with relatively high discount factors. Third, CIL is more likely where the probability of continued interaction between the participating states is high.

The important question to which we now turn is that raised by Goldsmith and Posner. Does increasing the number of participants make it tougher to sustain cooperation? For this to be the case it is necessary that the critical discount factor becomes higher—more difficult to meet—as the number of states involved increases.

¹⁰⁴ Grim trigger is not necessarily tougher than penance. While in some cases grim trigger will provide a greater incentive to cooperate than penance, the opposite will be true if the gain to defection when the other state cooperates is less than the gain to defection when the other state defects. We show this mathematically in Appendix I.

The simplest, but probably least likely, case is that in which none of the pay-offs are affected by the number of states involved. If this is the case, then the multilateral CIL rule is no more than the aggregation of a series of independent bilateral CIL rules and so is no more difficult to sustain than the individual bilateral rules.

It is more likely, however, that the multilateral context has some relevance in that the number of states that are effectively party to the multilateral CIL rule affects some of the payoffs.¹⁰⁵ This force could derive from the public goods or network aspect of the establishment of the CIL rule itself, a point to which we turn below, or more generally from the possibility that the greater the number of states that form a CIL rule, the greater will be the aggregate benefits that flow from that rule.

However, the situation that is most often considered has the following characteristics. First, the return to each state from cooperation decreases as the number of states in the CIL rule increases. Second, the return to a state from defection is greater when there are more states that continue to abide by the CIL rule. Third, non-formulation of a CIL rule leaves states in an autarkic situation, so that the returns to non-formulation or total breakdown of a proposed CIL rule are the same and independent of the number of states. Where this is the case, it follows that the critical discount factor is increasing with the number of states involved so that multilateral CIL is, indeed, harder to sustain than bilateral CIL.

This is, however, not the same as saying that such CIL rules are *impossible* to sustain or even highly unlikely to be sustained. First, we are more likely to see multilateral CIL among states that have, and are expected to have, interactions over an extended period. Second, multilateral CIL rules are more likely between “patient” states, meaning states that tend to value future payoffs more highly than others. Third, multilateral CIL rules are more likely to hold when the relevant interactions are frequent.

Moreover, there are three additional countervailing forces that can work to sustain multilateral CIL.

The first follows from our analysis above. It is not difficult to imagine circumstances in which the gains to unilateral defection fall with the number of states while the share to each state from cooperation increases with the number of states.

¹⁰⁵ For a recognition by the International Court of Justice that even protection of diplomats has an important multilateral dimension, see *Diplomatic and Consular Personnel Case*, *I.C.J. Reports*, 1980 p. 3 at pp. 42-3 (para. 92). In that case, the ICJ made the following statement:

In recalling yet again the extreme importance of the principles of law [protection of diplomats] which it is called upon to apply in the present case, the Court considers it to be its duty to draw the attention of the entire international community, of which Iran itself has been a member since time immemorial, to the irreparable harm that may be caused by events of the kind now before the Court. Such events cannot fail to undermine the edifice of law carefully constructed by mankind over a period of centuries, the maintenance of which is vital for the security and well-being of the complex international community of the present day, to which it is more essential than ever that the rules developed to ensure the ordered progress of relations between its members should be constantly and scrupulously respected.

Where this is the case the conventional argument is actually reversed. Multilateral CIL rules are *easier* to sustain than bilateral rules.

Under what circumstances might this arise? One possibility¹⁰⁶ is that the rule relates to the investment by each participating state in the provision of a public good.¹⁰⁷ As more states participate, the investment in provision increases as does the benefit to the individual state from cooperation. By contrast, the gains from defection can be expected to decrease with the number of states. The same is likely to be true for policies with strong network effects – for example, common international technology standards.¹⁰⁸ A single state may gain from going it alone but the potential gains are likely to be less the more isolated the state is, that is, the more that the state looks like a single stand-out where others adopt a common standard. A third possibility is that there are reputation effects. A state gains from defecting on a rule but the act of defection harms the state's reputation, making it less likely that the state will be able to make agreements with other states.

The second countervailing effect arises when states are involved in a series of international rules (including treaty rules) with overlapping groups of partners. In such a case, defection on one rule has potentially harmful effects for *all* of the rules to which a particular state is party. This is similar to the analysis of the potential for multimarket contact to maintain tacit agreements between firms.¹⁰⁹

In any multilateral context in which states enter into bilateral as well as multilateral rules, the critical discount factor decreases as the number of bilateral rules that each state enters into increases. This leads to a simple but compelling proposition. *A state can use slack enforcement power in bilateral rules to sustain multilateral rules.*

The third possibility is related to the second. Multilateral rules are more easily sustained when they involve frequent interactions between the member states.

These latter two possibilities can, of course, interact. States may be able to enter into multiple rules, some with frequent interaction and some with infrequent interaction. By the same argument as above, *a state can use slack enforcement power from rules with frequent transactions to sustain rules with infrequent transactions.*

Thus, there exists a significant set of cases in which it will be possible to form multilateral customary rules of international law. The likelihood of formation in any particular circumstance will depend on a number of factors, including (i) the relative

¹⁰⁶ See, e.g., Robert Pahre, *Multilateral Cooperation in an Iterated Prisoners' Dilemma*, 38:2 J. CONFL. RES., 326 (1994); Snidal, *supra* note 34, at 929.

¹⁰⁷ For a discussion of global public goods, see PROVIDING GLOBAL PUBLIC GOODS: MANAGING GLOBALIZATION (I. Kaul *et. al.*, eds. 2003).

¹⁰⁸ Other areas of international law, including the law of sovereignty, diplomatic protection, etc., may exhibit network effects. These may arise simply from efficiencies due to consistency of arrangements. For a relevant analysis in the corporate law field, see Michael Klausner, *Corporations, Corporate Law and Networks of Contracts*, 81VA. L. REV. 757 (1995).

¹⁰⁹ The classic reference on this is Bernheim & Whinston, *supra* note 87.

value of cooperation versus defection, (ii) the number of states effectively involved, (iii) the extent to which increasing the number of states involved increases the value of cooperation or the detriments of defection, including whether the particular issue has characteristics of a commons problem, a public good, or a network, (iv) the information available to the states involved regarding compliance and defection, (v) the relative patience of states to realize benefits of long-term cooperation compared to short-term defection, (vi) the expected duration of interaction, (vii) the frequency of interaction, and (viii) whether there are also bilateral relationships or other multilateral relationships between the involved states.

6. Implications

The model presented above and in Appendix I suggests that CIL may affect behavior, and that it will do so to varying extents under varying circumstances. This model has implications for international legal theory, for CIL doctrine, for policy and for research.

a. International Legal Theory Implication: CIL May Affect Behavior

The discussion in parts 2 and 3 and the model described in part 5 and in Appendix I suggest that there exists a significant set of cases in which CIL will affect behavior. As we cannot here assess in any particular case the actual value to states of cooperation, or their discount rates, or many of the other factors included in the model, it is impossible to say with certainty by how much, or how often, CIL affects behavior. But it is equally impossible to say that it does not affect behavior, that it seldom does so, or even that it only has marginal effects. So, we believe that CIL is plausible.

CIL seems no less plausible than social norms in the domestic context. In fact, there are reasons to believe that it may be more plausible, based on economies of scale and scope in information and multi-sector contact. On the other hand, there are countervailing factors, including the possibility that individuals in small communities have a greater sense of permanence than governments in international society, perhaps making them more patient. Interestingly, this analysis suggests that states generally have an interest in the development of domestic political institutions that will instill “patience” in other states.

b. Normative Implications: Institutional Modifications

The prior analysis suggests a number of potential implications for policy. “Perhaps game theory’s greatest potential for contributing to international law is to provide a rigorous means of describing and articulating important aspects of state interaction and cooperation. The hope is that fully developed game theoretic models will help states design law that creates or enhances the conditions for cooperation, if such cooperation is desirable.”¹¹⁰

¹¹⁰ Chinen, *Game Theory and Customary International Law*, *supra* note 1, at 155.

i. The Role of Regional or Plurilateral Custom

As demonstrated above, the number of states involved in forming a particular customary rule may have a significant effect on the ability to form a rule. As we have suggested, this effect will differ in direction, depending on the context. Therefore, states may find that they can develop regional or other plurilateral rules of CIL in circumstances where multilateral rules are more difficult to establish. Regional or other plurilateral intensification of relationships, such as in the European Community or the OECD, may establish the conditions for greater use of custom.

ii. Network Effects in Institutionalism

International cooperation in different sectors may be mutually supportive, and there may be a kind of network effect that makes each additional instance of cooperation more attractive than it would be absent existing instances. This game theoretic perspective provides support for the early neo-functionalism hypotheses regarding international economic integration, and suggests the potential value of cooperation “for its own sake” or in order to facilitate further cooperation. It also provides theoretical support for strategies of “constructive engagement” outside the CIL context. Network effects may be enhanced in regional or other plurilateral¹¹¹ contexts, by concentrating and intensifying relationships.

iii. The Information Role of NGOs and International Courts

By disseminating information regarding compliance with particular norms, NGOs or international organizations may play a critical role in improving information, facilitating the development of CIL. In cases of complex rules or facts, where under a regime of auto-interpretation states may argue over compliance, courts or other “independent” third parties may resolve this information problem more definitively.

iv. Custom and Treaty

This article applies the literature of law and social norms in a way that suggests a bridge over the gulf between law and social norms. It recognizes that law on the one hand, and social norms on the other hand, are alternative or sometimes complementary means of social control, or social cooperation. It thus suggests the utility of comparative institutional analysis¹¹² between law, on the one hand, and social norms, on the other.

Custom is a mechanism for international “legislation” that requires only a degree of consensus,¹¹³ not unanimity. Given the difficulty of establishing global treaties without significant holdouts, and given the need to avoid free riders, we might understand the CIL process as an alternative mechanism for global legislation. A rule of consensus

¹¹¹ By “plurilateral” we mean groups of states that are subsets of the broader group.

¹¹² See NEIL KOMESAR, *IMPERFECT ALTERNATIVES* (1996); HERVÉ MOULIN, *COOPERATIVE MICROECONOMICS: A GAME-THEORETIC INTRODUCTION* (1995).

¹¹³ “Consensus” is here understood as an absence of express objection.

acts as a default rule that promotes compliance, and increases the bureaucratic costs of “persistent objectors.” We recognize that this type of strategy would raise concerns regarding democratic legitimacy. However, this type of legislative technique is not more invasive than majority voting, and the “persistent objector” rule allows states to opt-out of rules that raise sufficient concern.

There is less of a distinction between CIL and treaty, on the one hand, than there is between social norms and either contract or law, on the other hand. As treaty is in legal theory only supported by CIL and by institutions created by treaty, rather than by an external state, it is something less than contractual. In fact, we may understand treaty in the international law context as an extension of custom, or of social norms. It is an important extension, with the capacity to specify required performance in greater detail, establishing greater agreement on the content of the relevant norm, and thereby overcoming important information problems. Of course, to the extent the treaty specifies binding dispute settlement, additional information problems may be overcome.

Furthermore, treaty has a greater capacity for concreteness than custom, and treaty lends itself more to specificity—to rules as opposed to broad standards. Treaty also is more amenable to domestic ratification, which is both a burden in terms of efficiency of agreement, and a benefit in terms of accountability. As in the rules and standards context,¹¹⁴ custom may serve as a pathfinder for later-established specific treaty rules. Conversely, treaty structures, including dispute settlement, may serve as an institutional setting to promote custom.

Further analysis of the relationship between treaty and custom, and the choice of instruments in particular contexts, is beyond the scope of this article.

c. Doctrinal Implications: A Contractual Approach to *Opinio Juris*

We have not yet directly addressed the argument by Goldsmith and Posner that CIL generally does no work—that state behavior is not motivated by CIL, but only by self-interest.¹¹⁵ The implication of this argument, not made explicit by Goldsmith and Posner, is that CIL does not exist, because CIL doctrine requires practice motivated by *opinio juris*—by CIL. This argument is subject to several responses.

i. CIL Rules May Affect Behavior

First, our discussion in parts 2, 3 and 5 shows that CIL may affect behavior. So our refutation of the assumption that the multilateral prisoner’s dilemma is unlikely to be resolved shows that it is plausible that state behavior is affected by CIL.

ii. A Contractual Approach to *Opinio Juris*

¹¹⁴ For an introduction to the rules versus standards discussion in law and economics, see Kaplow, *supra* note 30.

¹¹⁵ Goldsmith & Posner 1999, *supra* note 1 at 1131-1133.

Recall that under the Restatement Third formulation, CIL does not exist without *opinio juris*, or the sense of legal obligation. However, as D'Amato suggests,¹¹⁶ it seems on first analysis that this requirement is circular: the first state that complies "from a sense of legal obligation" must do so erroneously. There may be a solution to this paradox.

Analogizing CIL to social norms, we might postulate that instead of a "sense of legal obligation," the Restatement Third formulation might more correctly refer to an "intent to create or accept a rule of law."¹¹⁷ As suggested by the formulation contained in Article 38(1)(b) of the Statute of the International Court of Justice, we may in addition refer to a "sense of *incipient* legal obligation."¹¹⁸ CIL may arise in the international system in just the way that social norms arise in the domestic setting, with the same possible beneficial effects in terms of cooperation and coordination. The social norms analogy suggests, with McDougal, Mendelson, Swaine, and Thirlway¹¹⁹ that we may understand the initial act of "compliance" not necessarily as an error, but as an offer or an act of leadership. The offer and acceptance must generally take the form of practice, although our model does not require us to take a position on the amount of practice that will be necessary to constitute either an offer or an acceptance.

Consider the development of social norms in the municipal setting. The first time that in Shasta County, a cattle farmer returned a lost calf to its owner without charge, despite a lack of legal obligation, there existed no relevant social norm. Yet that action, perhaps recognized by its author to provide efficiencies that would eventually benefit him if multilateralized, or perhaps motivated by something else but interpreted as a proposal to initiate a rule, began a process that resulted in a "social norm."

If we understand the "sense of legal obligation" referenced in the Restatement Third not as a sense of a fully-formed legal rule, but as a perception or assertion that a *legal* rule would be beneficial, the circularity problem is resolved. Thus, custom must be

¹¹⁶ ANTHONY A. D'AMATO, *THE CONCEPT OF CUSTOM IN INTERNATIONAL LAW* 5 (Cornell 1971).

¹¹⁷ It is worth noting the relationship of this proposition to the first formulation of Kant's categorical imperative: "Act only on that maxim through which you can at the same time will that it should become universal law." IMMANUEL KANT, *GROUNDWORK OF THE METAPHYSICS OF MORALS* (Mary Gregor trans. & ed., Cambridge Univ. Press 1997) (1785).

¹¹⁸ See *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, 1986 I.C.J. 14, 109 (June 27) ("Reliance by a State on a novel right, or an unprecedented exception to the principle, might if shared in principle by other States, tend towards a modification of customary international law.").

¹¹⁹ "[A]t the initial stage of the development of the custom, it is sufficient that the States concerned regard the practice as what the Court . . . referred to as 'potentially norm-creating', as conforming to a rule which either already exists or is a useful and desirable rule which should exist." Hugh Thirlway, *The Law and Procedure of the International Court of Justice 1960-1989 (Part Two)*, 1990 BRIT. Y.B. INT'L L. 1, 43, citing *North Sea Continental Shelf (F.R.G. v. Den.; F.R.G. v. Neth.)*, 1969 I.C.J. 3, 42 (Feb. 20). See MCDUGAL ET AL, *STUDIES IN WORLD PUBLIC ORDER* 773-74 (1960); Mendelson, *supra* note 11; Swaine, *supra* note 1, at 615 (suggesting that states that initiate a custom might be analogized to offerors in a contractual setting, and that this conditional obligation is consistent with *opinio juris*). See also D'Amato, *supra* note 116, at 73-102. And what of the circumstances in which initial acts are unreciprocated? We might understand these acts as transaction costs: as costs of identifying feasible transactions or appropriate partners. This is not very different from gift-giving practices in many village and other social circumstances. See Posner, *supra* note 15, 49-67.

understood not as mere action, but as an initial or continuing proposal for collective action over time, with acceptance evidenced by compliance. A state may test a proposed rule of collective action informally, without the domestic or international costs that otherwise might attend the proposal.¹²⁰ “Run it up the flagpole and see if anyone salutes.” Supporting this approach, a recent report of the International Law Association explained that *opinio juris* requires practice “in circumstances which give rise to a legitimate expectation of similar conduct in the future.”¹²¹ This understanding also offers a plausible explanation of changing rules of CIL. In fact, there is no real difference between initiation and change: initiation of a rule is a change from a *laissez-faire* rule.

Thus, there is a rationale for the *opinio juris* requirement in terms of general state intent:¹²² mere regularity of action, or mere action based on motives that do not include the formation of a legal rule, cannot form a rule of CIL. This is recognized in Article 38 of the Statute of the International Court of Justice, which specifies “international custom, as evidence of a general practice accepted as law.”¹²³ Goldsmith and Posner seem correct that motivation by narrow coincidence of interest (the “required” behavior intrinsically benefits the actor) or narrow coercion (of a more direct and unilateral type than the retaliation included in our model) are inconsistent with *opinio juris* and so do not contribute to CIL. But even here, states are likely to act with varying and multiple motivations, and it is possible that a CIL rule could be formed despite the fact that some states adhere largely due to narrow coincidence of interest or narrow coercion. Moreover, it is incorrect to argue that where states act out of self-interest, no *opinio juris* can exist.

iii. CIL Rules May be Coterminous with Self-Interest

¹²⁰ See Parisi, *supra* note 1, at 18 (describing “articulation theories” of CIL in these terms). Parisi suggests special deference to rules chosen prior to conflict: “Articulations that are made prior to unveiling of conflicting contingencies can be analogized to rules chosen under a Harsanyiian veil of uncertainty.” *Id.* At 19. This condition is referred to as “stochastic uncertainty.” However, one might ask why there would be an incentive to produce new rules in the absence of conflict. Perhaps the answer to this problem is that the formation of custom is a process occurring over time, and states may participate before they know their individual contingent position. Furthermore, one might question whether the veil of uncertainty does not apply even after conflict arises, to the extent that states may not know their position in future instances of conflict.

¹²¹ COMM. ON FORMATION OF CUSTOMARY (GEN.) INT’L LAW, INT’L LAW ASS’N, FINAL REPORT: STATEMENT OF PRINCIPLES APPLICABLE TO THE FORMATION OF GENERAL CUSTOMARY INTERNATIONAL LAW §1(i) (2000), <http://www.ila.hq.org/pdf/CustomaryLaw.pdf>, citing Mendelson, *supra* note 11, at 399.

¹²² What of states that have no relevant intent, or that object to the formation of a legal rule? We will not deal with the doctrinal problem of the persistent objector. But it seems plausible that states would accept a rule of consensus legislation in international custom: a rule may be formed that binds states that do not object. For an analysis of consensus-based decision-making within the GATT/WTO system, see Richard Steinberg, *In the Shadow of Law or Power: Consensus-Based Bargaining and Outcomes at the GATT/WTO*, 56 INT’L ORG. 399 (2002).

¹²³ The International Court of Justice seems to ignore the text in favor of a reverse reading, seeking to find evidence of international custom in a general practice accepted as law. ROSALYNN HIGGINS, PROBLEMS & PROCESS: INTERNATIONAL LAW AND HOW WE USE IT 18-19 (1994). Of course, there are other formulations. See, e.g., IAN BROWNLIE, PRINCIPLES OF PUBLIC INTERNATIONAL LAW 4-11 (5th ed. 1998).

Furthermore, there is a terminological or doctrinal problem with Goldsmith and Posner's argument, making it seem tautological. Goldsmith and Posner argue that "[s]tates do not comply with CIL because of a sense of moral or legal obligation; rather, CIL emerges from the states' pursuit of self-interested policies on the international stage."¹²⁴ Unless Goldsmith and Posner mean merely to refute the natural law position that states comply with international law because of its normative appeal or legitimacy or because it is the right thing to do, this argument is a non-sequitur, as legal obligation and self-interest are not mutually exclusive categories. Law and economics has long utilized price theory to understand behavior under legal rules, and there is no question that law can affect behavior *through* self-interest. In the CIL setting, the motivating force is wider or potentially longer-term self-interest that flows from making and achieving compliance with a rule, or even with rules generally: narrow or diffuse reciprocity.¹²⁵

Goldsmith and Posner's main point here must then be understood simply as the application of an *assumption* of the rationalist model (and one that is subject to at least some contention): that agents only care about their own utility, and therefore would not follow a rule of CIL for *intrinsic* reasons: for its own sake.¹²⁶ However, compliance with international law resulting from a sense of legal obligation is consistent with *extrinsic* reasons: there may well be utility in upholding a particular rule, or in upholding the rule of law in general.

iv. Evidentiary Requirements

Third, a marginalist approach would reject the evidence presented by Goldsmith and Posner as a basis for their argument. Their case-based evidence merely suggests in a limited range of contexts that there are plausible non-CIL reasons for observed behavior.¹²⁷ Goldsmith and Posner adduce no data that suggests the relative magnitude of these reasons, or that suggests the absence of other reasons. So, they and we have no way of knowing that CIL was not a contributing, or alternative sufficient, cause of behavior.¹²⁸ Survey or interview data might be useful to fill this gap.

Of course, to the extent that a particular instance of compliance is fully and exclusively explained by true coincidence of interest or coercion, it cannot be argued that CIL did any work.¹²⁹ Our theory of CIL examines the effects of broader self-interest based on reciprocity, and accepts the possibility of a different kind of coercion by punishment for defection, as a normal part of the CIL process. While CIL is endogenous

¹²⁴ Goldsmith & Posner 1999, *supra* note 1, at 1115.

¹²⁵ See Robert O. Keohane, *Reciprocity in International Relations*, 40 INT'L ORG. 1 (1986).

¹²⁶ Recall that there is an important literature on social norms that considers the potential for preference modification, holding out the possibility of compliance because of modified preferences. See sources cited at note 25, *supra*. It may not be necessary to choose between static and dynamic preferences: changing preferences may work together with static preferences and self interest.

¹²⁷ *But see* Swaine, *supra* note 1 (raising concerns regarding Goldsmith and Posner's perspective on CIL doctrine).

¹²⁸ *Cf.* Guzman, *supra* note 1, at 1875 (asserting that Goldsmith and Posner's claim that CIL does not affect state behavior goes beyond what the evidence suggests).

¹²⁹ For some criticism of Goldsmith and Posner's description of CIL doctrine, see Vagts, *supra* note 2.

to states in the aggregate, once formed it is at least largely exogenous to the particular state. As explained above, we also find it plausible that there are important circumstances in which CIL would have significant effects on state behavior.

v. Violations or Proposals for Change?

Related to the problem of determining whether a custom has the requisite motivation under CIL doctrine is the question of how to deal with anomalous conduct. Does the anomaly constitute the initiation of a revised rule of CIL, or a simple violation?¹³⁰ It is important to recognize that no law, in any system, achieves perfect compliance. Thus we must determine another way to evaluate compliance. The best way is to evaluate the *extent* to which law affects behavior. Thus, the fact that wars occur does not alone mean that the international legal prohibition on the use of force is without effect, just as the fact that murders occur does not mean a domestic proscription of murder is meaningless. Goldsmith and Posner argue that variations in levels of compliance suggest that no multilateral rule exists, or affects state behavior.¹³¹ However, we would expect systematic circumstantial variations in compliance with respect to all laws. That is, in order to determine that CIL exists, or that we as social scientists should pay attention to it, it need not in every case *determine* behavior, so long as it may do so in some set of marginal cases. In game theoretic terms, even a small effect of international law could affect the payoffs from compliance, potentially tipping the balance in favor of compliance in a marginal case.

Furthermore, in order to maintain the dynamic, evolutionary character of CIL, it is necessary that any theory of CIL allow for some violations of existing rules to be understood as proposals for the establishment of new rules.¹³² This is indeed a subtle and difficult distinction, but it is necessary in any decentralized system. Thus, we should not demand that CIL command absolute compliance, or be inflexible. In fact, one might argue, as some do about the common law, that one of its great virtues is its dynamism.

d. Research Implications: An Empirical Research Agenda

Theory alone tells us little about the world. The next step is to develop and test hypotheses based on the theory of CIL elaborated here.

i. Assessing the Patience/Horizon of States

In empirical research, it would be useful to determine parameters or proxies by which to assess the patience (including the horizon) of states and governments, and to regress these parameters against measures of compliance. Is political stability associated with patience, and do we see greater compliance with CIL by states with greater political stability? Are democratic states, or states with better developed financial markets, more

¹³⁰ See D'Amato, *supra* note 116, at 97 (stating that "an 'illegal' act by a state contains the seeds of a new legality").

¹³¹ Goldsmith & Posner 1999, *supra* note 1, at 1157.

¹³² See, e.g., MALCOLM N. SHAW, *INTERNATIONAL LAW* 69 (4th ed. 1997).

patient? Are autocratic states more patient because they are not concerned about election cycles? Do we see greater amounts of violation of law before an election than after? Is patience determined by a bell curve, in which strong autocracies and stable democracies are patient, and those in between are not? Are corrupt governments “impatient”? What about more-or-less independent bureaucracies that may be charged with action that determines compliance with specific rules of CIL? Can these independent bureaucracies exhibit greater patience than their elected governments? These conjectures are linked to liberal states theory,¹³³ as well as to theories of transnational governmental networks.¹³⁴ In fact, the correct level of analysis for compliance with certain rules of CIL, in terms both of patience and information transmission, may be the bureaucratic division, rather than the state.

ii. Proxies for Payoffs

Once we have developed empirical methods of measuring patience, it would be necessary to develop proxies for payoffs from violation and compliance, in order to know determine whether compliance occurs as predicted by the model described here.

iii. Payoff Structure

Is CIL more likely to be formed under circumstances involving public goods or networks, where the value of cooperation may rise with the number of players?

iv. Network Externalities in International Law: More Relationships Make Each Relationship More Reliable

Do we see an acceleration of custom, or a tipping point at which sufficient relationships are established to make compliance with CIL more likely? Is there a synergy between treaty and CIL? This hypothesis might be tested by examining the relationship between entry into treaties and compliance with CIL. Again, the possibility of network externalities is inconsistent with a disaggregation of the state into independent functional components.

v. Regional Custom

To what extent can we identify regional or other plurilateral CIL, and how does it compare in terms of compliance to universal CIL?

vi. Information

Is there a relationship between the establishment of NGOs that enhance information regarding compliance and defection, or the establishment of adjudicative bodies, and compliance with CIL?

¹³³ See Anne-Marie Slaughter, *International Law in a World of Liberal States*, 6 EUR. J. INT’L L. 538 (1995).

¹³⁴ See Anne-Marie Slaughter, *The Real New World Order*, 76 FOR. AFF. 183 (1997).

7. Conclusions

This article shows the rational plausibility of CIL. It thus serves as a basic defense of the international law system. Indeed, this article presents not just a theory of CIL, but also a theory of international law more generally. The theory presented here is based on methodological and normative individualism. It thus departs substantially from the airy idealism of natural law theory. It serves as a positive social scientific theoretical basis for international legal positivism and a nuanced rationalism, basing law on the action of states or their agents in pursuit of their self-interest, broadly understood.¹³⁵ It shows that CIL, and with it treaty law, is something of a levitation feat. It therefore rests not on a rock-solid natural law basis of divine principles, but on a fabric of rational acts, woven through a multiplicity of relations over time.

The goal of this article was to develop a model that would generate interesting hypotheses about compliance with CIL. In this regard, our model should be compared with Goldsmith-Posner's approach, which would suggest that there are no interesting hypotheses about CIL because it cannot affect behavior. More particularly, Goldsmith and Posner provide no analysis of the circumstances under which the multilateral prisoner's dilemma, or the bilateral prisoner's dilemma, might yield stable and efficient equilibria. In that sense, our article is quite different, and presents the possibility for a progressive research program in CIL.

Some may ask, is this article about law, or is it only about social order, labeled "law." In a sense, of course, this article does not need to distinguish CIL from other forms of cooperation--and indeed the basic model is generic, although we make assumptions to accord with the CIL system, such as the general doctrinal rule of bilateral retaliation.

So, while it is true that our basic model is a model of cooperation, and it applies to CIL, to treaty, and to other forms of international cooperation, there are some distinctive and important aspects of the CIL supergame that do not apply to general cooperation. First, CIL rules may serve as equilibrium-selection devices, providing a greater possibility for a stable equilibrium. Second, designation as CIL serves to incorporate a substantial set of default rules that fill in a large portion of the "incomplete contract" regarding many aspects of the obligation, including the scope of remedies for violation. Third, it may be that designation as CIL serves to link compliance with any particular CIL rule to others in order to widen the possible scope of retaliation to fields that might not be considered "legitimate" otherwise. In this last connection, we might say that designation as CIL increases the returns to compliance by placing at stake the general sense of international legality. That is, if state A can be a scofflaw in one sector, what prevents state B from being a scofflaw in an area that injures state A? In this sense, there is a possibility for implicit multilateral retaliation, even if formal CIL doctrine does not permit multilateral retaliation.

¹³⁵ This article does not by any means challenge the theory that law can affect behavior by modifying preferences. It merely presents a theory that does not depend on modifying preferences.

This article also serves as a refutation of the central claim of structural realists in political science with respect to international law: that it is epiphenomenal. This article shows that law may well be a social expression of rules that achieve real collective goals, are backed by real sanctions, and have real behavioral effects. It is a strange realism that would ignore this.

If social norms theory in the domestic sphere finds social norms attractive as a mechanism for production and enforcement of rules, perhaps CIL holds promise as an alternative to treaty. Perhaps the main distinguishing feature, and potential value, of CIL is systemic. That is, although we have assumed sectoral divisions for modeling purposes, international law may also be understood as a set of linked games, or one extensive game. Once a particular rule is absorbed into the CIL system, or is established through treaty, it may benefit from linkage to other rules of CIL, and of treaty.¹³⁶ The special nature of legal rules may derive simply from their reception into this linked system. It is order and law.

Like all positive social science, this theory has normative implications. CIL has advantages and disadvantages as a process for making rules. As states identify these advantages and disadvantages in particular contexts, they may decide in some contexts to facilitate the development of CIL through institutional modifications.

* * *

¹³⁶ See Swaine, *supra* note 1 (“states do not, in fact, interact solely with respect to one rule or the other, and it is also possible to understand their interaction with respect both to an individual rule and to the system of customary international law”).

Appendix I

The repeated prisoner's dilemma game has two important features, in common with all repeated games. First, the strategies played by each state at time t can be made conditional on the history of play up to time t . Second, these strategies can include the possibility of punishment if a state deviates from an implicit agreement to follow a particular strategy.

Assume that state i is a representative state in an n -state group that has implicitly agreed to cooperate on some aspect of its international relations—for example, fishing rights, protection of diplomats, expropriation, or technology standards. Further suppose that a subgroup m of the universe of states n , such that $m \leq n - 1$, other than state i , have also promised cooperation. If state i sticks by its promise to cooperate it receives $C_i(n; m)$. If state i defects on its promise against some group g of the cooperating states it receives $G_i(n, g; m)$. We further assume that if state i defects it triggers punishment only by those states adversely affected by the defection.¹³⁷ If state i cooperates while group g defects against i then state i receives $L_i(n, g; m)$. Finally, if i defects against group g and this group moves to its punishment strategy against i , then i receives $D_i(n, g; m)$.¹³⁸

This game is an n -state prisoner's dilemma game when the following conditions hold for each state i in the group and for all $m \leq n - 1$ and $g \leq m$:

- (i) $C_i(n; m) > D_i(n, g; m)$: Cooperation with group g is better for state i than mutual defection no matter the number of other states that are cooperating or the size of group g ;
- (ii) $G_i(n, g; m) > C_i(n; m)$: State i is better off defecting against group g than cooperating with g no matter the number of other states that are cooperating or the size of group g ;
- (iii) $L_i(n, g; m) < D_i(n, g; m)$: If group g defects against i then i is better off defecting than continuing to cooperate with g .

Given that conditions (i) – (iii) are satisfied, we have the standard result that defection is the dominant strategy for every state, no matter the number of states that are cooperating. The Nash equilibrium to this game if it is played only once, or a finite number of times known in advance to all the participants, is that all states defect. In other words, even if a CIL rule is articulated it will not be followed.

Assume instead that the game is repeated indefinitely. To make this more concrete, assume that in any period t state i believes that this game will be played again in period $t + 1$ with probability $\rho_i \leq 1$. Further suppose that state i has discount factor R_i , where $0 < R_i < 1$ is a measure of impatience: the smaller is R_i the less patient state i is –

¹³⁷ This keeps the notation reasonably simple and is not unreasonable in the context of international law. The formal analysis is unchanged if we assume that defection against g induces punishment by some group h , which may include g plus other states. Most analyses of the n -person prisoner's dilemma assume, in fact, that h is the full set of cooperating states other than i .

¹³⁸ For the sake of notational simplicity we assume in this part of the discussion that returns to A are determined by the number of cooperating, punishing and defecting states but not by their identities. We return to this assumption below.

the more that state i favors current benefits over future benefits. We can then construct a “probability adjusted” discount factor $\delta_i = \rho_i R_i$ for each state.

Given that we are interested in the conditions that will sustain full cooperation of all n states, we confine our attention to the case in which $m = n - 1$. We denote the group “all states other than i ” by n_{-i} . Now consider whether the strategy combination {Cooperate by i , Cooperate by n_{-i} } can be sustained as an equilibrium for each state i when the game is repeated indefinitely. For this to be the case, {Cooperate by i , Cooperate by n_{-i} } must be a subgame perfect equilibrium. One strategy profile that has the potential to support such an equilibrium is the grim trigger strategy for each state:

1. Cooperate in period t if all states have cooperated in all periods up to and including $t - 1$;
2. Defect against a defecting state in period t and all subsequent periods if the defecting state has defected in any period prior to t .

As discussed in Part 2, an alternative strategy, addressing the problem of collective rationality, is “penance.” Fudenberg and Tirole show that the following “penance” strategy profile for each state i is “weakly renegotiation proof” (WRP): “Begin in the cooperative phase where all states play Cooperate. If a single state A switches to Defect against i , then i switches to the punishment phase for A . In this phase, state A plays Cooperate and state i plays Defect against A . Play remains in this phase until the first time state A plays Cooperate with i , at which point play returns to the cooperative phase.”¹³⁹ The logic of this strategy is that a state having defected from a rule can have the rule reinstated only by accepting a period of punishment, in which it cooperates while the states against which it has defected defect against it.

Whether we use the grim trigger or the penance strategy profile, standard analysis indicates that we need only consider a single-period defection by a state.¹⁴⁰ Suppose, then, that all states other than i are Cooperating and that state i chooses to Defect against group g in period 0.¹⁴¹ Its return from doing so is:

$$S_D = \begin{cases} G_i(n, g : n - 1) + D_i(n, g : n - 1) \frac{\delta_i}{1 - \delta_i} & \text{with grim trigger} \\ G_i(n, g : n - 1) + \delta_i L_i(n, g : n - 1) + C_i(n : n - 1) \frac{\delta_i^2}{1 - \delta_i} & \text{with penance} \end{cases}$$

If, by contrast, state i chooses Cooperate in every period, its return is:

¹³⁹ Fudenberg & Tirole, *supra* note 383838, at 180. See also the “defect for deviate” strategy proposed by Mahoney and Sanchirico, *supra* note 262626.

¹⁴⁰ We could also assume with grim trigger that states switch to Defect for a finite number of periods T without affecting the qualitative conclusions.

¹⁴¹ Again, standard analysis indicates that we can always choose the first period in which Defect is played as the beginning of the game.

$$S_C = \frac{C_i(n:n-1)}{1-\delta_i}$$

For Cooperate to be preferable for state i , and therefore sustainable, we must have $S_C > S_D$ which, after some manipulation, requires:

Patience Condition:

$$S_C > S_D \text{ if and only if } \delta_i > \begin{cases} \delta_g^* = \frac{G_i(n, g:n-1) - C_i(n:n-1)}{G_i(n, g:n-1) - D_i(n, g:n-1)} \text{ with grim trigger} \\ \delta_p^* = \frac{G_i(n, g:n-1) - C_i(n:n-1)}{C_i(n:n-1) - L_i(n, g:n-1)} \text{ with penance.} \end{cases}$$

The ‘‘Patience Condition’’ is that the actual discount factor (δ_i) exceeds the critical discount factor (δ_g^* or δ_p^*).

The first point to note is that, for the Patience Condition to hold, punishment of a defector must be sufficiently large to affect the defector’s behavior. Consider, for example, a case in which state i is ‘‘strong’’ while the states in g are ‘‘weak’’, with the result that the states in g have little ability to punish any defection by i : formally, $D_i \approx C_i \approx L_i$. Then the Patience Condition is unlikely to be satisfied unless the states are ‘‘very’’ patient. This suggests one of two outcomes. Either CIL will not be formed between such states or the states will have to find some way of strengthening the punishment of deviation. One possibility, for example, might be to create an independent arbiter who is capable of coordinating multilateral punishment beyond g in the event that i defects.¹⁴²

A necessary but not sufficient condition for cooperation to be sustainable is that the critical discount factor is less than unity. Condition (iii) above gives the familiar result that, with the grim trigger strategy, provided the participants to the potential rule are sufficiently patient and the probability of continuation is sufficiently high, multilateral cooperation is sustainable. This need not be the case with the penance strategy. This is because $\delta_p^* < 1$ if and only if $C_i(n:n-1) > (G_i(n, g:n-1) + L_i(n, g:n-1))/2$. In other words, the penance strategy will sustain cooperation by state i only if the single-period return to cooperation exceeds the average of the single-period return to defection and return while being punished.

It should not be thought, however, that grim trigger is necessarily a tougher punishment regime than penance and so more able to sustain a cooperative rule. We note that $\delta_p^* < \delta_g^*$ if and only if $G_i(n, g:n-1) - C_i(n:n-1) < D_i(n, g:n-1) - L_i(n, g:n-1)$, in other words, if the gain to defection against g when all other states are cooperating is

¹⁴² We are grateful to an anonymous referee for identifying this possibility. Maggi (1999) identifies the World Trade Organization as such an arbiter of trade agreements. Giovanni Maggi, *The Role of Multilateral Institutions in International Trade Cooperation*, 89 AM. ECON. REV. 190 (1999).

less than the gain to defection when the states in g retaliate. The Patience Condition indicates that penance imposes a tougher punishment if the gain to defection when the other states cooperate is less than the gain to defection when the other states defect.

The Patience Condition can be interpreted in other ways that are directly relevant to our analysis. First, CIL is more likely to emerge and be sustainable when the returns to cooperation are high relative to non-cooperation and when the returns to unilateral defection are low. Second, CIL is more likely to be formed between relatively patient states: those with relatively high discount factors. Third, CIL is more likely where the probability of continued interaction between the participating states is high.

The important question to which we now turn is that raised by Goldsmith and Posner. Does increasing the number of participants make it tougher to sustain cooperation? For this to be the case it is necessary that δ^* rises as n increases.¹⁴³

With some manipulation we can show (suppressing n and g in the interests of brevity) that

$$\text{sign}\left(\frac{\partial\delta_g^*}{\partial n}\right) = \text{sign}\left((C_i - D_i)\frac{dG_i}{dn} + (G_i - C_i)\frac{dD_i}{dn} - (G_i - D_i)\frac{dC_i}{dn}\right)$$

$$\text{sign}\left(\frac{\partial\delta_p^*}{\partial n}\right) = \text{sign}\left((C_i - L_i)\frac{dG_i}{dn} + (G_i - C_i)\frac{dL_i}{dn} - (G_i - L_i)\frac{dC_i}{dn}\right)$$

We are interested in the signs of these derivatives: whether they are positive or negative determines whether δ^* rises as n increases. This is important since, the higher the critical discount factor the less likely it is that the CIL rule is sustainable multilaterally. Note that from (i) – (iii) above, all of the bracketed terms $(C_i - D_i)$ and so on are positive. It follows that the signs of the derivatives $\partial\delta^*/\partial n$ are determined by the signs of the derivatives of the payoffs with respect to the number of states in the multilateral rule.

The simplest, but probably least likely, case is that in which none of the pay-offs are affected by n , that is $dG_i/dn = dD_i/dn = dC_i/dn = dL_i/dn = 0$. If this is the case, then the multilateral CIL rule is no more than the aggregation of a series of independent bilateral CIL rules and so is no more difficult to sustain than the individual bilateral rules.

It is more likely that the multilateral context has some force in that the number of states that are effectively party to the multilateral CIL rule affects at least some of the payoffs. This could arise if there is a public goods or network aspect to the establishment of the CIL rule itself, a point to which we turn below, or more generally from the idea

¹⁴³ There is the related question: What is the optimal deviation for i ? In other words, what is the optimal choice of g ? This question, while interesting, takes us beyond the specific interests of this paper and so is not considered.

that the greater the number of states that form a CIL rule, the greater will be the aggregate benefits that flow from that rule.

The situation that is most often considered has the following characteristics. First, the return to each state from cooperation decreases as the number of states in the CIL rule increases, so that $dC_i/dn < 0$. Second, the return to a state from defection is greater when there are more states that continue to abide by a CIL rule, so that $dG_i/dn > 0$. Third, non-formulation of a CIL rule leaves states in an autarkic situation, so that the returns to non-formulation or total breakdown of a proposed CIL rule are the same and independent of the number of states ($dD_i/dn \approx 0$). Where this is the case, it follows that the critical discount factor is increasing in n so that multilateral CIL is, indeed, harder to sustain than bilateral CIL.

This is, however, not the same as saying that such CIL rules are *impossible* to sustain or even highly unlikely to be sustained. The critical condition is, as noted above, that each state i that is party to the CIL rule has probability adjusted discount factor $\delta_i > \delta^*$. This is more likely to arise if each state believes that the probability of continuation ρ_i is high and has a high discount factor R_i . The implication is first, that we are more likely to see multilateral CIL among states that have, and are expected to have, interactions over an extended period. Second, multilateral CIL rules are more likely between “patient” states, meaning states that tend to value future payoffs more highly than others. Third, multilateral CIL rules are more likely to hold when the relevant interactions are frequent. This is because R_i is a “per period” discount factor and ρ_i is a “per period” probability. The shorter the time period between transactions the higher is the effective probability adjusted discount factor and so the more likely it is that the $\delta_i > \delta^*$ condition will be satisfied.

Moreover, there are at three additional countervailing forces that can work to sustain multilateral CIL.

The first follows from our analysis above. It is not difficult to imagine circumstances in which the gains to unilateral defection fall with n while the share to each state from cooperation increase with n --where $dG_i/dn < 0$ and/or $dC_i/dn > 0$. Where this is the case the conventional argument is actually reversed. Multilateral CIL rules are *easier* to sustain than bilateral rules.

Under what circumstances might this arise? One possibility is that the rule relates to the investment by each participating state in the provision of a public good. As more states participate, the investment in provision increases as does the benefit to the individual state from cooperation. By contrast, the gains from defection can be expected to decrease with n . The same is likely to be true for policies with strong network effects – for example, common international technology standards. A single state may gain from going it alone but the potential gains are likely to be less the more isolated the state is, that is, the more that the state looks like a single stand-out where others adopt a common standard. A third possibility is that there are reputation effects built into $G_i(n)$. A state gains from defecting on a rule but the act of defection harms the state’s reputation,

making it less likely that the state will be able to make agreements with other states. In the context of the game in Table 1, this implies that the gains from current defection will tend to decrease with n .

The second countervailing effect arises when states are involved in a series of international rules (including treaty rules) with overlapping groups of partners. In such a case, defection on one rule has potentially harmful effects for *all* of the rules to which a particular state is party. This is similar to the analysis of the potential for multimarket contact to maintain tacit agreements between firms and is referred to as *issue linkage* in the political science literature.¹⁴⁴ Rather than provide a general analysis, the basic idea can be illustrated by using a variant of one of the examples from Goldsmith and Posner,¹⁴⁵ assuming bilateral punishment and the grim trigger strategy profile.¹⁴⁶ In this game, cooperation by all states on a particular CIL rule gives each state a share $6/n$. Defection by all states gives each a share $4/n$. Defection by a single state gives that state 4 and leaves the remaining states a share $2/n$.

Suppose that there are 2 states, i.e. $n = 2$. Then the critical probability adjusted discount factor above which a bilateral rule is sustainable is, from the Patience Condition:

$$\delta_g^*(2) = \frac{4 - 3}{4 - 2} = \frac{1}{2}$$

By contrast, if there are 4 states the critical discount factor to maintain a multilateral rule between all four is:

$$\delta_g^*(4) = \frac{4 - 6/4}{4 - 1} = \frac{5}{6}$$

This case is, in other words, one in which the multilateral rule is more difficult to sustain than the bilateral rule.

Now suppose that there are four states and that each state enters into a bilateral rule or agreement with one of the other three states as well as a multilateral rule with all

¹⁴⁴ The classic reference on this is Bernheim & Whinston, *supra* note [878787](#). See also Susanne Lohmann, *Linkage Politics*, 41 J. CONFL. RES. 38 (1997) for an analysis of issue linkage in the context of a bilateral prisoner's dilemma game.

¹⁴⁵ Goldsmith & Posner 1999, *supra* note [111](#), at 1125, table 2. Goldsmith & Posner's table 2, relating to the protection of coastal fishing boats, appears as follows:

		State i	
		Attack	Ignore
State j	Attack	2, 2	4, 1
	Ignore	1, 4	3, 3

¹⁴⁶ The more general analysis is available on request.

four. Each rule has the pay-offs given in Table 2. Consider the following grim trigger strategy.

1. Cooperate in period t if all states have Cooperated in all periods up to and including $t - 1$;
2. Defect in period t and all subsequent periods on both rules if any state has Defected on either rule in any period prior to t .

Clearly, if any state is going to Defect it will Defect on both the bilateral and multilateral rules, since the strategy calls for its partner to react to defection by defecting on both rules. The returns to the defecting state from defection are:

$$4 + \frac{2\delta}{1-\delta} + 4 + \frac{\delta}{1-\delta} = 8 + \frac{3\delta}{1-\delta}$$

The returns from continued cooperation are:

$$\frac{3}{1-\delta} + \frac{3}{2(1-\delta)} = \frac{9}{2(1-\delta)}$$

It follows that the critical probability adjusted discount factor above which the multilateral and single bilateral rules are sustainable is

$$\delta^*(4,1) = \frac{7}{10}$$

By the same argument, if a state enters into *two* bilateral rules and a four-state multilateral rule, the probability adjusted critical discount factor is

$$\delta^*(4,2) = \frac{9}{14}$$

Note that this example has the property that the critical probability adjusted discount factor increases as the number of potential partner states increases, making the multilateral CIL rule in this example more difficult to sustain as the number of participants to the rule increases, as Goldsmith and Posner argue. However, our example also illustrates the countervailing power that derives from states being able to enter into bilateral as well as multilateral rules. In any multilateral context, the critical discount factor decreases as the number of bilateral rules that each state enters into increases. This leads to a simple but compelling proposition. *A state can use slack enforcement power in bilateral rules to sustain multilateral rules.*

The third possibility is related to the second. Recall that the discount factor relates to a particular period of time: the time between “transactions”. In other words, if transactions occur every three months then δ^* is a three-month discount factor, whereas if

transactions occur every month then δ^* is a one-month discount factor. To see what this means, consider once again the multilateral game with 4 states and assume that this game is played every quarter. Then the critical quarterly probability adjusted discount factor above which cooperation is sustainable is $5/6$, or 83%. Now suppose that the game is played every month. The critical monthly probability adjusted discount factor to sustain cooperation is again $5/6$ or 83%, but this is equivalent to a quarterly probability adjusted discount factor of $(5/6)^3$, or 58%. In other words, multilateral rules are more easily sustained when they involve frequent interactions between the member states.

These two possibilities can, of course, interact. States may be able to enter into multiple rules, some with frequent interaction and some with infrequent interaction. By the same argument as above, *a state can use slack enforcement power from rules with frequent transactions to sustain rules with infrequent transactions.*¹⁴⁷

The implication of this analysis is that there exists a significant set of cases in which it will be possible to form multilateral customary rules of international law. The likelihood of formation in any particular circumstance will depend on a number of factors, including (i) the relative value of cooperation versus defection, (ii) the number of states effectively involved, (iii) the extent to which increasing the number of states involved increases the value of cooperation or the detriments of defection, including whether the particular issue has characteristics of a commons problem, a public good, or a network, (iv) the information available to the states involved regarding compliance and defection, (v) the relative patience of states to realize benefits of long-term cooperation compared to short-term defection, (vi) the expected duration of interaction, (vii) the frequency of interaction, and (viii) whether there are also bilateral relationships or other multilateral relationships between the involved states.

¹⁴⁷ This is the case analyzed by Lohmann, *supra* note 144.