

2004

Self-Enforcing International Agreements and the Limits of Coercion

Robert E. Scott
Columbia Law School, rscott@law.columbia.edu

Paul B. Stephan
PSTEPHAN@LAW5.LAW.VIRGINIA.EDU

Follow this and additional works at: https://scholarship.law.columbia.edu/faculty_scholarship



Part of the [Contracts Commons](#), and the [International Law Commons](#)

Recommended Citation

Robert E. Scott & Paul B. Stephan, *Self-Enforcing International Agreements and the Limits of Coercion*, WISCONSIN LAW REVIEW, VOL. 2004, P. 551, 2004; COLUMBIA LAW & ECONOMICS WORKING PAPER No. 246 (2004). Available at: https://scholarship.law.columbia.edu/faculty_scholarship/1315

This Working Paper is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact cls2184@columbia.edu.

Columbia Law School
The Center for Law and Economic Studies
435 West 116th St.
New York, NY 10027-7201

Working Paper No. 246

Self-Enforcing International Agreements and
the Limits of Coercion

Robert E. Scott and Paul B. Stephan, III

March 2004

*This paper can be downloaded without charge from the
Social Science Research Network electronic library at:*

<http://ssrn.com/abstract=511362>

*An index to the working papers in the Columbia Law
School Working Paper Series is located at:*

http://www.law.columbia.edu/center_program/law_economics

SELF-ENFORCING INTERNATIONAL AGREEMENTS AND THE LIMITS OF COERCION

*Robert E. Scott**
*Paul B. Stephan***

International law provides an ideal context for studying the effects of freedom from coercion on cooperative behavior. To be sure, almost all academic discussions of the subject begin by asking whether international law constitutes “law.” We will duck this problem by observing that the category of all “international law” is too big and heterogenous to permit useful analysis. Whether to regard, say, the rules governing the conduct of war or international humanitarian law as “law”¹ presents radically different issues than does analyzing the legal character of the Treaty of Rome (the constitutive instrument of the European Community)² or the Warsaw Convention (the instrument governing contracts for the carriage of goods in international air transit).³ Instead, we will focus on a subset of international law, namely enforcement mechanisms for treaties and other agreements among states, and in particular agreements that involve the joint production of social welfare.⁴

By limiting our inquiry to welfare-enhancing international agreements, we necessarily exclude customary international law. The existence and significance of this body of norms and obligations is a hardy perennial among specialists.⁵ But the lack of any clear consensus as to what customary

* David and Mary Harrison Distinguished Professor of Law, University of Virginia School of Law.

** Lewis F. Powell, Jr. Professor of Law, University of Virginia School of Law. We are grateful to Pamela Clark, Allen Lynch, Eric Posner, Elizabeth Scott, John Setear, and participants in a Virginia Legal Theory workshop, the Conference on Freedom from Contract, and a University of Chicago International Law workshop for helpful comments, and to Xinh Luu for invaluable research assistance.

¹ See James D. Morrow, *The Laws of War, Common Conjectures, and Legal Systems in International Politics*, 31 J. LEG. STUD. S41 (2002); James D. Morrow, *The Institutional Features of the Laws of War*, 55 INT’L. ORG. 971 (2001); Oscar Schachter, *The Right of States to Use Armed Force*, 82 MICH. L. REV. 1620 (1984).

² Consolidated Version of the Treaty Establishing the European Community, O.J. (C325) 1 (Dec. 24, 2002).

³ Convention for the Unification of Certain Rules Relating to International Transportation by Air, Oct. 12, 1929, 49 Stat. 3000, T.S. No. 876.

⁴ We focus our attention on welfare-enhancing international agreements in order to examine systematically the enforcement questions that are at the heart of our inquiry. See notes 18-27 *infra* and accompanying text.

⁵ On the meaning and function of customary international law, see Anthony A. D’Amato, *THE CONCEPT OF CUSTOM IN INTERNATIONAL LAW* (1971); Michael Byers, *CUSTOM, POWER, AND THE POWER OF RULES: INTERNATIONAL RELATIONS AND CUSTOMARY INTERNATIONAL LAW* (1999); Karol Wolke, *CUSTOM IN PRESENT INTERNATIONAL LAW* (2nd rev. ed. 1993); Curtis

international law means or does, along with the extremely spotty practice of courts regarding its invocation, discourages us from attempting to explore its functional effects.⁶ Moreover, to the extent customary international law has any coherence, it seems to comprise an amalgam of tort and property rules, and we intend to explore the contributions that contract theory specifically can make to our understanding of the field. Accordingly, we will consider the enforcement only of express instruments entered into by states.⁷

Second, we distinguish among international agreements not according to their subjects or objects, but according to the enforcement mechanisms attached to them. This approach allows us to concentrate on the instrumental aspects of agreements, as distinguished from their moral or symbolic functions. We defend this focus by arguing that the existence of observable instrumental effects has distinct expressive value even to those who find only the moral or rhetorical aspects of law interesting.⁸

A. Bradley & Jack L. Goldsmith, *Customary International Law as Federal Common Law: A Critique of the Modern Position*, 119 HARV. L. REV. 815 (1997); Jonathan I. Charney, *Universal International Law*, 87 AM. J. INT'L L. 529 (1993); Jack L. Goldsmith & Eric A. Posner, *A Theory of Customary International Law*, 66 U. CHI. L. REV. 1113 (1999); Phillip R. Trimble, *A Revisionist View of Customary International Law*, 33 U.C.L.A. L. REV. 665 (1986).

⁶ Another reason for preferring agreements to custom is that the former reduce the “counterfactual problem” that arises when departures from custom occur, namely determining whether the departure constitutes a defection or adherence to a new custom. See Paul G. Mahoney & Chris William Sanchirico, *Norms, Repeated Games, and the Role of Law*, 91 CALIF. L. REV. 1281 (2003). On the role of the treaty formation process in addressing the counterfactual problem, see John K. Setear, *Law in the Service of Politics: Moving Neo-Liberal Institutionalism from Metaphor to Theory by Using the International Treaty Process to Define “Iteration”*, 37 VA. J. INT'L L. 641 (1997). For an extension of the point to noncoercive third-party dispute resolution based on international agreements, see Tom Ginsburg & Richard McAdams, *Coordination at the Core: An Expressive Theory of International Adjudication* (May 2003).

⁷ We acknowledge that the distinction between an international agreement that looks to the joint production of social welfare and other types of international law might not always be clear, inasmuch as the scope of some agreements may be more tacit than formal, and most instances of customary law are said to rest on consent. For the vast majority of cases, however, the distinction seems sufficiently clear to provide a useful basis for analysis. We focus on explicit state commitments, in contrast to obligations inferred from behavior regularities or expressions of conventional wisdom. We mean to include in our analysis agreements to which at least two states are parties, including agreements to which non-state actors, such as private firms or international organizations, also join. We do mean to exclude multilateral agreements that have a purely coordination as opposed to a productive function. We also will not consider contracts between a single state and private actors, although some aspects of our analysis may be applicable to such agreements. For a general discussion, see Daniel R. Fischel & Alan O. Sykes, *Governmental Liability for Breach of Contract*, 1 AM. L. & ECON. REV. 313 (1999).

⁸ Our focus on enforcement strategies distinguishes this paper from research on the question of whether the parties to an international agreement seek to create a binding obligation under international law. See Andrew T. Guzman, *The Design of International Agreements* (Jul. 2003); Jack L. Goldsmith & Eric A. Posner, *International Agreements: A Rational Choice Approach*, 44 VA. J. INT'L L. 113 (2003); Kenneth W. Abbot, Robert O. Keohane, Andrew Moravcsik, Anne-Marie Slaughter & Duncan Snidal, *The Concept of Legalization*, 54 INT'L ORG. 401 (2000); Kal Raustiala, *Form and Substance in International Agreements* (Draft 2004). Some agreements that do not create an obligation under international law nonetheless invoke self enforcement mechanisms, while finding that an agreement creates an obligation under international law does little to determine what enforcement mechanisms apply.

Our thesis is straightforward: Framers of international agreements, no less than the authors of private contracts, can choose between self enforcement and coercive third-party mechanisms to induce compliance with the commitments they make. Studies of individual contracting provide some evidence that coercive sanctions may crowd out self enforcement, implying that too great a propensity by external actors to intervene in the contractual relationship may produce welfare losses. We explore the possibility that too much coercive third-party enforcement similarly can reduce the value of international agreements.

A casual acquaintance with the field may leave the impression that international agreements can only be self-enforcing. No international sheriff works in the background to compel compliance and punish breaches of treaties.⁹ Agreements may fashion their own enforcement mechanisms, but these have no greater authority than the instrument that creates them. Consider a nonaggression treaty, such as that Hitler's Germany signed with Stalin's Soviet Union in 1939.¹⁰ Nominally, the Molotov-Ribbentrop Pact called for arbitration of all disputes between the parties, but the only sanction for violation was retaliation, a power that existed regardless of the Pact's provisions.¹¹ Without international law enforcement institutions, it would seem that all international agreements must stand or fall on the uncoerced good behavior of the parties.

A somewhat more sophisticated approach would find an analogy between international agreements and the many kinds of communally enforced contracts. Medieval trade fairs, contemporary cotton,

Just as scholars contend that choosing to make an agreement "legal" signals information about the parties' intent, Jack L. Goldsmith & Eric A. Posner, *supra*, at 132-34; Andrew T. Guzman, note 8, at 47-50, we maintain that the choice of enforcement mechanism reveals valuable information about the parties and their perception of the future.

⁹ See H.L.A. Hart, *THE CONCEPT OF LAW* 209 (1961):

It is indeed arguable, as we shall show, that international law not only lacks the secondary rules of change and adjudication which provide for legislature and courts, but also a unifying rule of recognition specifying "sources" of law and providing general criteria for the identification of its rules.

¹⁰ Treaty of Non-Aggression, Aug. 23, 1939, F.R.G.-U.S.S.R., 1939 RGBI. II, No. 38, translated in 7 DOCUMENTS ON GERMAN FOREIGN POLICY, 1918-1945, at 245 (Series D) (U.S. Dept. of State 1956); Secret Additional Protocol, Aug. 23, 1939, in 7 DOCUMENTS ON GERMAN FOREIGN POLICY, *supra*, at 246. For a more general discussion of Nazi Germany's use of the forms of international law, see Jack L. Goldsmith & Eric A. Posner, *Moral and Legal Rhetoric in International Relations: A Rational Choice Perspective*, 31 J. LEG. STUD. S115, S116-17 (2002).

¹¹ Article V of the Treaty contemplated "if necessary" the establishment of arbitration to resolve disputes between the parties, but the Soviet Union chose other means to respond to the German flouting of the agreement by its invasion in force on June 22, 1941.

diamond and rubber dealers, and lenders in U.S. immigrant communities, for example, have employed social sanctions based on communal solidarity as a substitute for third-party enforcement of commitments.¹² Similarly, some international lawyers argue, the “international community” induces compliance with agreements by maintaining a running assessment of the law-abiding quality of nation states and using social signals to deter violations of international commitments.¹³

A deeper examination reveals, however, that international agreements come with an array of enforcement mechanisms based on various institutional arrangements. Parties may choose to rely only on self-enforcement mechanisms, buttressed by social sanctions wielded by the international community, and most do so. But countries also can turn to external mechanisms backed by credible coercive authority in order to resolve their disputes. Agreements may invoke these mechanisms explicitly, or third-party enforcement bodies may decide on their own to enforce an agreement. International agreements, much like private contracts, necessitate choices about enforcement structure, and particularly the mix between external enforcement and self enforcement.

The architects of international law long have remarked on the superficial similarity between international agreements and private contracts.¹⁴ But functional analyses of these formally similar instruments traditionally has come from different directions. A substantial literature explores self enforcement of international agreements, for the most part assuming that no other means exist for inducing compliance.¹⁵ On the other hand, legal scholars of private contracts until recently have

¹² For studies of these phenomena, see Janet Tai Landa, TRUST, ETHNICITY, AND IDENTITY: BEYOND THE NEW INSTITUTIONAL ECONOMICS OF ETHNIC TRADING NETWORKS, CONTRACT LAW, AND GIFT EXCHANGE 112 (2001); Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 U. PA. L. REV. 1765 (1996); Lan Cao, *Looking at Communities and Markets*, 74 NOTRE DAME L. REV. 841 (1999); Avner Greif, *Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders*, 49 J. ECON. HIST. 857 (1989); Eric A. Posner, *The Regulation of Groups: The Influence of Legal and Nonlegal Sanctions on Collective Action*, 63 U. CHI. L. REV. 133 (1996).

¹³ For a synthesis of the argument that compliance with international obligations depend critically on the perceptions of other nations about compliance, see Andrew Guzman, *A Compliance-Based Theory of International Law*, 90 CALIF. L. REV. 1823 (2002).

¹⁴ See, e.g., *Foster v. Neilson*, 27 U.S. (2 Pet.) 253, 314 (1829) (Marshall, C.J.): “A treaty is in its nature a contract between two nations, not a legislative act.”

¹⁵ See, e.g., Kyle Bagwell & Robert W. Staiger, *Multilateral Tariff Cooperation During the Formation of Free Trade Areas*, 38 INT'L ECON. REV. 291 (1997); Robert Keohane, *Reciprocity in International Relations*, 40 INT'L ORG. 1 (1986); Giovanni Maggi, *The Role of Multilateral Institutions in Trade Cooperation*, 89 AMER. ECON. REV. 190 (1999); John K. Setear, *Responses to Breach of a Treaty and Rationalist International Relations Theory: The Rules of Release and Remediation in the*

focused mostly on the coercive aspects of enforcement, and have slighted both the role of self enforcement in explaining contract compliance and the interplay between self enforcement and coercion. It would seem that, however compelling is the contract metaphor in explaining bargaining behavior and expectations, these two categories of agreements occupy different universes in terms of their functions and behavioral effects.

But in the past few years, scholars have begun to explore how coercive enforcement of private contracts affects self enforcement.¹⁶ Although these inquiries are too early and tentative to support firm conclusions, a growing body of evidence indicates that, as to individuals, the relationship between self enforcement and coercion is complex and, in at least some areas, rivalrous. The individual characteristics that facilitate valuable self enforcement may be suppressed once one introduces coercive sanctions.

It is not too great a leap to ask whether the insights into individual decisionmaking and incentives suggested by these studies of private contracts might also inform our understanding of state-to-state behavior. We argue that, in spite of the obvious differences between state and individual decisionmaking, enough similarities exist to make the inquiry worthwhile. Using analytic moves worked out in the context of private contracts, we make two general claims about international agreements, one conventional and one controversial. First, we maintain that one usefully can evaluate efforts to frame and implement international agreements in terms of optimal enforcement structure. Choosing from a broad range of normative criteria, one still can distinguish between better and worse enforcement strategies. Second, we argue that the optimal enforcement structure for any particular international agreement will depend on both the goals of the agreement and the context in which it designed and implemented. Because these goals and contexts are diverse, the set of optimal enforcement structures is heterogenous. Some optimal enforcement structures will depend largely on self enforcement, while others will not.

Law of Treaties and the Law of State Responsibility, 83 VA. L. REV. 1 (1997); Paola Conconi & Carlo Perroni, Self-Enforcing International Agreements and Domestic Political Credibility (CESifo Working Paper No. 988) (Jul. 2003).

¹⁶ See Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, 103 COLUM. L. REV. 1641 (2003); Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, 113 YALE L. J. 541 (2003); Robert E. Scott, *The Rise and Fall of Article 2*, 62 LA. L. REV. 1009 (2002).

Central to our claim is an appreciation of the interaction of self enforcement and third-party coercion including binding arbitration, use of international courts, and enforcement by domestic actors. We maintain that in a far from trivial number of instances subject to international agreement, self enforcement and coercive enforcement may be rivalrous and the optimal enforcement structure would preclude or limit coercive enforcement. In particular, we argue that good theoretical arguments buttress the general tendency of domestic courts not to extend their coercive powers to implement an international agreement without a clear signal from the framers of the agreement that this coercion is desired.¹⁷

The article proceeds in three parts. In Part I we develop an informal model that draws on contract theory to specify the key structural determinants of the formation and enforcement of international agreements. The model introduces the theoretical underpinnings of self-enforcing agreements and proposes a framework for optimal enforcement design under idealized conditions based on rational reciprocity. Part II extends this model by applying it to international agreements under more realistic conditions. It describes the array of external and self-enforcement mechanisms available to states affected by an international agreement and applies the rational-reciprocity framework to the different enforcement strategies chosen by states. Part III examines the relationship between self-enforcing and coercive methods of inducing agreement compliance. We examine the conditions under which self-enforcement and coercion may be rivals rather than complements and explore the implications of rivalry for the design of optimal enforcement structures.

We draw a weak and a strong conclusion. We tentatively conclude that rational preferences for reciprocity expand the domain of self-enforcing international agreements and provide a plausible

¹⁷ We recognize that this tendency has prominent exceptions and that a significant scholarly debate exists over the role of domestic courts in enforcing international law. Some international law scholars argue that judicial uninvolvement, absent a legislative signal, is required as a matter of constitutional structure and policy. See Curtis A. Bradley, *International Delegations, The Structural Constitution, and Non-Self-Execution*, 55 STAN. L. REV. 1557 (2003); John C. Yoo, *Laws as Treaties? The Constitutionality of Congressional-Executive Agreements*, 99 MICH. L. REV. 757 (2001). We also note that the same issue manifests itself in the form of statutory interpretation, in particular whether jurisdictional statutes should be seen as authorizing judicial creation of enforcement mechanisms for certain treaties. The latter issue currently is before the Supreme Court. *Alvarez-Machain v. United States*, 331 F.3rd 664 (9th Cir.), cert. granted sub nom. *Sosa v. Alvarez-Machain*, 124 S. Ct. 807 (2003). We profess agnosticism as to the constitutional and statutory debates. Our argument, rather, is that a default of judicial nonenforcement has desirable welfare effects. We leave it to others to explore whether constitutional or statutory interpretation should incorporate welfare consequences.

explanation for the reluctance of states to adopt available mechanisms for coercive enforcement. By narrowing the domain of coercive enforcement, states can preserve space for parties to exploit opportunities to reciprocate. Our strong conclusion is that the rational-reciprocity approach offers rich prospects for further research that will enhance our understanding of international law.

I. A RATIONAL-RECIPROCITY APPROACH TO INTERNATIONAL AGREEMENTS

In this Part, we first ask the threshold question of why states choose to enter agreements with other states that constrain their future behavior rather than rely on the present exchange of entitlements or other valuable property rights. We conclude that it is the contractual character of these agreements that constitute their appeal. As with private contracts, an international agreement has an intertemporal aspect: Parties agree today to do something tomorrow. But why do states enter into such commitments? After all, international relations often are conducted without them. States, just as private parties, often make simultaneous exchanges of entitlements for corresponding concessions, rather than exchange promises for the later trade of these entitlements. Recall the exchanges of spies that periodically took place at Checkpoint Charley during the Cold War. As long as simultaneity applies, a respect for each state's entitlements is sufficient to encourage commerce between the states, because a state will only part with an entitlement if it values more highly what is offered in exchange. Contract enforcement would be unnecessary to support welfare-enhancing trades (of any sort) between nations.

But in contrast to simultaneous exchanges, a contract is a set of promises regarding future behavior. Such promises are costly to make and to memorialize. To understand the role of enforcement in relation to contracting behavior, one must explain why enforcement will induce states to incur these costs.

A. Why Write International Agreements?

To answer the question of why states choose to enter enforceable agreements that constrain their future actions, we begin with several simplifying assumptions (that we then relax in Part II). Assume a world of nation states, each governed by a class of rational elites that seek to maximize the welfare

of their citizenry.¹⁸ Further assume that states have assets that they can trade and invest, including “property” inherent to sovereign states such as regulatory jurisdiction, military power and revenue-collection capabilities.¹⁹ Finally, assume that some “trades” of this “property” may enhance welfare – put simply, assume that complete autarchy of all nations is not the most desirable end state.

As with private contracting parties, our hypothetical nation states face the canonical contracting problem of ensuring efficient *ex post* trade and efficient *ex ante* investment.²⁰ We can assume that each state will ensure efficient trade without the need for enforceable agreements. For example, if State A agreed to trade ten units of an economic good to State B, but it turns out that trading twenty units would maximize joint gains, then the parties can modify the agreement to provide for delivery of the larger quantity. But ensuring efficient investment is more difficult. The investments we have in mind would include the production of specialized goods or services, the development of human capital specific to a particular relationship, and research to acquire information about future economic or political conditions. Imagine, for example, an arms control treaty that requires a state to pass up the design and implementation of a particular weapons system, or a trade regime that makes the development of a particular product (say, a type of aircraft) valuable as long as other states will allow their firms to

¹⁸ State welfare is obviously a mushy term, but we use it here as synonymous with shared conceptions of “the national interest.” The key element of this move is that we assume the absence of agency costs. We relax this assumption at notes 73-77 *infra* and accompanying text.

¹⁹ For exploration of the analogy of sovereign regulatory jurisdiction to a private property right, see Joel P. Trachtman, *An Economic Analysis of Prescriptive Jurisdiction*, 42 VA. J. INT’LL. 1 (2001); Joel P. Trachtman, *Trade in Financial Services under GATS, NAFTA and the EC: A Regulatory Jurisdiction Analysis*, 34 COLUM. J. TRANSNAT’LL. 37 (1995). We mean “revenue-collection capability” to encompass not just the traditional subject of tax treaties, but also the imposition of duties, quotas and analogous constraints on international trade.

²⁰ Implicit in our analysis is a set of assumptions more often associated with what international relations specialists would call rationalist institutionalism. To some extent this puts us at odds with those (self-described) realists that see state insecurity, the drive for relative rather than absolute advantage, and exercises of coercive power as the central mechanisms in international relations. *See, e.g.*, John J. Mearsheimer, *The False Promise of International Institutions*, 19 INT’LSECURITY 5 (1994-95). Similarly, because we believe that there exists objective factors (the possibility of gains derived from cooperation, information asymmetries, *etc.*) that explain important aspects of international relations, we slight the insights of those who would emphasize the role of ideological predispositions and culture in conditioning international relations. *See, e.g.*, Alexander Wendt, *SOCIAL THEORY OF INTERNATIONAL POLITICS* (1999). Finally, because we emphasize joint production that presents collective actions problems, we neglect other forms of international cooperation and uses of international law, in particular coordination problems that do not contain opportunism problems. *See, e.g.*, Duncan Snidal, *Coordination versus Prisoners’ Dilemma: Implications for International Cooperation and Regimes*, 79 AM. POL. SCIENCE REV. 923 (1985). We will not seek to defend our choices here, other than note that some work in international relations theory seems more congenial to mainstream contract theory than does others.

purchase the good when produced. What defines these investments is their relation-specific character, in the sense that they have little or no value if put to an alternative use.

We propose a simple example to show how enforcement is essential to ensuring these value-enhancing investments.²¹ Each state in our story can either produce generic economic goods that are useful to many other states or produce specialized goods that are specific to a particular relationship with another state. Assume that State A can produce a generic version of a particular economic good and sell it on the world market at a price that equals cost (including a return on the selling state's investment). Imagine, for example, that this good consists of a particular exercise of sovereign power, such as a decision to regulate through competition rules all transactions anywhere in the world that have an observable economic effect on the state's economy, that other states "purchase" this product by adopting their own rule of regulatory jurisdiction in light of State A's choice, and that the costs and benefits of this product derive from the effects of all states' regulation on global welfare.²² State A also can produce a specialized version of the good for State B. Imagine, for example, that the specialized "good" consists of allowing State B's regulatory decisions to displace State A's for a specified range of transactions, such as the application of competition rules to producers operating in State B.²³ Assume that the cost of the generic good is \$1000 and purchasing states value it at \$1,500. The cost of the investment to produce the specialized good is \$2,000 and State B values it at \$3,000.

²¹ The example that follows draws on the discussion in Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, note 16 *supra*, at 559-562.

²² The United States generally applies the "effects" test in the application of its antitrust laws. *Hartford Fire Ins. Co. v. California*, 509 U.S. 764 (1993). For some of the finer points involving the extent of this jurisdiction, see *Empagran S.A. v. F. Hoffman-LaRoche, Ltd.*, 315 F.3d 338 (D.C. Cir.), *cert. granted*, 124 S. Ct. 966 (2003). For purposes of analysis, we assume that the regulation in question increases aggregate welfare, and thus does not involve significant economic rents. For debate about these assumptions and extension of the "regulation as property" analysis, see Andrew T. Guzman, *Choice of Law: New Foundations*, 90 GEO. L. J. 883 (2002); Erin Ann O'Hara, *Economics, Public Choice, and the Perennial Conflict of Laws*, 90 GEO. L. J. 941 (2002); Paul B. Stephan, *The Political Economy of Choice of Law*, 90 GEO. L. J. 957 (2002); Andrew T. Guzman, *Public Choice and International Regulatory Competition*, 90 GEO. L.J. 971 (2002). If these assumptions seem unrealistic, one could conceive a different example, say a more conventional economic good such as advanced weapons systems produced by a public-private collaboration.

²³ *Cf. Rewe-Zentral AG v. Bundesmonopolverwaltung für Branntwein (Cassis de Dijon)* (Case 120/78), [1979] E.C.R. 649 (explaining obligation under Treaty of Rome for one state to defer to regulatory decisions of other states). For further analysis of regulatory deference in the European Community, see Francesca Bignami, *The Challenge of Cooperative Regulatory Regulations after Enlargement* in *LAW AND GOVERNANCE IN AN ENLARGED EUROPE* (George A. Bermann & Katharina Pistor eds. 2004). For discussion of specific proposals to allocate internationally the authority to exercise competition policy jurisdiction, see Paul B. Stephan, *Competitive Competition Law? An Essay Against International Cooperation*, in *ANTITRUST POLICY: COMPETITION AND COOPERATION* (Richard A. Epstein & Michael S. Greve eds. 2004).

But State A's investment to make the specialized good cannot be redeployed; thus, State A will lose its entire investment if the agreement breaks down. For example, imagine that State A would commit to follow State B's competition rules and that if State B then exercised its regulatory authority to encourage the formation of monopolies designed to exploit consumers in State A, State A could not quickly enact new legislation reclaiming its regulatory power because of its constitutional arrangements governing lawmaking.

Under the assumptions above, A and B would prefer to produce the specialized good when that would maximize the contractual surplus. Using our hypothetical values, therefore, the parties would agree to produce the specialized good: This agreement would generate a surplus of \$1,000, while the alternative would generate only a surplus of \$500. Assuming that the two states are equally patient bargainers, the parties under certain plausible assumptions about bargaining behavior would divide the \$1,000 surplus equally with a contract price of \$2,500.²⁴

Assume initially that this agreement is not enforceable. In this world, the price at which the parties ultimately will transact would not be \$2,500, because State B's incentive to cooperate disappears after State A invests \$2,000 in the deal. After State A has made its investment, B has an incentive to demand renegotiation of the price. It could, for example, insist that State A make trade or regulatory concessions, such as agreeing to purchase a weapons system produced in State B, as compensation for B's discouragement of monopolies targeted at consumers in State A. At that point, State A would have sunk the investment cost, which the renegotiated bargain therefore would ignore. The only question for State A at this point is whether to trade the specialized good at some price or not to trade at all. Because trade would produce a gross gain of \$3,000, while the decision not to trade produces no value, the parties would proceed with the renegotiated transaction, dividing the \$3,000 gain equally. Under a renegotiated price of \$1,500, of course, State A would lose \$500 (\$1,500 less its

²⁴ We suppose that States A and B are equally patient, because they have similar costs of capital. Thus, assuming the parties engage in "deal me out" bargaining, the parties would share the surplus equally unless one party's next best option exceeds half of an equal split. For discussion of the possible bargaining game and why an even split is plausible, see Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, note 16 *supra*, at 552-54.

investment cost of \$2,000). Because State A would lose its entire investment if the parties failed to trade, it would (reluctantly) agree to the new price.²⁵

The point of this example, of course, is that if the agreement were unenforceable, State A would refuse at the outset to produce the specialized good even though the relation-specific investment would maximize expected surplus. State A would anticipate the hold-up potential inherent in the investment situation and accordingly would elect to produce generic goods instead. This outcome reduces the joint welfare of the parties, because the generic alternative generates a social surplus of \$500 while the specialized investment would have generated a surplus of \$1000. But if State B's promise to pay the \$2,500 contract price were enforceable, the parties could cooperate in producing the value-enhancing investment. State A would then anticipate being compensated for its investment, and B would always prefer the specialized investment with its \$500 payoff rather than the general alternative which only generates a payoff of \$250.

What conclusions can we draw from this example? First, the conventional view is that contract enforcement protects injured promisees, such as State A, who otherwise will lose their reliance interest in the prospective deal. But this view misses the main point. If international agreements were not enforceable, State A would not risk its reliance in the first place. Instead, as in our example, it would elect to produce generic economic goods rather than subject itself to exploitation. The key insight is that enforcement benefits *promisors*; it enables them to make credible promises to perform. State B, in our example, wishes that it could make an enforceable promise to pay State A the \$2,500 contract price. In sum, enforcement enables states to make credible promises to each other to secure relation-specific investments that will enhance the contractual surplus.

²⁵ We assume that the parties ignore State A's sunk \$2,000 investment cost when renegotiating the contract. There is experimental evidence that individuals will sometimes take sunk costs into account. Thus, an investing party's payoff in a bargain will increase if the other party knows that the investor has spent money to prepare. See Colin Camerer, BEHAVIORAL GAME THEORY 85-90 (2003); Lorne Carmichael & W. Bentley MacLeod, *Caring About Sunk Costs: A Behavioral Solution to Holdup Problems with Small Stakes*, 19 J.L. ECON. & ORG. 106 (2003). As we discuss below, it is an open question whether the fairness concerns regarding sunk costs that individual experimental subjects act upon also motivate states.

Second, this example illustrates a primary motivation for states to write enforceable international agreements: Enforcement is essential to ensuring welfare-enhancing investments that are specific to bilateral or multi-state relationships. Our story suggests, *inter alia*, why private parties have made so little foreign direct investment (FDI) in most of the former Soviet states, relative to the levels seen in many third-world countries.²⁶ Most FDI is relation specific (*e.g.*, building a factory far from the home country, developing a mine or an oil field). Potential investors will not deal unless the host country or local firm can make credible promises to adhere to the terms originally agreed upon, rather than renegotiating those terms after investments had been made. Although many of the former Soviet states have signed international agreements promising to protect foreign investment, investors have reason to regard these promises as insufficiently credible.²⁷

B. Enforcement and Optimal Contract Design

If enforcement enhances welfare, how should countries pursue it? What are the enforcement choices available to states seeking to encourage relation-specific investments? A range of possibilities exist. The options run from self-enforcement mechanisms to various forms of third-party coercion, including a commitment to arbitration, the embedding of enforcement in valuable multilateral organizations and direct enforcement in the promisor's domestic courts. For present purposes, however, we need to distinguish only between self-enforcing agreements and third-party, coercive enforcement.

1. Self-Enforcement

²⁶ See George T. Abed & Hamid R. Davoodi, *Corruption, Structural Reform, and Economic Performance in the Transition Economies*, IMF Working Paper (Jul. 2000).

²⁷ We discuss these agreements, some of which provide for a form of coercive enforcement, at notes 119-122 *infra* and accompanying text. The post-Soviet experience suggests that the existence of some coercion *simpliciter* is not as significant as the degree of coercive third-party enforcement. The degree of coercive enforcement that these agreements establish seems to have been insufficient to meet the concerns of foreign investors. See also authorities cited at note 47 *infra*.

Self-enforcing agreements provide credibility to commitments where parties contemplating noncompliance face effective sanctions that do not depend on the actions of third-party enforcers.²⁸ Retaliatory threats and reputation undergird the traditional analysis of this process. Consider first a long-term relationship between a supplier and a producer in which both firms make investments in the contract and face significant costs in switching to different partners. Both parties appreciate their own and their counterparty's vulnerability to opportunism. Both have a credible capacity to retaliate against opportunism based on their capacity to withhold the expected benefits from future dealings.

A significant constraint on the efficacy of retaliatory threats, however, is the extent of the parties' expectation of future benefits. As the term of an agreement draws to a close, the parties have fewer future benefits that retaliation can deny them. This end-game problem plagues all relationships that have definite concluding points. End games aside, exogenous changes that lead one or more parties to devalue what it would gain from the agreement also diminishes the effectiveness of retaliatory threats.

As an alternative to contractor-specific relations, agreements may invoke community-based sanctions based on reputational losses. This mechanism involves transactors who belong to a community that invests in the monitoring of transactor behavior, disseminates information generated by monitoring to community members at a reasonable cost, and whose members retaliate against noncompliers. Reputation is an effective means of self-enforcement whenever a transactor values its reputation in this community, such as where migration to other communities entails significant costs, and community members make the effort to impose sanctions on parties who break their agreements.²⁹

²⁸ There is an extensive literature on self-enforcing contracts. *See, e.g.*, Benjamin Klein & Keith B. Leffler, *The Role of Market Forces in Assuring Contractual Performance*, 89 J. POL. ECON. 615, 617 (1981); Lester G. Telser, *A Theory of Self-Enforcing Agreements*, 53 J. BUS. 27 (1980); Oliver E. Williamson, *Assessing Contract*, 1 J. L. ECON. & ORG. 177, 201-02 (1985); Robert E. Scott, *Conflict and Cooperation in Long-Term Contracts*, 75 CALIF. L. REV. 2005, 2031-34, 2039-44 (1987). For antecedent work with different methodological commitments that provides rich insights into self enforcement, see Stewart Macauley, *Non-Contractual Relations in Business: A Preliminary Study*, 28 AM. SOC. REV. 55 (1963); Stewart Macauley, *An Empirical View of Contract*, 1985 WISC. L. REV. 465.

²⁹ Reputation is most effective as a means of enforcement in homogeneous communities, where the behavior of individual parties is a matter of general knowledge. *See* Janet Landa, *A Theory of the Ethnically Homogenous Middleman Group: An Institutional Alternative to Contract Law*, 10 J. Legal Stud. 349 (1981); Avner Grief, *Informal Contract Enforcement: Lessons from Medieval Trade* in 2 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND LAW 287 (Peter Newman, ed. 1998). Lisa Bernstein has documented as well the ability of trade associations to disseminate information about individual misbehavior and to enforce

Like repeat play, there are constraints on the use of reputation as a self-enforcement mechanism. Reputation cannot operate effectively if the parties are strangers to each other, when a party intends to withdraw from the community (as distinct from ending the contract-specific relationship), or when the benefits from opportunism exceed reputational costs.³⁰ Thus the range of contexts in which reputation can enhance promisor credibility, although significant, is limited.

Contract theorists recently have augmented their study of self enforcement by considering the tendency of people to value reciprocal fairness.³¹ New experimental work suggests that, in certain contexts, reciprocity is a potent additional means of self-enforcement.³² These studies have produced three key findings: (1) Many people deviate from purely self-interested behavior in a reciprocal manner. Reciprocity means that in response to friendly actions, many people are much more cooperative than predicted by the axioms of rational choice. Conversely, in response to hostile actions many people are much more nasty and vengeful;³³ (2) People repay gifts and take revenge even in one shot interactions with complete strangers and even if such action is costly for them and yields neither present nor future material rewards;³⁴ and (3) This is a heterogeneous world. Some people exhibit reciprocal fairness and others are selfish. Taking all the experiments together from countries as diverse as Austria, Indonesia, Russia and the U.S., the fraction of fair subjects ranges from 40 to 60% as does the fraction of subjects who are selfish.³⁵

boycotts and other collective sanctions. The contracting behavior of the members of the association thus becomes part of the collective memory of the trade group. See Lisa Bernstein, *Private Commercial Law in the Cotton Industry: Creating Cooperation through Rules, Norms and Institutions*, 99 MICH. L. REV. 1724 (2001); Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 U. PA. L. REV. 1765 (1996).

³⁰ See Andrew T. Guzman, note 13 *supra* (discussing reputational effects and compliance with international law).

³¹ For fuller discussion, see Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1663-65.

³² See, e.g., Matthew Rabin, *Incorporating Fairness into Game Theory and Economics*, 83 AM. ECON. REV. 1281 (1993); David K. Levine, *Modeling Altruism and Spitefulness in Experiments*, 1 REV. ECON. DYNAM. 593 (1998); Ernst Fehr & Klaus Schmidt, *A Theory of Fairness, Competition and Cooperation*, 114 Q. J. ECON. 817 (1999); Ernst Fehr, Simon Gächter, & Georg Kirchsteiger, *Reciprocity as a Contract Enforcement Device: Experimental Evidence*, 65 ECONOMETRICA 833 (1997). For a review of the literature, see Ernst Fehr & Klaus Schmidt, *Theories of Fairness and Reciprocity – Evidence and Economic Applications*, University of Zurich, Institute for Empirical Research in Economics, Working Paper No. 75, at 2-3 (2001).

³³ Ernst Fehr & Simon Gächter, *Fairness and Retaliation: The Economics of Reciprocity*, University of Zurich Institute for Empirical Research in Economics, Working Paper No. 40, at 1 (2000).

³⁴ *Id.*

³⁵ Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1665. This finding of heterogeneity provides a convincing explanation for the apparent anomaly of the robust evidence of reciprocal fairness in bilateral interactions and the equally robust evidence from experiments in competitive markets where almost all subjects behave as if they were self-interested.

This evidence suggests that third-party coercive enforcement may be unnecessary to enhance the credibility of promises even in circumstances where retaliatory threats and reputation effects might not work. As long as the fraction of reciprocally fair individuals in the relevant population is consistent with the experimental evidence, even strangers and one-time transactors will make relation-specific investments in reliance on the promise of the other to pay.³⁶ We thus must consider the existence and scope of the preference for reciprocal fairness.

Notwithstanding the predictive power of this preference in experimental settings, the theory has yet to be seriously tested in real world contexts. Thus, any use of the fairness concept raises the question of external validity. One critique of the experimental evidence is particularly relevant to the application of reciprocal fairness theory to explain international treaties. All of the experimental subjects are untrained individuals and not professional bureaucrats or politicians. Thus, it is unclear to what extent the observed behaviors, even if they apply to the general population, also apply to the elites that affect compliance with treaty obligations. Individuals in laboratory experiments may respond differently than state officers, for example, because the experimental subjects do not face the same pressures to make maximizing decisions.

We also note that the experimental evidence does not establish that the observed preferences for reciprocity are a deeply entrenched, intrinsic motivation (and thus inconsistent with the assumption of rational self-interest). Rather, the observed preference for reciprocity may be either a learned or a normative behavior. It would be hardly surprising if individuals learn to devise strategies – or

The central insight is that the observed behavior is a function of the economic environment. In bilateral experiments, the presence of a fraction of reciprocally fair individuals can create incentives for selfish types to make fair offers. Alternatively, in a competitive market a few selfish players can drive the price to the competitive level and no single fair person can effect that price. Ernst Fehr & Klaus Schmidt, *Theories of Fairness and Reciprocity*, note 32 *supra*, at 38-40.

³⁶ The reciprocal fairness experiments show that it pays to write trust contracts – that is, agreements based on observable but not verifiable conduct – even where the promisee is uncertain whether the promisor is a fair or selfish type. So long as the population is heterogeneous (*i.e.*, there is a significant fraction of fair types in the population), then even where the transactors are anonymous, reciprocity yields better enforcement outcomes in experimental settings *on average* than does the alternative of coercive enforcement. This is because even selfish parties will respond reciprocally to an offer to enter into a trust contract because of the positive probability that the promisor will behave fairly. One reason that self-enforcement is the better strategy on average is that reciprocity is significantly less costly than third party enforcement. Another reason is that parties to self-enforcing agreements can condition performance on observable factors that might not be verifiable to a third party, thus enhancing joint welfare. For discussion, see Robert E. Scott, *Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1682-85.

heuristics – that *do* work in real-world transactions that generally present a possibility of repeat play and reputational effect, and then fail to adjust those strategies to the pure single iteration game in the laboratory.³⁷

Moreover, there is evidence that cultures do generate norms of reciprocity. These norms are consistent with individual self-interest because, *over time*, parties will be better off if they behave fairly. Following the “over time” heuristic consistently, not making distinctions for what appear to be one-shot interactions with strangers, may be a successful, maximizing strategy.³⁸ After all, sometimes one might mistake a repeat-play game for a single-iteration game and get punished or pay an unexpected reputational price. Thus, it is possible that a rational individual could behave in the way the experimental economists describe as reciprocally fair (and not utility maximizing) simply because the experiments do not take into account the costs of categorizing, and mis-categorizing, transactions.³⁹

These speculations do not challenge the experimental evidence so much as suggest a context for appreciating its significance. The existence of a preference for reciprocal fairness may not undermine the rational self-interest hypothesis, but rather extend its reach. Repeat play, reputation, and a preference for reciprocal fairness all explain why actors rationally may police their own behavior, obviating the need for external coercion. We do not argue that the potential of self enforcement is limitless, but rather that, working within the conventional economics framework of interested decisionmaking, self enforcement can operate to enhance promisor credibility. We next turn to considering more fully the relationship between repeat play and reputation, the more fully studied mechanisms of self enforcement, and the newer reciprocal fairness research. Together these

³⁷ On heuristics generally, see *BOUNDED RATIONALITY: THE ADAPTIVE TOOLBOX* (Gerd Gigerenzer & Robert Selten eds., 2000).

³⁸ Rational actors can profit from a precommitment strategy to guard against short-term deviations from behavior that would be inconsistent with their long term preferences. Thus, a rule such as: “respond reciprocally even in one shot interactions with strangers” may be a successful self-command strategy to guard against mistakes in mischaracterizing repeat play and isolated interactions. For discussion of the relevance of self-command to contract theory, see Robert E. Scott, *Error and Rationality in Individual Decisionmaking: An Analysis of the Relationship Between Cognitive Error and the Management of Choice*, 59 S. CALL. REV. 329 (1986).

³⁹ Robert E. Scott, *Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1673-74.

mechanisms provide the basis for a theoretical approach to agreements based on what we will call rational reciprocity.

2. Rational Reciprocity as Adding Value to Self Enforcement

We discussed above areas where repeat play and reputational effects would not motivate rational actors to honor their obligations. More general problems attend these mechanisms even within their effective domains. Parties who rely on either to enforce their agreements face fundamental difficulties both in detecting a failure of performance and in responding proportionately once nonperformance is observed. The success of self-enforcement depends significantly on the clarity and predictability of the threatened responses to non-performance. Selecting an appropriate response to, say, an instance of shirking becomes more complicated when the other party's behavior cannot readily be understood. Parties rarely shirk by directly announcing their unwillingness to perform as promised. They typically affirm solidarity, protest helplessness in the face of intractable problems, or act in subtle ways that are difficult to evaluate. In other words, nonperformance is a "noisy" signal and systematic misperception of the other's actions may cause inappropriate responses.⁴⁰

Misunderstanding of an actor's behavior can result from many sources, including reliance on a small sample size.⁴¹ Moreover, even if participants can observe shirking behavior, third parties may be unable to detect it.⁴² Thus, reputation will work to make promissory commitments credible only if other parties can conveniently learn about the reasons why any particular transaction broke down. Consequently, a reputation for trustworthiness can be difficult to establish, especially in heterogeneous environments where most market participants are unfamiliar with any particular contracting party. Without moral clarity, the mere fact of breakdown is not sufficient to impose a reputational cost on either party.

⁴⁰ For a fuller discussion of these issues, see John K. Setear, note 15 *supra*, at 86-90.

⁴¹ Robert Jervis, PERCEPTION AND MISPERCEPTION IN INTERNATIONAL POLITICS 319-406 (1976); John K. Setear, note 15 *supra*, at 93-98.

⁴² See notes 49-51 *infra* and accompanying text.

Even if shirking can be detected, the imposition of self-enforcing sanctions must be carefully calibrated. Overdeterrence or excessive retaliation by the counterparty functions as the equivalent of a breach and presents serious moral hazard problems. In addition to the risk of a disproportionately harsh response, the parties also must consider the risk of underenforcement, which will leave opportunism inadequately deterred. High variance in the costs of performance thus poses a major threat to an agreement that relies on the self-enforcing discipline of repeated interactions. This variance in expected costs correlates to a party's temptation to shirk its obligations to perform under the agreement.

The preceding analysis underscores the fact that self-enforcement through the threat of a loss of reputation or repeat transactions is not costless. It requires substantial monitoring and a punitive sanction for non-performance. Even if the breaching party understands and "accepts" the punishment, retaliation imposes stress on any on-going relationship that may threaten its survival. All parties therefore have an interest in augmenting their relationship within an embedded framework based on reciprocity. The existence of such a norm makes enforcement, when necessary, easier to administer and accept. For example and as we discuss in greater detail below, states can use trust agreements based on reciprocity to encourage each other to behave cooperatively long enough for them to discover a project's long-run benefits, thus bridging the gap between short-term and long-term payoffs.

In sum, reciprocal fairness offers a particularly stable foundation for a strategy of conditional cooperation. The strategy seems credible because it relies on behavioral responses that "go without saying." As a consequence, it ameliorates problems of detection and proportionality. A trust contract requires no monitoring other than a measured response to the observable actions of the other party.⁴³ When a selfish response is observed, a reciprocally fair type will retaliate appropriately.⁴⁴ The contracting parties can observe and assess each other's behavior within the context of the contract, rather than relying on general observations.

⁴³ For a discussion of trust contracts, which depends on behavior that parties can observe but not verify to a third party at a reasonable cost, see Robert E. Scott, *Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1681-82.

⁴⁴ Reciprocity, whether learned, normative or intrinsic is deeply embedded behavior. Thus, for example, the "eye for an eye and tooth for a tooth" formulation in the Hammurabi Code was intended to *restrict* revenge by requiring a measured, proportional response. See also the thirteenth century Norse epic verse, the EDDA: "A man ought to repay gift with gift and lies with treachery."

3. Coercive Third-Party Enforcement

The preceding discussion demonstrates both the possibilities of self enforcement to enhance the credibility of promises and the existence of inherent limitations to any self-enforcement regime, even conceding that reciprocity may extend its domain. As we have observed, reputation and the discipline of repeated interactions suffer from significant constraints. And while reciprocity may be an effective means of self-enforcement *on average*, it will also have a high variance. Because there is *both* self interest and reciprocity in the world, an agreement that relies on reciprocal fairness would be inefficient whenever a single act of non-performance might lead to serious disruptive effects. Moreover, where transactions are complex and the respective promises interrelated, a failure to perform may not be observable *at the time*, and thus reciprocity, which depends on a linkage between action and response, may not work to make the promises credible. At the time that parties exchange complex and interactive performances, they may have difficulty determining whether one party's refusal to respond cooperatively in a particular instance represents unfair or selfish behavior or is an appropriately measured, retaliatory response to an earlier instance of shirking by the other. These complex interactions are the sorts of agreements that private commercial parties typically reduce to legally enforceable obligations.⁴⁵

For several reasons, then, coercive enforcement by third parties is a desirable option in the optimal design of international agreements.⁴⁶ Self enforcement, to be sure, involves coercion, and it also may entail third parties who are empowered to determine the rights of the parties. Coercive third-party enforcement as we mean the term, however, requires both a disinterested arbiter and the arbiter's ability directly to impose sanctions on violators. This kind of arbiter can add value not only by unraveling the chain of causation that produced the breakdown, but also by investing in resources to punish a breach itself rather than leaving it to the parties to fashion a response. Moreover, the two

⁴⁵ Robert E. Scott, note 28 *supra*, at 2050-51.

⁴⁶ One of us has argued previously that third-party enforcement and self-enforcement regulate different aspects of the contractual relationship. On this view, third-party enforcement functions much as a nuclear umbrella, deterring breach in those states of the world where the payoffs from breach are substantial and exceed the range of self-enforcement. The other side of the argument is that, where the payoffs are relatively low, and reputation and repeated interactions are effective, self-enforcement is a more efficient "conventional" deterrent. *See id.* at 2044-48. *See also* Karen Eggleston, Eric A. Posner & Richard Zeckhauser, *The Design and Interpretation of Contracts: Why Complexity Matters*, 95 NW. U. L. REV. 91, 116 (2000).

function are complements: The ability to impose sanctions lends credibility to the unraveling. As we discuss below and contrary to conventional assumptions, states do have the option of designing external and coercive enforcement into their agreements.

The necessary conditions for external enforcement, then, are a disinterested third-party referee and a stakeholder who can impose costs or provide compensation for breach.⁴⁷ In the case of international agreements, the choice of external enforcement can range from a multilateral tribunal (think of the European Court of Justice) to third-party arbitration to access to domestic courts known for their independence and endowed with credible coercive authority.⁴⁸ Of course, the same institution can serve both functions. Because the value of enforcement derives from its capacity to make credible the promises of the respective parties, the imposition of costs can substitute for the awarding of compensation. Thus it is sufficient that the referee can impose proportionate sanctions on the breaching party, even if the referee cannot provide compensation for the breach once it occurs.

4. Designing Optimal Enforcement

Assume that our hypothetical states have a choice of enforcement options, ranging from self-enforcement to coercive enforcement. How do they design an enforcement regime for any particular agreement that maximizes the contractual surplus? The optimal means of enforcement in any given case will depend on the effects of asymmetric information on the relative costs of the enforcement options (including both direct costs and error costs).

a. The effects of asymmetric information. The existence of asymmetric information truncates the set of agreements that states can write. This point requires more careful unpacking. Consider an

⁴⁷ Honesty and independence are perhaps the most salient attributes of a disinterested referee. Its absence helps to explain the relatively low levels of foreign direct investment that private parties have made in countries with weak judicial systems, which prevents the local parties from making believable promises. See Douglass C. North, INSTITUTIONS, INSTITUTIONAL CHANGE, AND ECONOMIC PERFORMANCE (1990); Simeon Djankov, Rafael La Porta, Florencio Lopez-de-Silanes & Andrei Shleifer, *Courts*, 118 Q.J. ECON. 453 (2003); Stijn Claessens & Luc Laeven, *Law, Property Rights and Growth*, 58 J. FIN. 401 (2003); Rafael La Porta, Florencio Lopez-de-Silanes, Andrei Shleifer, & Robert W. Vishny, *Investor Protection and Corporate Valuation*, 57 J. FIN. 1147 (2002); Paul Mahoney, *The Common Law and Economic Growth: Hayek Might Be Right*, 30 J. LEG. STUD. 503 (2001).

⁴⁸ We discuss the details of these institutions in Part II. See notes 118-151 *infra* and accompanying text.

information taxonomy found in the contract theory literature. Information economics attaches analytic importance to the distinction between observable and verifiable information, but both concepts remain somewhat imprecise. According to standard theory, a datum of information is “unobservable” if the other contracting party cannot perceive it. A datum of information is “observable but not verifiable” if the other party can perceive it, but cannot prove the fact to a court or other third party at an acceptable cost. A datum of information thus is “verifiable” if a party both can observe it and efficiently prove its existence to a third party.

The field of international relations abounds with examples of both unobservable and unverifiable information.⁴⁹ Consider secret laboratory research on weapons systems or intelligence collection, which states work very hard to keep unobservable. At a less dramatic level, states rarely can determine another state’s costs in maintaining a particular tariff or exercising some range of regulatory authority, in part because of the difficulty of ascertaining the “policy set” of alternative regulatory choices. As for nonverifiability, consider a frequent problem in trade relations, namely an acute industry downturn that leads to “safeguard” protection against import competition. Victims of such barriers may be able to distinguish genuine claims of domestic injury from lax commitment to trade obligations, but third parties have great difficulty separating the two.⁵⁰ Alternatively, some evidence exists that human rights advocates concentrate their efforts (which largely involves the production of

⁴⁹ Goldsmith and Posner question the significance of information asymmetries in explaining international relations. They argue that: Most states these days are open, or do a poor job of keeping their secrets (and those insulated states, like North Korea, are for that reason assumed to be “bad types,” in an example of the classic unraveling result in games of asymmetric information), and one can obtain a fair indication of a state’s political stability by consulting the market’s valuation of its bonds.

Jack L. Goldsmith & Eric A. Posner, note 8 *supra*, at 137. Their argument focuses only on observability not verifiability. Market valuation of bonds exemplifies a process by which observers reliably may observe information, in the sense that they are willing to place bets on it in the bond market, but have great difficulty convincing a third party of the link between a particular event and a particular change in bond prices. See Kim Oosterlinck, *Why Do Investors Still Hope? The Soviet Repudiation Puzzle (1918-1919)* (Sep. 10, 2003) (documenting unverifiability of events affecting bond prices).

⁵⁰ For an instance illustrating the difficulty in ascertaining the strength of domestic injury justifies the use of rules that force a protecting state to signal its circumstances by absorbing higher costs, see Alan O. Sykes, *Protectionism as a “Safeguard”: A Positive Analysis of the GATT “Escape Clause” With Normative Speculations*, 58 U. CHI. L. REV. 255 (1991). For evidence that third parties have difficulty verifying the strength of domestic forces triggering safeguard measures, see, *e.g.*, United States – Safeguard Measures on Imports of Fresh, Chilled or Frozen Lamb Meat from New Zealand and Australia, (AB-2001-1) (reviewing legitimacy of U.S. safeguard measure under Uruguay Round Agreement on Safeguards and relying on second-order observations to justify determination).

reputational sanctions) on state acceptance of human rights obligations, rather than on state compliance, because of the difficulty of verifying the latter relative to the former.⁵¹

From the foregoing it follows that parties will contract for third-party enforcement only where the relevant measures of performance are verifiable. Otherwise, the uninformed party will be subject to risks of hidden action and hidden information. Take, for example, the hypothetical agreement between State A and State B for the production of a specialized good. Assume that State B is unable to observe whether State A has, in fact, invested \$2000 in production costs. Because of the risk of hidden action (or moral hazard), State B will decline to condition its promised payment on whether the promised investment was made. Instead, State B will write a contract that conditions its performance on the verifiable quality of the good produced by State A's investment. The problem for the parties, however, is that the quality required by State B may not be verifiable. If so, the parties must write a second best contract that conditions performance on a less relevant but verifiable quality.

An inability to observe a relevant condition will further limit the ability of states to obtain maximum value from potential transactions. In our example, State B might be willing to pay a premium for the highest quality specialized good if it knew that State A had the skill to produce it. But if B is unable to observe whether the producers in A are skilled or unskilled, B is vulnerable to the risk of hidden information (or adverse selection). Because B cannot observe A's type, B will pay only a blended price that reflects the probability that A's workers are unskillful. Unless states endowed with skilled workers can somehow communicate that information, they will be unable to compete with the unskilled workers in other states at the blended price.⁵² In short, if the value one party derives from performance depends on the characteristics or actions of the other party, then the inability to observe those characteristics or actions will render the performance in question not contractible.

⁵¹ See Oona Hathaway, *Do Human Rights Treaties Make a Difference?* 111 YALE L.J. 1935 (2002).

⁵² This is based on the assumption that it is more costly to train skilled workers than unskilled workers. See generally George Akerloff, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q. J. ECON. 355, 366 (1970).

b. Screening and signaling with preliminary agreements. In this paper, as in standard contract theory, we treat verifiability as an exogenous variable.⁵³ Observability, however, is an endogenous variable. Parties can overcome some information deficits by using screening or signaling mechanisms. For example, assume that State A and State B are contemplating a long-term investment agreement for a specialized good. Assume that the quality of the specialized good is observable to State B but cannot be efficiently verified to a third party. Under these circumstances, the parties have an incentive to write a self-enforcing agreement, because the specialized performance otherwise is not contractible. But assume that State A has no history with State B that provides the basis for a reputation for trustworthiness; nor does the prospect of future interactions discipline A's behavior. State A might propose a trust contract that relies on B's generosity to respond in kind once A invests, but A is unable to observe whether B is reciprocally fair or selfish. The assumption of heterogeneity implies, therefore, that State A risks losing a major investment by responding with enhanced efforts to a selfish party who subsequently fails to perform.

One solution is for the parties to design a screening mechanism so that State A can determine State B's preference for reciprocity. For example, State A and State B can enter into a preliminary agreement in which B promises to pay A for investing the necessary efforts to produce one unit of the specialized economic good and to pay A a bonus if the product proves satisfactory to B. This preliminary agreement, simple in form, offers a clearly-defined opportunity to reciprocate. It thus permits the parties to learn more about each other's taste for reciprocal fairness.⁵⁴ In this case, potential transactors are not only subject to observation, but they must spend considerable time

⁵³ The assumption that verifiability is exogenous is made for convenience, but, in fact, whether a measure of performance is verifiable or not is subject to some party control. *See, e.g.,* Franklin Allen & Douglas Gale, *Measurement Distortion and Missing Contingencies in Optimal Contracts*, 2 *ECON. THEORY* 1 (1992); Chris Sanchirico & George G. Triantis, *Evidence Arbitrage: The Fabrication of Evidence and the Verifiability of Contract Performance*, University of Virginia School of Law, Law and Economics Research Paper Series, No. 02-17 (Dec. 2002).

⁵⁴ Robert E. Scott, *Self-Enforcing Agreements*, note 16 *supra*, at 1683. The fact that State B is willing to enter an agreement that conditions on non-verifiable factors does not necessarily signal that it is reciprocally fair. A selfish state could copy the signal, as the invitation to reciprocate would induce greater efforts from State A and thus greater returns to the informed party. But the preliminary agreement itself creates several opportunities to reciprocate *in advance* of the formalization of a long-term relationship between the parties. First, a preliminary agreement gives State A the opportunity to acquire information about the character of State B by observing its behavior in response to opportunities to reciprocate. Second, the preliminary agreement separates in time the opportunity to reciprocate from the subsequent transaction that is ultimately contemplated. *Id.* The expenditure of time itself communicates a valuable signal of a preference for reciprocity. *See* A. Michael Spence, *Time and Communication in Economic and Social Interaction*, 87 *Q. J. ECON.* 651 (1973).

carrying out an agreement that is only self-enforcing. Because reciprocally fair individuals can capture the returns to general information about their type through an enhanced reputation for cooperation, they are more willing to spend resources to provide this information.⁵⁵

We do not wish to overstate the efficacy of preliminary agreements. If the cost of exhibiting cooperative behavior were less than the potential gains to be achieved subsequently from exploiting counterparty reliance, parties rationally would discount the value of the signal. Our point is simply that in some circumstances self enforcement can extend the scope of contractual space by compensating for the deficiencies of third-party enforcement.

c. The direct and error costs of enforcement. The distinction between observable and verifiable facts is key to understanding what motivates parties to choose between self-enforcement and third-party enforcement. Enforcement, as a conceptual matter, entails two categories of costs: (1) direct costs, comprising (a) investments in detecting and verifying breach and in imposing sanctions, and b) foregone opportunities from writing second best contracts that condition performance on less relevant but verifiable factors; and (2) error costs, entailing losses in welfare derived from both failures to recognize and sanction nonperformance (Type I errors) and incorrect determinations that breaches have occurred (Type II errors). Parties should seek to minimize the sum of these costs given any particular distribution of observable and verifiable facts relevant to the obligation to be enforced.

To illustrate how problems of private information influence the tradeoff between the direct costs and the error costs of enforcement, assume that the elites acting on behalf of the states in our hypothetical world behave as if they have the preferences for reciprocity revealed in laboratory experiments. This assumption implies that self-enforcement will be effective in a wide range of transactions not involving close-knit communities, or established on-going relationships. If reciprocal fairness is a potent complement to these traditional means of self-enforcement, then it might extend to “one-shot” investment agreements between states with no prior transactional history.

⁵⁵ A reciprocally fair party will not only earn a portion of the enhanced surplus in this transaction, but, by revealing her type (the bonus is paid), she will be able to develop a reputation for fairness that can be exploited at lower cost in future transactions. See Joseph E. Stiglitz, *The Theory of Screening, Education and the Distribution of Income*, 65 AM. ECON. REV. 283, 287 (1975).

In any transaction in which self-enforcement is an available option, the direct costs of self-enforcement typically will be less than the direct costs of coercive enforcement. This is because self-enforcement only requires an actor to expend costs to observe the behavior of the counterparty, while coercive enforcement requires the parties to expend *additional* resources in verifying that behavior to a third party.⁵⁶ Moreover, where significant measures of performance are observable but not verifiable, the direct costs of enforcement will include both the resources expended in verification as well as the welfare losses from writing second best contracts. Self-enforcement, on the other hand, permits parties to make credible promises regarding non-verifiable measures of performance, thus increasing joint surplus.⁵⁷ If direct enforcement costs were the only consideration, therefore, self-enforcement would generally be both cheaper and better than third-party enforcement.

But the advantages of self-enforcement are limited by two constraints. First, the further parties are from environments in which reciprocity can be predicted to work, the greater the risk that a self-enforcing sanction will not be credible. Second, and perhaps more significant, self-enforcement requires what we have termed “moral clarity.” Each party must be able to observe and properly characterize instances of primary misbehavior or other nonperformance and distinguish them from justifiable retaliatory responses to earlier instances of defection by the counterparty.⁵⁸ Accordingly, declines in the moral clarity of breaches will generate higher error costs for self-enforcement relative to external enforcement.

⁵⁶ There are conditions under which the parties’ direct costs of self-enforcement may exceed that of coercive enforcement. This would occur where the promisee has to expend resources in undertaking a punitive sanction for observable defection (say, a proportionate retaliation or a refusal to deal) and where the imposition of a coercive sanction was subsidized by the state. Under these conditions the cost of imposing a self-enforcing sanction may exceed the costs of verifying the breach to a third party.

⁵⁷ Self-enforcing agreements may also help to solve a multi-tasking problem. Assume, for example, that State A’s performance involves both verifiable and non-verifiable tasks. An agreement that was conditioned only on the verifiable tasks would be inefficient. Linking verifiable performance measures to compensation will cause State A to substitute away from the non-verifiable tasks to the compensated verifiable tasks, thus impairing overall performance. Bengt Holmstrom & Paul Milgrom, *Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design*, 7 J.L. ECON. & ORG. 24 (1991) (Special Issue). Self-enforcing agreements can avoid this problem by making any reward for cooperation dependant on the over-all performance of all tasks. Ernst Fehr, Alexander Klein and Klaus M. Schmidt, *Fairness, Incentives and Contractual Incompleteness*, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 72 (2001), at 26-29.

⁵⁸ See John K. Setear, note 6 supra, at 666-75.

Whenever the transactions are complex and the sequence of performances interrelated, moral clarity dissipates. Either party can mistakenly label a justifiable retaliation as a defection and vice versa. In these environments, third-party referees serve a valuable function by “calling fouls.” A disinterested observer may be in a better position to sort out complex behavior and, by “blowing the whistle,” can both empower the aggrieved party to respond in kind and forestall further retaliation by the breacher. A constraint on this capacity, however, would be the presence of observable but unverifiable factors. The third-party observer adds value only in cases where moral clarity dissipates at a greater rate than does verifiability.

Any calculation of the costs of enforcement thus must balance the higher direct costs of coercive enforcement against the potentially higher error costs of self-enforcement. Those error costs are particularly salient in the cases of complex agreements between nations with no history nor a necessary future. When expected error costs exceed the expected direct costs, the optimal enforcement regime will opt for some form of coercive, third-party enforcement.

5. Complementarity and Rivalry in Enforcement Mechanisms

Finally, one must consider the possible interaction of multiple enforcement mechanisms on each other. Three possibilities exist: The mechanisms may be independent, in that increasing one will have no effect on the benefits derived from increasing the others; they may be rivalrous, in that increasing one will decrease the returns from outlays on others; or they may be complements, in that increasing one will increase the benefits from others at no additional cost.

It seems plausible that the three aspects of self-enforcing mechanisms (retaliatory threats based on repeat transactions, reputation, and reciprocity) are complements. Even in particular cases where the pull of retaliatory threats and reputation may be weak, such as with the emergence of new regimes, states may learn to reciprocate because reciprocation pays off in many transactions over time. Reciprocation may also induce a virtuous cycle, in which engaging in cooperative behavior increases a regime’s preference for more cooperative behavior. Successful cooperation that generates a reputation for trustworthiness or produces returns in ongoing transactions both furthers a regime’s self-

interest and causes parties to learn to care more about the other party's payoff. This, in turn, may strengthen a regime's willingness to reciprocate voluntarily even where the prospect of retaliation is quite low.⁵⁹

Some experimental evidence regarding individuals does support this claim of complementarity.⁶⁰ Experiments have compared the effort levels of subjects who were given a single, anonymous opportunity to respond to generous offers with the effort levels in a similar game in which repeated interactions created the additional opportunity for retaliation.⁶¹ The results are, that while a significant fraction of individuals were motivated by reciprocity in the one-shot, anonymous transaction, repeated interactions caused a significant increase in the effort level.

Theorists have not settled on a single explanation for why this complementarity exists between reciprocity and retaliation, and these studies remain subject to the same reservations we expressed above about the experiments that produced evidence of a preference for reciprocal fairness.⁶² Nonetheless, the evidence seems sufficient to justify speculation about what might drive the observed complementarity. One conjecture notes that the properties of incentives created by repeated interactions are similar to the properties of incentives created by invitations to reciprocate. Self-enforcing incentives are imposed implicitly and ex post. Thus, for example, in a repeat game, tit-for-tat framework a cooperator can punish a defection ex post without risking offense to a potential cooperator by announcing in advance a sanction for defection.⁶³

⁵⁹ See F. Van Dijk, J. Sonnemans, & F. van Winden, *Social Ties in a Public Good Experiment*, 85 J. PUBLIC ECON. 275 (2002).

⁶⁰ Ernst Fehr & Klaus M. Schmidt, *Theories of Fairness and Reciprocity*, *supra* note 32, at 6-7; Ernst Fehr, Simon Gächter & Georg Kirchsteiger, note 32 *supra*; Ernst Fehr, Simon Gächter & Georg Kirchsteiger, *Gift Exchange and Reciprocity in Competitive Experimental Markets*, 42 EUR. ECON. REV. 1 (1998); Gary Charness, *Responsibility and Effort in an Experimental Labor Market*, 42 J. ECON. BEHAV. & ORG. 375 (2000); Ernst Fehr & Armin Falk, *Wage Rigidity in a Competitive Incomplete Contract Market*, 107 J. POL. ECON. 106 (1999); Simon Gächter & Armin Falk, *Reputation or Reciprocity*, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 19 (1999).

⁶¹ Ults Martin Brown, Armin Falk & Ernst Fehr, *Incomplete Contracts and the Nature of Market Interactions*, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 38 (2001).

⁶² See text at note 37 *supra*.

⁶³ See Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1681.

By contrast, incentives based on coercive enforcement are explicit and ex ante.⁶⁴ The difference in their nature might make these incentives rivalrous with those that operate implicitly and ex post. Experimental data indicate that, when offered a trust contract, a substantial number of individuals will both pay higher prices and extend higher levels of effort than narrow self-interest would dictate. When offered the same choices *plus* the possibility of obtaining a monetary sanction if the promisor shirks, the average price offered by buyers and the average effort given by sellers was *lower*. Without coercive enforcement, reciprocal fairness generates high levels of performance. But once the interaction is backed by coercion, reciprocity declines and overall performance is reduced.⁶⁵ These experimental results suggest that self-enforcing motivations based on reciprocity and explicit, coercive incentives may indeed be in conflict with each other. In particular, coercive enforcement may “crowd out” behavior based on reciprocal fairness.⁶⁶

Why might reciprocal fairness and repeated interactions complement each other while reciprocal fairness functions as the rival of coercive enforcement? One conjecture focuses on the fact that coercive third-party enforcement is structured as a zero sum game in which the promisee threatens ex ante to sanction the promisor for subsequent nonperformance. The explicit, ex ante nature of coercive

⁶⁴ Ernst Fehr & Armin Falk, *Psychological Foundations of Incentives*, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 95, at 19-20 (2001).

⁶⁵ There are two primary reasons why the amount of voluntary cooperation declined when the contract was enforceable coercively. First, shirking by sellers increased. This was true even where the expected costs of shirking exceeded the expected returns to the seller. Second, reciprocity either in the form of generous offers by buyers and/or reciprocating efforts by sellers vanished almost completely. Where shirking was expected to payoff for the sellers, they chose the minimum effort in the vast majority of cases. In addition, in those instances where buyers offered more generous prices above the minimum, sellers did not reciprocate with greater efforts. *See* Ernst Fehr & Simon Gächter, note 56 *supra*, at 15-18.

There are other experiments that have reported similar effects from the introduction of coercive enforcement. *See, e.g.*, Iris Bohnet, Bruno Frey, & Steffen Huck, *More Order with Less Law: On Contract Enforcement, Trust and Crowding*, John F. Kennedy School of Government, Harvard University (KSG Working Paper No. 00-009 (2000)); Uri Gneezy & Aldo Rustichini, *A Fine is a Price*, 29 *J. LEGAL STUD.* 1 (2000); Ernst Fehr & Bruno Rockenbach, *Incentives and Intentions – The Hidden Rewards of Economic Incentives*, University of Zurich (mimeo 2000). An extensive literature in social psychology also considers the crowding out of intrinsic motivations. *See, e.g.*, Edward L. Deci, R. Koestner, & Richard M. Ryan, *A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation*, 125 *PSYCH BULL.* 627 (1999).

⁶⁶ Robert E. Scott, *Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1689-91. The experimental research on crowding out is still at a preliminary stage. Thus, we do not currently understand precisely *why* or *when* reciprocity and voluntary cooperation will be undermined by coercive enforcement. There is some evidence, however, that framing effects influence the crowding out phenomenon. For example, if the explicit incentive is framed as a bonus from a base offer rather than as a fine for nonperformance, the levels of reciprocity are considerably greater. *See* Ernst Fehr & Simon Gächter, *Do Incentive Contracts Undermine Voluntary Cooperation?* (Inst. for Empirical Research in Econ., Univ. Of Zurich, Working Paper No. 34, 2002), at 26-30.

sanctions may undermine the instinct to reciprocate. Fair types may regard the threat of coercive enforcement as simply unfair, as they are willing to reciprocate voluntarily, while selfish types may interpret the threat of sanction through third-party enforcement as a signal that the promisee is unlikely to be a reciprocator.⁶⁷ The same explicit threat does not exist in the case of repeated interactions where the implicit sanction (terminating the relationship) is imposed *ex post* and only after defection has been observed. In that sense, *ex post* punishment may be perceived as “fairer” than the *ex ante* announcement of sanctions for breach of an obligation.

Another conjecture, derived from the apparent preference for fairness and studies of reputational sanctions, proposes that group solidarity may reduce the free riding problem associated with retaliatory enforcement of collective informal sanctions but play no role in formal coercive actions. Parties may consider it fair to absorb some costs in the monitoring and sanctioning of others if they regard this burden as borne generally. Assignment of a special role to a specific entity, however, may undermine a sense of fairly shared collective responsibility. Establishing a person or process responsible for enforcing an agreement may induce others to specialize away from enforcement. Finally, the reputational benefits derived from being seen to cooperate may diminish if the audience perceives the behavior as the product of coercion.

C. Summary

In this Part, we have argued that states design their bilateral and multilateral agreements to be enforceable when they wish to make relation-specific investments that otherwise would not be made. It is well-understood that self-enforceable agreements are an alternative to third-party, coercive enforcement. Where self-enforcement is effective, it is more efficient than third-party enforcement because it better responds to the effects of asymmetric information.

⁶⁷ Another speculation that might explain crowding out is the perception that coercion is a “hostile” action. A coercive sanction is always framed as a threat (“if you don’t perform, then I will invoke authority. . .”). To the extent that intentions matter in motivating reciprocity, the *ex ante* threat may be interpreted as a hostile intention. A liquidated damages clause in a contract may thus be perceived as an indication of distrust. If sellers perceive the damages clause as a hostile act they may be less willing to put forth the same quality of efforts as compared to a situation in which the first mover sends a trusting signal. Ernst Fehr & Simon Gächter, note 66 *supra*, at 14.

Contract theorists traditionally have believed that self-enforcement is limited to contexts where reputation or repeated interactions are sufficient to make promises credible. Recent work in experimental economics suggests that reciprocal fairness is a potent supplement to the traditional means of self-enforcement.⁶⁸ But this expansive view of the potential of self enforcement does not mean that parties to international agreements have no need to create mechanisms that rely on third-party enforcement. Self-enforcing agreements have a common feature: The agreements are simple in form, clear in commitment and are structured to create opportunities for parties to reciprocate in ways that expand the contractual surplus.⁶⁹ When inter-state investment depends on complex agreements where performance is interactive, the high error costs of self-enforcement argue for mechanisms that provide a means of verifying performance and nonperformance to a third-party referee.

The preceding analysis has sought to demonstrate that a contract theory approach to international agreements between states offers a useful framework for both positive explanation and normative critique. A particular feature of the framework we have developed focuses on what we will call rational reciprocity. In the next part, we relax the restrictive assumptions of our informal model in order to apply the insights of the rational-reciprocity approach in a more realistic setting.

II. THE ENFORCEMENT OF INTERNATIONAL AGREEMENTS

We observed above that self enforcement is a pervasive, but not exclusive, mechanism for giving instrumental effect to international agreements. Here we examine this point in some detail. We explain how international agreements invoke self enforcement, but we also rebut the claim that treaties do not rely on other strategies to promote compliance. We explore three such alternatives: third-party

⁶⁸ Whether reciprocal fairness is a learned behavior that derives from the benefits of cooperation in repeated interactions, or instead is an intrinsic motivation, remains an open question. But the evidence suggests that the domain of self-enforcing agreements extends to environments in which a reputation for trustworthiness and the discipline of ongoing relationships are relatively weak forces. See Robert E. Scott, *Self-Enforcing Indefinite Agreements*, note 16 *supra*, at 1674-75.

⁶⁹ Where transactions are complex and the respective promises interrelated, a failure to perform a promise may not be obvious and thus reciprocity may not serve to make the promises credible. Given the highly interactive nature of the respective obligations of each party, it may be difficult to know, for example, whether one party's refusal to respond cooperatively in a particular case represents unfair or selfish behavior or an appropriately measured, retaliatory response to an earlier instance of non-cooperation by the other. Robert E. Scott, note 28 *supra*, at 2050-51.

arbitration of claims brought by nongovernmental actors; accepting the jurisdiction of a dispute resolution body embedded in important institutional relationships; and authorizing independent domestic courts to entertain claims based on an agreement. Each of these mechanisms possesses the critical elements of coercive third-party enforcement: Persons affected by violations (and not just the states who are parties to the agreement) may present their claim to a disinterested third party that has the authority and capacity itself to impose substantial sanctions on the violator. Each may serve as a substitute for, or a supplement to, self enforcement. We illustrate how the particular enforcement strategies used by these agreements is broadly consistent with the rational-reciprocity approach.

A. Self Enforcement

A wide array of international arrangements induce cooperation without invoking external coercion to induce compliance. Arms control agreements, joint ventures for the production of advanced military technology, OECD recommendations on international competition policy, the Basel Accords on capital adequacy standards for financial institutions and myriad other instruments and concordats constitute a body of commitments that operate without regard to formal means of enforcement.⁷⁰ Other agreements contain elaborate dispute resolution provisions but rely ultimately on self enforcement to induce compliance. We examine in detail the mechanisms of self enforcement implicated by these arrangements and explain why states agree to them, rather than other enforcement mechanisms.

1. Counterparty Retaliation in International Agreements

Observers have noted that “almost all nations observe all principles of international law and almost all of their obligations almost all of the time,” but a comprehensive explanation of this

⁷⁰ Our examples in text include both agreements that the parties intend to have “legal force” but invoke no third-party coercive enforcement mechanism, and “soft law” that by its own terms creates no legal obligation. *See* Jack L. Goldsmith & Eric A. Posner, note 8 *supra* (surveying the field of soft law and providing a positive theory for choice of no formal enforcement). Goldsmith and Posner focus on the decision whether to invoke a sense of legal obligation under international law. We, in by contrast, examine specific enforcement mechanisms, and in particular self enforcement. We observe that the practice of states tends to relegate both soft law and a wide array of formal treaty commitments to self enforcement, although some mechanisms of self enforcement (such as WTO dispute resolution) entail a higher degree of formality than do others.

phenomenon remains elusive.⁷¹ In a wide range of instances, compliance occurs without the operation of any formal third-party enforcement mechanism. Before we consider what third-party enforcement does, we must have some understanding of how self-enforcing agreements work in the context of international commitments.

To draw on contract theory to understand international bargaining, we must specify how state-to-state negotiations resemble private contracting. In the case of private contracts, theory starts with the assumption that contractors seek to maximize their joint welfare. In the case of states, we have to specify more carefully who the contractors are and under what institutional constraints they operate.

First, we relax the assumption made in the previous part that governing elites seek to maximize national welfare. Taking our lead from rational choice theory in international relations, we instead will assume that each state has a political elite that seek to maximize its own welfare and faces uncertainty about the future.⁷² Importantly, however, we suppose that the institutional constraints on the ability of political elites to pursue self-interested goals varies according to the characteristics of the domestic regime.⁷³ States ruled by dictatorships or other authoritarian regimes, for example, constrain the leaders only in the sense that some choices might lead to an unacceptably high risk of domestic or foreign overthrow. Constitutional democracies with separation of powers, in contrast, constrain the political elite both through mandated power-sharing and check-and-balances arrangements and by posing a risk of electoral defeat.

The recognition that political elites are agents who pursue their own self-interest necessarily raises the question of whether rent-seeking may distort their behavior. In Part I we assumed that political elites acted to maximize the joint welfare from their international commitments. But in this

⁷¹ Louis Henkin, *How Nations Behave* 47 (1979). Cf. Thomas M. Franck, *Legitimacy in the International System*, 82 AM. J. INT'L L. 705 (1988):

The surprising thing about international law is that nations ever obey its strictures or carry out its mandates. This observation is made not to register optimism that the half-empty glass is also half-full, but to draw attention to a pregnant phenomenon: that most states observe systemic rules much of the time in their relations with other states.

⁷² See Stephen D. Krasner, *SOVEREIGNTY: ORGANIZED HYPOCRISY* 7, 43-44 (1999).

⁷³ Cf. Robert Putnam, *Diplomacy and Domestic Politics: The Logic of Two-level Games*, 42 INT'L ORG. 427 (1988) (exploring interactions between international relations and domestic politics).

more realistic framework, the accuracy of that assumption may be challenged. If elites cannot be controlled by their citizens whose interests they represent, then they may be maximizing their own welfare at the expense of social welfare. No one doubts that political elites act in ways that are inconsistent with the collective interests of their citizens. Nevertheless, we will assume that elites do select enforcement mechanisms that are designed to maximize the surplus from their international commitments. After all, political elites seek rents either by diverting wealth to themselves or to the interests groups they represent. In any case, elites have no interest in degrading the quality of their international agreements; these agreements create the wealth they may then attempt to divert.⁷⁴

We focus on elites and regimes, rather than states, to capture an important phenomenon in international relations: Regime changes occur and with them both the internally known preferences and the externally manifested characteristics of the political elite.⁷⁵ In form, international agreements bind states, but more realistically they bind political elites. A regime change may alter the elite's preferences, and thus its susceptibility to retaliatory threats.

Furthermore, we assume that reputations attach to regimes more than to states. A change in regime thus can dissipate the reputational effects of a prior elite's behavior. Without this assumption, one would have to assume that reputational effects are pervasive and strong in almost all international

⁷⁴ Alan Schwartz & Robert E. Scott, note 16 *supra*, at 550-1. We acknowledge that we have only partly relaxed the assumption of no agency costs in regime behavior, inasmuch as we do not specifically consider the possibility of a regime acting as an agent for particular groups, such as producers, and how domestic principals might approach enforcement. One of us has explored this problem in general, but not its effect on the choice of enforcement mechanisms. Paul B. Stephan, *Barbarians Inside the Gates: Public Choice Theory and International Economic Law*, 10 AM. U. J. INT'L L. & POL'Y 745 (1995); Paul B. Stephan, *Accountability and International Lawmaking: Rules, Rents and Legitimacy*, 17 NW. J. INT'L L. & BUS. 681 (1996-97); Paul B. Stephan, *The Futility of Unification and Harmonization in International Commercial Law*, 39 VA. J. INT'L L. 743 (1999); Paul B. Stephan, *Courts, Tribunals and Legal Unification – The Agency Problem*, 3 CHI. J. INT'L L. 333 (2002). Our present intuition is that the agency costs in the regime-domestic constituency relationship do not detract from our argument, but we leave it to later work to explore this issue.

⁷⁵ We use "regime" to refer to the domestic constellation of personnel and structures that generate a state's choices about international relations. We emphatically do not wish our usage to be confused with the conventional use of "regime" by international relations theorists to refer to the international structures that condition the choices of states in their dealings with each other. For outstanding examples of regime theory in the international relations sense of the term, see INTERNATIONAL REGIMES (Stephen D. Krasner ed. 1983); Robert O. Keohane, *Neoliberal Institutionalism: A Perspective on World Politics*, in INTERNATIONAL INSTITUTIONS AND STATE POWER I (Robert O. Keohane ed. 1989); Robert O. Keohane, *The Demand for International Regimes*, 36 INT'L ORG. 325 (1982).

agreements, as states-qua-states demonstrate considerable durability.⁷⁶ In a world of semi-permanent states, almost all states would have to anticipate future interactions with other states, even for those states not engaged in extensive repeat transactions with other states. Turnover of regimes explains why anything more than reputation-based sanctions by the “international community” is needed to enhance the credibility of promises made through international agreements.

For purposes of our analysis, we will not seek to define either elite or regime too tightly, and instead will rely on soft sociological intuitions about these concepts. On the one hand, a political elite can remain stable even as it undergoes gradual evolutionary change. On the other hand, a state can exhibit continuity in its type of government but experience regime change. Consider, for example, an authoritarian state where one ruling clique ousts another, in the process purging the political and technocratic leadership. The new clique, we assume, has some capacity to disavow the actions of its predecessors and thus avoid at least some of the benefits and burdens of its predecessor’s reputation, even though the form of government remains the same. Conversely, a stable political system can change its leaders without altering either its regime or the governing elite. Thus we assume that the normal replacement of one political party by another in a democratic parliament or executive does not dissipate a state’s reputation regarding international relations, any more than, historically, changes in the identity of the Party General Secretary in and of themselves affected the Soviet Union’s core preferences or reputation.⁷⁷

Once one considers these distinguishing attributes of states, it seems evident that the threat of future retaliation can induce parties to an international agreement to honor their obligations. As the rational-reciprocity approach predicts, many international pacts do involve mutual commitments to forego actions that would benefit the promisor and harm the promisee, such as increasing armaments or raising tariffs. Where the benefits and costs foregone are sufficiently symmetrical between the parties and project reliably into the future, a promisee can credibly threaten to punish a defaulting promisor

⁷⁶ See Tanisha N. Fazal, *State Death in the International System*, __ INT’L ORG. __ (2004); John K. Setear, Taking Both Biology and International Law Seriously: Evolutionary Biology, Neo-Realist Theories of International Relations, and the Promise(s) of International Law, UVA School of Law, Public Law Working Paper No. 03-19 (observing low rates of state death in modern era).

⁷⁷ For an insightful study of the role of enduring attributes of the Soviet Union in influencing its foreign policy, in addition to the impact of leadership changes, see Morton Schwartz, *THE FOREIGN POLICY OF THE USSR: DOMESTIC FACTORS* (1975).

by claiming the benefits and inflicting the costs that the agreement otherwise precludes. In a bilateral arms treaty, such as the several nuclear arms agreements between the Soviet Union and the United States, each party had an incentive to adhere to its obligations to the extent that the other party had the capacity to respond to breaches by augmenting its arms beyond the treaty limits.⁷⁸

The efficacy of retaliatory threats in these agreements depends on two factors – the ease of detection of defections from the agreement and the ability of the parties to make credible and symmetrical retaliatory threats. The Molotov-Ribbentrop Pact, for example, did not specify any observable, much less verifiable, conditions short of armed invasion that would constitute a breach, and thus pushed back defection to the point where retaliation would be indistinguishable from the self-preserving actions that would occur in the absence of any agreement.⁷⁹ Moreover, exogenous changes

⁷⁸ Richard Posner has offered a clever example of how retaliatory threats can work independently of reputational incentives. He speculates as to why states generally comply with international conventions on the treatment of prisoners of war. When engaged in war, states may wish to have a reputation for dangerous, even irrational ferocity but still adhere to minimum standards of decency regarding prisoners:

If both sides hold the same number of prisoners, each has a simple and effective means of retaliation if its opponent mistreats its prisoners. If one has very few prisoners relative to the other, this means it is probably losing the war and so will fear punishment if it mistreats its few prisoners; also there will be few benefits, since the cost of maintaining only a few prisoners will be small. The winning side, which holds a disproportionate number of prisoners, can afford to maintain them, precisely because it is winning, and so has little to gain from mistreating them, especially since there is some, though perhaps only a small, risk that its opponent will retaliate against the prisoners that it holds.

Richard A. Posner, *Some Economics of International Law: Comments on Conference Papers*, J. LEG. STUD. S321, S325 (2002). For virtually the same argument in a judicial opinion, see *Hamdi v. Rumsfeld*, 316 F.3rd 450, 468-69 (4th Cir.) (Wilkinson, J.), *rehearing denied*, 337 F. 3rd 335 (4th Cir. 2003), *cert. granted*, 124 S. Ct. 981 (2004). This conjecture opens up another puzzle, however: Why do states bother to codify the rules regarding prisoners of war if they know they will face retaliation if they lose? During World War II, for example, Japan was not a signatory to the 1929 Geneva Convention Relative to the Treatment of Prisoners of War, but after the war the United States still held war crime trials and executed many Japanese soldiers for mistreating U.S. prisoners. One possibility is to regard some instruments not as contracts, but rather as a form of coordination. From this perspective, a state adheres to the Geneva Conventions not only to bind itself, but to announce to potential future adversaries what standards it will apply to their conduct. A state might join such an agreement to limit its obligations to those contained in the text, rather than facing the uncertain standards of victors' justice. As this speculation takes us away from the analysis of agreements and into the realm of so-called customary international law, *see note 5 supra*, we will not pursue it further.

⁷⁹ The secret protocol associated with the Pact did permit the parties to take some concrete steps to signal cooperation. This instrument divided up Poland between the two powers and committed Germany not to interfere with Soviet occupation of the then-independent Baltic republics. Treaty of Non-Aggression, Aug. 23, 1939, F.R.G.-U.S.S.R., 1939 RGBI. II, No. 38, *translated in* 7 DOCUMENTS ON GERMAN FOREIGN POLICY, 1918-1945, at 245 (Series D) (U.S. Dept. of State 1956); Secret Additional Protocol, Aug. 23, 1939, *in* 7 DOCUMENTS ON GERMAN FOREIGN POLICY, *supra*, at 246. It appears that the clarity of the dispositions contained in these protocols, unlike the ongoing nonaggression obligation, lent themselves to self enforcement. We concede, however, that the creation of a common border between the two powers where buffer states had existed did have some of the elements of a trust agreement. For the follow-on agreements that fixed the border after the joint liquidation of Poland, see German-Soviet Boundary and Friendship Treaty, with Confidential Protocol and Two Secret Supplementary Protocols and Declaration, September 28, 1939, v.1.5.1940 (RGBI. II S. 4-8), *translated in* 8 DOCUMENTS ON GERMAN FOREIGN POLICY 1919-1945, ser. D (U.S. Dept. of State 1956), at 164-67 & 208-212.

– Germany’s easy conquest of Western Europe and the consequent removal of an immediate military threat to its west – led Hitler to devalue the benefits of avoiding an armed conflict with the Soviet Union. In other words, as of the spring of 1941 the cost of Soviet retaliation no longer seemed sufficient to induce German compliance with its obligations under the Pact. The Molotov-Ribbentrop Pact thus illustrates an international agreement with extremely weak self-enforcement. Put differently, the Pact in retrospect exhibited great vulnerability to end-game problems.

The events that resulted from the Molotov-Ribbentrop Pact suggest another point. One should not confuse self-enforcement based on reciprocal threats of retaliation with unilateral imposition based on coercion. A state that surrenders to superior force is not behaving cooperatively, but rather demonstrating a preference for survival or a distaste for discomfort.⁸⁰ Coercion signals no reliable information about either the durability of coercing party’s preferences or the credibility of the coerced party’s promises, except that it suggests an inability to cooperate reciprocally on the part of both. If one interprets the Molotov-Ribbentrop Pact not as a mutual commitment to refrain from aggression, but rather as a unilateral obligation by the Soviet Union not to take precautions against Germany, then the Pact teaches very little about self-enforcement. Similarly, we would not regard Czechoslovakia’s acceptance of its absorption into Germany through the *Anschluss* as reflecting compliance with an international agreement, because only in a formal and meaningless sense of the word did Czechoslovakia “agree” to become part of Germany.⁸¹

Trade agreements offer another example of self enforcement through retaliatory threats. A state that lowers its tariffs or otherwise reduces trade barriers gives up potential revenue and diminishes

⁸⁰ For development of the point, see Jack L. Goldsmith & Eric A. Posner, note 5 *supra*, at 1123-24.

⁸¹ Consider the following finding of the Nürnberg Tribunal:

On the 14th March, 1939, the Czech President Hacha and his Foreign Minister Chvalkovsky came to Berlin at the suggestion of Hitler and attended a meeting at which the defendants Ribbentrop, Göring and Keitel were present, with others. The proposal was made to Hacha that if he would sign an agreement consenting to the incorporation of the Czech people in the German Reich at once, Bohemia and Moravia would be saved from destruction. He was informed that “German troops had already received orders to march” and that any resistance would be broken with physical force. The defendant Göring added the threat that he would destroy Prague completely from the air. Faced by this dreadful alternative, Hacha and his Foreign Minister put their signatures to the necessary agreement at 4.30 in the morning, and Hitler and Ribbentrop signed on behalf of Germany.

Judgment of the International Military Tribunal for the Trial of German Major War Criminals,

protection desired by its producers to the benefit of exporting producers based in other states. If one party defects, the other may respond symmetrically. As long as each party has a domestic market that other parties covet, its threat to protect its market will encourage others not to protect theirs. By “a market,” of course, we mean not just the immediate trade balance, but the discounted value of a series of anticipated future interactions.⁸²

Yet another illustration of the function of retaliatory threats in international agreements involves sovereign indebtedness. Throughout history and manifestly today, sovereigns manage to acquire significant assets based only on their promise to repay the debt, in spite of the capacity of many unilaterally to repudiate their debts or to degrade them through currency manipulation.⁸³ What most powerfully explains the willingness of lenders to treat these promises as credible is a conviction that governments will seek private financing of public debt into the indefinite future. Thus, although government default seems about as frequent an occurrence as private failures to honor debts, typically lenders and states negotiate rescheduling rather than invoke the coercive powers of third parties. The defaulting state pays a penalty in terms of high interest rates rather than by absorbing a lump-sum sanction. The exception to this pattern reinforces the point: Revolutionary regimes often declare a break with the past, repudiate past obligations and initially pursue a policy of financial autarky. Over time, these regimes seek to return to the capital markets, and when they do so belatedly negotiate settlements of their predecessors’ obligations.

Credible and reciprocal threats of direct retaliation explain much of the enforceability of international agreements. First, a large portion of extant international agreements are bilateral, and thus

⁸² Conversely, trade agreements also can illustrate nonreciprocity and coercion. A country that derives greater benefit from access to another country’s domestic market than it can offer to exporters from that country (again accounting for future expectations) may not have a credible retaliatory threat. When the United States in the early 1950s violated its obligation to the Netherlands to allow cheese imports, the Netherlands did not retaliate in spite of an impartial finding that the United States had infringed its rights under the General Agreement on Tariffs and Trade (GATT). *Netherlands Action Article XXIII:2 to Suspend Obligations to the United States*, GATT BISD 1S/62 (1952). We can surmise that, at least at that time, the Netherlands rationally believed that the United States did not regard the discounted value of present and future access to the Netherlands market as sufficiently valuable to induce any change in its actions. Put plainly, the United States could coerce the Netherlands not to retaliate for its unilateral departure from its GATT obligations.

⁸³ See generally Charles P. Kindleberger, *A FINANCIAL HISTORY OF WESTERN EUROPE passim* (2nd ed. 1993) (documenting practice of sovereign debt).

identify precisely the persons who have an interest in inducing enforcement. Bilateral agreements also are the type of accord most likely to produce substantial internalization of the agreement's benefits by the parties. Second, many agreements seek to augment welfare through a mutual exchange of concessions. Where the expectations underlying the agreement are realized, withdrawing the concession would harm the other party.

At the same time, the prospect of party retaliation is insufficient to maximize the value of all international agreements. First, a significant fraction of agreements are multilateral, sometimes extensively so, leading to a free rider problem. If retaliation is costly and the benefits of adherence are shared by a large number of parties, no one party may have a sufficient incentive to bear the burden of retaliation. Knowledge of this fact diminishes the threat value of retaliation.⁸⁴ Second, some agreements generate benefits to persons besides the parties. We should not expect any party to absorb costs to protect those externalized benefits. And third, all agreements remain vulnerable to the end-game problem. As state preferences change, due either to exogenous events or a transformation of a political elite, one or more parties may have less to lose from retaliation and thus become prone to defection. Understanding this, all parties may underinvest in compliance.⁸⁵ To the extent we nonetheless observe some enforcement of such agreements, we need to consider complementary self-enforcing mechanisms.

2. Reputation in International Agreements

Reputation is an additional mechanism for self-enforcing international agreements. As we observed in Part I, there exists extensive documentation of the effectiveness of community-based sanctions as a means of inducing the compliance of individuals with an obligation.⁸⁶ This dynamic applies to states as well.⁸⁷ A reputation for honoring commitments benefits a state by increasing the

⁸⁴ For a discussion of the incentives created by breaches of multilateral treaties, see John K. Setear, note 15, at 33-38.

⁸⁵ Robert E. Scott, note 28 *supra*, at 2033.

⁸⁶ See note 12 *supra*.

⁸⁷ See Andrew Guzman, note 13 *supra* (analyzing the economics of international reputation and arguing for its central role in explaining adherence to international law). Other scholars have recognized the role of reputation but contend that it has a more limited function than Guzman acknowledges. See George W. Downs & Michael A. Jones, *Reputation, Compliance, and International*

possibilities of future beneficial cooperation. Monitoring compliance need not be costly (although the Molotov-Ribbentrop Pact provides a counterexample), and information about compliance, if verifiable, may be easy to share among states. The set of potential future parties presumably is larger than the set of present parties to any particular international agreement, so more states have an incentive both to acquire information about the behavior of other states and to generate a positive reputation for international probity.

In some instances reputational incentives can provide a better explanation for compliance with international commitments than does reciprocal retaliation. A state facing an end-game problem in one agreement must understand that its behavior will affect its future prospects for international cooperation through other agreements. Thus, even if it wishes to dispose of its current relationship opportunistically, it must consider how such opportunism will affect its future prospects. Moreover, all potential states can make reputation-based responses to a state's behavior, not just those with which a state currently enjoys formal treaty relations. A concern for its reputation thus will induce a state to cooperate even in cases where it is indifferent to retaliation by its counterparty.

Reputation, however, is not a complete explanation of international behavior. One complication derives from our observation above that reputations attach to regimes rather than to states. Because reputation has value only to the extent it has an impact on a state's future prospects, one must introduce a discount factor to reflect the present significance of those prospects. It seems reasonable to assume, however, that discount factors vary among regimes. A state which has a stable history and no significant internal or external threat should have a regime that values future prospects nearly as much as present ones. The ruling elite in a state undergoing radical transformation with an uncertain outcome, in contrast, should devalue all future payoffs, including both the benefits and harms associated with its reputation. To take an example from recent events, the reputational effects to the United States and the United Kingdom regarding their obligations under the United Nations Charter with respect to the invasion of Iraq are likely to be long-lasting, while the effects on the soon-to-depart Baathist regime probably were less significant. More generally, disputes between foreign investors

Law, 31 J. LEG. STUD. S95 (2002); Jack L. Goldsmith & Eric A. Posner, note 8 *supra*, at 136-37.

and host countries tend to arise mostly after a regime change, whether the product of a violent revolution (as in America in 1783 and Russian in 1917) or due to more ordinary domestic turmoil.⁸⁸

A second problem results from the difficulty of specifying the elements of a regime's reputation for compliance with its international obligations. Andrew Guzman, in the course of an extensive analysis of the role of reputation in explaining compliance, suggests that a state's reputation may be both compartmentalized and dependent on the character of the international obligation at issue. He hypothesizes, for example, that violations of an arms control treaty may not affect how others perceive a state's propensity to honor trade agreements, and that dishonoring "soft law" obligations may count for less than transgressing a commitment that is explicitly legally binding.⁸⁹

We agree with Guzman that a state's reputation for complying with international obligations, and hence the role of reputation as a self-enforcement mechanism, may not be monolithic or linear. We are skeptical, however, about his particular conjectures. He comes close to arguing that reputational effects vary precisely with the scope and content of the specific international obligation at issue. This collapses the question of the reputational effects and formal international law doctrine in a way that assumes that international law experts are the only relevant audience. We suspect that reputations can vary in ways that do not coincide with formal compliance with international law obligations. To take a concrete example, the effect on Japan's reputation due to its treatment of its prisoners during World War II depended not at all on the fact that Japan had not joined the 1929 Geneva Convention.⁹⁰

This exposes a third complexity. If reputations are messy and the impact of particular actions on them are unclear, then the significance of reputation as a constraint on behavior comes into question. Once one disposes of the extremes of outlaw recklessness and obsessive punctiliousness, more finely grained assessments of reputation become problematic. What should one make, for example, of the decision of the United States in 1985 to withdraw its submission to the compulsory jurisdiction of the

⁸⁸ Cf. *Dames & Moore v. Regan*, 453 U.S. 654, 678-84 (1981) (reviewing history of investment disputes between United States and new regimes).

⁸⁹ Andrew Guzman, note 13 *supra*, at 1880, 1883.

⁹⁰ See note 78 *supra*.

International Court of Justice? One prominent scholar described this as a “wanton act,” although the United States had scrupulously complied with its obligations regarding the Court’s jurisdiction.⁹¹ Did the United States encounter greater costs in its international negotiations after that date? The growth of its multilateral activity afterwards, and particularly the building of the coalition that fought the first Gulf War, suggests not.

A device that can clarify the connection between action and reputation is a third-party dispute resolution system.⁹² The GATT and its successor, the WTO, provide important instances of an international agreement specifying a self-contained process to declare compliance and violations. Only parties to the agreement (that is to say, states) can invoke the process, and the dispute settlement body has no coercive powers beyond a capacity to declare whether actions do or do not comply with the agreement.⁹³ The agreement thus uses dispute settlement to generate specific information about compliance without expending resources to coerce behavior.⁹⁴

But even self-contained dispute resolution has its limits. First, the quality of the information about members rests largely on the dispute resolution body’s reputation for accurate and disinterested assessments. It seems unrealistic to take the latter for granted. Second, we doubt that observers systematically separate information based upon dispute settlement determinations from information about how states respond to compliance determinations. Evidence of indifference to the resolution of a dispute, as manifested both by a failure of noncompliers to alter their behavior and the absence of

⁹¹ Abram Chayes, *Nicaragua, the United States and the World Court*, 85 COLUM. L. REV. 1445, 1445 (1985).

⁹² See Eric A. Posner & John C. Yoo, *A Theory of International Adjudication* (draft 2004) (exploring role of third party adjudicators as generating information about both the meaning of and compliance with international agreements).

⁹³ On GATT and WTO dispute settlement generally, see Andrew Guzman & Beth A. Simmons, *To Settle or Empanel? An Empirical Analysis of Litigation and Settlement at the World Trade Organization*, 31 J. LEG. STUD. S205 (2002); Andrew Guzman, *The Political Economy of Litigation and Settlement in the WTO* (working paper); Paul B. Stephan, *Sheriff or Prisoner? The United States and the World Trade Organization*, 1 CHI. J. INT’LL. 49 (2000). Strictly speaking, the WTO dispute settlement body has the authority to determine the legality of both complained of conduct and proposed retaliation by a complaining party. It cannot award damages or authorize independent, *i.e.*, nonretaliatory, sanctions.

⁹⁴ For a more extensive discussion of formal dispute resolution based on self-enforcing agreements, see Tom Ginsburg & Richard H. McAdams, note 6 *supra*.

significant retaliation, would undermine the reputational impact of those determinations.⁹⁵ Third, as discussed above, not all issues of importance to contracting states are verifiable at an acceptable cost. As to nonverifiable issues, any pronouncements the dispute settlement system might make would be meaningless.

As with retaliatory threats, then, reputation provides an important but incomplete explanation of successful self-enforcement of international agreements. In particular, it does not provide a satisfactory account of the behavior of either new or soon-to-depart regimes or of the transitivity of reputation across international agreements. These lacunae might suggest that self-enforcement has serious limits as a positive theory to explain compliance with international agreements. There remains to be considered, however, whether states may have a preference for compliance that is independent of retaliatory threats and reputation. We now turn to that question.

3. Reciprocal Fairness as a State Preference

We noted above that laboratory studies provide substantial evidence of the existence of a preference for reciprocal fairness on the part of many, but by no means all, individuals. To what extent does a preference for reciprocal fairness extend to states? We argue that the individual characteristics identified by the laboratory studies may also extend to regime preference formation. Our argument rests on both inferences about regimes and some casual empiricism.

At the most fundamental level, regimes comprise individuals, and it is not wholly implausible to believe that those people who shape a regime's preferences have something in common with the persons studied in the laboratory experiments. We suspect, however, that the representativeness problem that underlies all experimental data may have special salience when the group of interest – elite decisionmakers responsible for conducting foreign policy – is atypical of the general population. We therefore do not insist on the relevance of the experimental data, other than as suggesting a line of investigation.

⁹⁵ For a preliminary but significant inquiry into the relationship between an international tribunal's structure and compliance with its decisions, see Eric A. Posner & John C. Yoo, note 92 *supra*.

It may be that elites must take into account a widely held preference in the general population for reciprocal fairness, even if elite members do not share this preference, due to democratic constraints on elite conduct. We doubt this argument has much purchase, however. First, a significant number of regimes face reduced or nonexistent democratic constraints. Some have no democratic elections; others, such as the members of the European Union, delegate significant discretion regarding international relations to supranational institutions. Second, we suspect that national foreign policy elites will have considerable capacity to mediate accounts of their international behavior with national electorates. Depictions of reciprocal fairness to some extent may substitute for actual conduct.

Notwithstanding these qualifications, there are still good reasons to believe that many regimes exhibit a preference for reciprocal fairness. First, states typically carry out international relations through bureaucracies. Ministries of foreign affairs, trade and the armed forces shape policy in the overwhelming majority of states. It seems plausible that these bureaucracies might display the tendency generally observed in such structures, namely, *ceteris paribus*, maximization of discretionary authority.⁹⁶ And exhibiting reciprocal fairness seems a good strategy for optimizing interactions with counterpart bureaucracies, which in turn should maximize each bureaucracy's power *vis-à-vis* its political masters.

A response to this conjecture might be that tension rather than cooperation may optimize bureaucratic power. Military bureaucracies in particular might increase their authority by minimizing interactions with their counterparts as part of a broader strategy of stoking tensions and feeding insecurity. One strand of late twentieth century thought, for example, maintained that a U.S. national security complex fattened its budgets and expanded its influence by contriving a permanent sense of crisis.⁹⁷

⁹⁶ Max Weber, *THE THEORY OF SOCIAL AND ECONOMIC ORGANIZATION* (A.M. Henderson & Talcott Parsons trans. 1947); William A. Niskanen, Jr., *BUREAUCRACY AND PUBLIC ECONOMICS* (1994); William A. Niskanen, *Bureaucrats and Politicians*, 18 *J.L. & ECON.* 617 (1985).

⁹⁷ Noam Chomsky, *HEGEMONY OR SURVIVAL: AMERICA'S QUEST FOR GLOBAL DOMINANCE (THE GLOBAL EMPIRE PROJECT)* (2003); Gar Alperwitz, *ATOMIC DIPLOMACY: HIROSHIMA AND POTSDAM: THE USE OF THE ATOMIC BOMB AND THE AMERICAN CONFRONTATION WITH SOVIET POWER* (1985).

On balance, however, this response seems far fetched. By and large, uncooperative behavior that fuels international tension creates disproportionately great risks for foreign policy elites. A crisis brings their performance under more scrutiny than usual, may set in motion dangerous events with consequences beyond the elite's control, and otherwise unsettles the stable rationality that bureaucracies generally seek. We suspect that the link between conventional bureaucratic incentives and reciprocal fairness is strong, although by no means absolute.

Anecdotal evidence that even military bureaucracies exhibit some preference for reciprocal fairness includes the existence of confidence-building measures to which states commit themselves. The Treaty on Conventional Armed Forces in Europe exemplifies such arrangements.⁹⁸ One might think that professional military organizations would oppose an agreement that constrains weapons deployments and the size of deployed forces. But the Treaty has substantial reporting and inspection requirements and rests on confidence-building measures such as embedded observers that earlier agreements had promoted.⁹⁹ These techniques engage military personnel in a range of cooperative behaviors with their counterparts, interactions that those involved seem to find desirable.

A further suggestion of preferences for reciprocal fairness among many states can be inferred from the growth in number, scale and scope of international organizations devoted to facilitating cooperative behavior. Growth may reflect many factors, of course, but it seems reasonable to associate an organization's reputation, which accumulates over time, with increased attractiveness to regimes already disposed to reciprocal fairness. To cite some important but by no means exclusive examples of growth, the International Monetary Fund and the World Bank had 29 members in 1946, 173 in 1992, and 184 in 2003; the GATT had 23 founding members in 1947, 102 in 1979, and its successor, the WTO, had 146 in 2003; the European Communities had 6 founding members in 1957, grew to 15 in 1994 and 25 in 2004. Each of these institutions also evolved from a specific-purpose

⁹⁸ Treaty on Conventional Armed Forces in Europe, Nov. 19, 1990, S. Treaty Doc. No. 102-8 (1991). *See generally* S.V. Kortunov, *Basic Principles of Reduction and Limitation of Conventional Forces* in INTERNATIONAL LAW AND INTERNATIONAL SECURITY—MILITARY AND POLITICAL DIMENSIONS 186 (Paul B. Stephan & Boris M. Klimenko eds. 1991) (analysis of Treaty and its implications by leading Soviet expert).

⁹⁹ Abram Chayes & Antonia Handler Chayes, *THE NEW SOVEREIGNTY—COMPLIANCE WITH INTERNATIONAL REGULATORY AGREEMENTS* 147 (1995).

entity (postwar reconstruction, currency stability, tariff reduction) to a much broader governance institution. We are prepared to believe that these organizations may stand for less than they seem.¹⁰⁰ But even discounting for the gap between the ambitions and the accomplishments of these organizations, their proliferation suggests that an increasing number of regimes prefer the kind of reciprocal and cooperative relations that membership in the organizations promotes.

4. Self-Enforcing International Agreements and Contract Theory

As we observed above, self enforcement is the norm in international agreements, not the exception. Before exploring departures from this norm, we consider whether the patterns in self-enforcing international agreements conform to the predictions of the rational-reciprocity approach. Many of these agreements involve fairly simple and clear commitments, either as stated or as a result of the dispute resolution process contained in the agreement. At least as importantly, they typically implicate observable but unverifiable conditions, such as perceptions of national security or “serious injury” to economic interests. Moreover, many agreements provide the parties with ongoing opportunities to signal a preference for reciprocity.

Consider first the broad class of arms control agreements. By forgoing technologically feasible procurement in instances where research, development and deployment take time, a party to these agreements exposes itself to the risk of opportunistic behavior by others. Increasing the credibility of a promise not to acquire a weapons system clearly would enhance the promisor’s expected return. At the same time, each promise carries an implicit reservation, namely that its adherence will not come at the price of destruction of the regime. The modern trend in these agreements has been toward increasingly elaborate inspection and verification, sometimes carried out by independent monitors, as in the case of the (ultimately unsuccessful) UN arms inspectors in Iraq between 1991 and 2003 and the various interventions of the International Atomic Energy Commission from Chernobyl in 1986 to Iran in 2004. Thus, although no arms control agreement of which we are aware employs formal, quasi-

¹⁰⁰ For expression of this skepticism, see Paul B. Stephan, *Accountability and International Lawmaking: Rules, Rents and Legitimacy*, 17 NW. J. INT’LL. & BUS. 681 (1996-97).

judicial third-party coercive enforcement, they do provide some space for noncoercive third-party dispute resolution.¹⁰¹

We believe the reasons why arms control regimes rely solely on self enforcement are reasonably clear and are consistent with contract theory. Parties have an incentive to make their commitments to arms control credible so as to induce reliance in others, but also require sufficient flexibility to abandon their commitments when grave security interests arise. To a greater extent than other kinds of international agreements, the subject of arms control involves regime security. Moreover, regimes plausibly do not regard the question of whether any state of affairs threatens its security as verifiable. Rather, regimes should regard the calculus as to what constitutes an unacceptable threat to their survival as uniquely nondelegable. At the same time, regime security seems reasonably observable. As a result, we should expect regimes to strive to enhance clarity and augment opportunities for reciprocal action. Observers report that this is the case.¹⁰²

A much studied example of a self-enforcing international agreement is the Uruguay Round Agreements and the World Trade Organization that administers them.¹⁰³ Like arms control agreements, enhancing the credibility of the respective promises augments the value of trade commitments. Also like arms control agreements, these agreements implicate fundamental issues of regime security, although of an economic rather than a military nature. Finally, and to a much greater extent than in arms control, the formalization and elaboration of noncoercive third-party dispute resolution has grown over time. The early GATT, ancestor of the Uruguay Round Agreements, contained no formal procedure as such, although one based on arbitration buttressed by self enforcement evolved to meet the needs of the parties. A 1979 Understanding formalized these customary practices.¹⁰⁴ The 1994 version involves greater institutional development, including the creation of a standing appellate body and strengthening of the statement of obligation associated with its determinations. But, as noted above, the WTO retains the fundamental structure of self enforcement. Only states who are parties to

¹⁰¹ Abraham Chayes & Antonia Handler Chayes, note 99 *supra*, at 179-83.

¹⁰² *Id.* at 189-96.

¹⁰³ For the agreements, see http://www.wto.org/english/docs_e/legal_e/legal_e.htm.

¹⁰⁴ Understanding Regarding Notification, Consultation, Dispute Settlement and Surveillance, reprinted at http://www.wto.org/english/docs_e/legal_e/tokyo_notif_e.pdf.

the Uruguay Round Agreements can invoke the dispute resolution process, and they may raise only their own economic injuries. The dispute resolution process informs the complainant what levels of retaliation would be proper under those agreements, but does not itself impose any sanctions. Violations result in party retaliation and reputation effects, and not third-party coercion.¹⁰⁵

The link between moral clarity and self-enforcement seems evident in the case of WTO dispute settlement. One particular trend is noteworthy. During the initial phase of the GATT system, from its 1947 creation until the 1979 changes embodied in the Tokyo Round Agreements, the obligations largely dealt with tariff reduction and comparable straightforward and determinable commitments. Observers documented high compliance and relatively infrequent resort to formal dispute resolution. Beginning with the Tokyo Round, and particularly with the 1994 Uruguay Round, the agreements extended to considerably more complex and interdependent commitments, especially reduction in nontariff trade barriers such as health and safety regulation. As the nature of the obligations have lost their clarity, reported violations have increased, arguably at a greater rate than under the old GATT, and we have seen a higher incidence of formal dispute resolution, more frequent refusal to end practices found to be inconsistent with WTO obligations, and greater criticism of the process by trade experts.¹⁰⁶

We offer these examples as representative of the kinds of self enforcement that pervades the field of international agreements. Further examples would not add much to an appreciation of the predictive

¹⁰⁵ Understanding on the Rules and Procedures Governing the Settlement of Disputes, http://www.wto.org/english/docs_e/legal_e/28-dsu.pdf. For discussion, Tom Ginsburg & Richard H. McAdams, note 6 *supra*; Andrew Guzman & Beth A. Simmons, note 93 *supra*; Andrew Guzman, note 93 *supra*; Eric A. Posner & John C. Yoo, note 92 *supra*. Guzman, working from the premise that WTO dispute resolution generates positive externalities in terms of greater regime clarity and deterrence of noncompliance that produces general harms, argues that coercive powers should be added to the system to induce greater resort to the dispute resolution process. Andrew T. Guzman, *The Cost of Credibility: Explaining Resistance to Inter-State Dispute Resolution Mechanism*, 31 J. LEG. STUD. 303 (2002). This argument ignores both the possibility that many Uruguay Round obligations may present verification issues and, more importantly, the evidence that coercive enforcement may crowd out self enforcement.

¹⁰⁶ Robert L. Howse & Elisabeth Tuerk, *The WTO Impact on Internal Regulations – A Case Study of the Canada-EC Asbestos Dispute*, in THE EU AND THE WTO: LEGAL AND CONSTITUTIONAL ISSUES 283 (Gráinne de Búrca & Joanne Scott eds. 2001); Eric A. Posner & John C. Yoo, note 92 *supra*; Alan O. Sykes, *Domestic Regulation, Sovereignty, and Scientific Evidence Requirements: A Pessimistic View*, 3 CHI. J. INT'L L. 353 (2002) (technical barriers to trade and WTO dispute resolution process); Daniel K. Tarullo, *Norms and Institutions in Global Competition Policy*, 94 AM. J. INT'L L. 478, 494 & n.57 (2000) (citing disputes over Cuba sanctions, meat hormones, preferences for former colonies, and dolphin protection).

value of the rational-reciprocity approach. There may well be instances where self enforcing international agreements lack clarity or rest on verifiable criteria, but these also tend to be cases where the importance of investments and the corresponding need to enhance promisor credibility is slight.¹⁰⁷ A real test of the relevance of the contract theory approach must come from a study of international agreements that go beyond self enforcement and incorporate some form of third-party coercion. We now turn to this inquiry.

B. *Coercive Third-Party Enforcement*

With respect to private parties, coercive third-party enforcement entails the state (or some comparable higher authority) threatening sanctions against defaulters. As to international agreements, third-party coercion seems problematic because there are no international entities capable of carrying out threats against those who dishonor the precepts of international law. As we will show, however, some enforcement mechanisms associated with international agreements do correspond to those underlying private contracts.

1. *The Elements of International Coercion*

Coercive third-party enforcement of international agreements has three critical elements. First, affected private parties, and not just states that sign an agreement, have the right to pursue violations. Second, a disinterested third party has the jurisdiction to consider the claim. Third, the third party has the authority and capacity to impose directly substantial sanctions on a violating state.¹⁰⁸

¹⁰⁷ Characteristic of this class of agreements are nonreciprocal commitments by states to respect the human rights of their own subjects. These agreements have considerable moral attraction, but they do not involve the reliance by one state on the commitment of others. As Goldsmith and Posner have observed, “[A]bsent special circumstances like the minority rights situation, a nation otherwise inclined to abuse its citizens gains nothing from declining to do so in return for a reciprocal commitment from another nation to do the same.” Jack L. Goldsmith & Eric A. Posner, *Understanding the Resemblance Between Modern and Traditional Customary International Law*, 40 VA. J. INT’LL. 639, 669 (2000). See Reservations to the Convention on the Prevention and Punishment of the Crime of Genocide, 1951 I.C.J. 15, 23 (characterizing human rights conventions as not based on reciprocity and exchange).

¹⁰⁸ Our definition of coercive third party enforcement corresponds in part to what some call transnational dispute resolution. See Robert Keohane, Andrew Moravcsik & Anne-Marie Slaughter, *Legalized Dispute Resolution: Interstate and Transnational in LEGALIZATION AND WORLD POLITICS* 73, 84-85 (2001). Unlike those authors, however, we focus not on the independence of the adjudicatory body (which in the case of investment arbitration is quite limited, at least using the criteria advanced by Eric A. Posner

We recognize that isolating these characteristics as internationally coercive implies a commitment to a particular theory of regime motivation. We will not unpack all of the implications of this commitment here, but we note several salient points. First, extending to private parties the capacity to invoke a tribunal's jurisdiction solves an agency problem that otherwise might lead to underutilization of the forum. A person affected by regime action (perhaps an investor injured by a regulatory decision of the host state) has a direct interest in obtaining redress. The regime administering the injured person's state might act on his behalf, but it also will take into account other aspects of its relationship with the injuring regime and may fail to prosecute the claim because of extraneous considerations.¹⁰⁹ Private party standing thus increases the likelihood that a tribunal will act against the interests of a regime bound by an international agreement.¹¹⁰

Second, when it comes to tribunals, disinterested does not necessarily mean independent. International agreements sometimes constitute tribunals in a manner that forces them to compete for business and therefore provides incentives for reaching outcomes that the disputants will regard as desirable. Such bodies differ from permanent tribunals that have mandatory jurisdiction over a set of disputes that can be resolved in ways that reflect the normative preferences of the tribunal members.¹¹¹

& John C. Yoo, note 91 *supra*), but on their coercive capabilities. To take an example, the U.S.-Iran Claims Tribunal has great coercive authority due to the fact that it presides over a large sum of money, which can be disbursed only pursuant to its decisions, but the body's jurisdiction is limited to the *ex post* resolution of a particular set of disputes, thereby limiting its independence. *See id.*

¹⁰⁹ A well known example of an agency problem leading to nonutilization of a tribunal is *Rich v. Naviera Vacuba S.A.*, 295 F.2nd 24 (4th Cir. 1961). Plaintiff sued an entity belonging to Venezuela in a commercial dispute. Under U.S. law at that time, federal court jurisdiction over a foreign sovereign entity required the consent of the U.S. government, and the State Department had announced that it automatically would consent to suits based on commercial activity. In this particular case, however, the State Department blocked the suit, evidently as compensation to Venezuela for its cooperation in resolving an airline hijacking dispute.

¹¹⁰ We note that the absence of private enforcement also excludes international tribunals responsible for the prosecution of war crimes and similar international offenses. We concede that these bodies have coercive authority as long as any one national state is willing to imprison those convicted by these tribunals. We nevertheless omit them from our discussion, largely because of doubts about the general significance of international agreements creating these tribunals. Most of involved post hoc responses to conflicts that had ended, based on the Nürnberg model. The only exception, the International Criminal Court based on the 1998 Rome Statute, has yet to hear a case and, we suspect, is unlikely ever to acquire a substantial docket. Posner and Yoo predict that side agreements reached by peacekeeping countries, such as those already reached by the United States, will prevent the International Criminal Court from doing much real work. Eric A. Posner & John C. Yoo, note 92 *supra*.

¹¹¹ *Id.* For further discussion of the distinction between dispute-resolution and expressive functions of tribunals, including those applying international law, see Paul B. Stephan, *A Becoming Modesty – U.S. Litigation in the Mirror of International Law*, 52 DEPAUL L. REV. 627 (2002).

Disinterestedness requires only that the decisionmakers will not allow extraneous considerations such as bribes or threats to influence their decision once they acquire jurisdiction over a dispute.

Third, we assume that direct sanctions imposed by a tribunal have greater salience than the reputational costs that an adjudicatory body can inflict simply by declaring one side in the wrong. We recognize that regimes might balance reputational and financial interests differently than would a private person, and in particular might respond to monetary sanctions in a more complex fashion.¹¹² But the power to apply these sanctions nevertheless implies an independent capacity to affect the behavior of parties to a dispute and augments the reputational effects of a tribunal's decision by reinforcing its consequential significance.

A wide range of international agreements specify a third-party dispute resolution mechanism, but not many of these meet our definition of coercive.¹¹³ Even the Molotov-Ribbentrop Pact, a zenith of international cynicism and futility masked as law, contained an arbitration clause. GATT and WTO dispute resolution in particular has evolved into a formalized, institutionally well developed and academically well studied system. As we have observed, however, what lies behind all these treaty-based dispute resolution mechanisms is a dynamic of state self-enforcement.¹¹⁴ Under these agreements, only states may seek dispute resolution, and the entity designated by the agreements as the dispute resolver – the WTO dispute resolution body in the case of the WTO, the International Court of Justice in the case of other instruments, or ad hoc arbitration as in the Molotov-Ribbentrop Pact – has only the authority to declare the legal rights and obligations of the parties, and not to impose any sanctions. None of these dispute settlement mechanisms, however formalized, employs coercion as a means of inducing compliance.

¹¹² See Daryl J. Levinson, *Making Government Pay: Markets, Politics, and the Allocation of Constitutional Costs*, 67 U. CHI. L. REV. 345 (2000).

¹¹³ For a survey of the practice that does not distinguish between coercive and noncoercive third-party adjudication, see Eric A. Posner & John. C. Yoo, note 92 *supra*.

¹¹⁴ Some might argue that the WTO dispute resolution process in particular presents a complex case, because WTO adjudicators have the capacity to authorize retaliatory actions by aggrieved states against countries that violate various Uruguay Round Agreements. But the WTO does nothing more than provide an opinion as to what retaliation would conform to its understanding of international law: The decision to retaliate remains that of the aggrieved state. At bottom, then, even sanctions generated by the WTO are reputational rather than direct and substantial. See authorities cited in note 15 *supra*.

But an increasing number of international agreements do invoke coercive sanctions through binding arbitration. Some involve ad hoc dispute resolution rather than permanent international tribunals; some do use such tribunals; and some employ independent domestic courts and enforcers. We discuss each below.

2. Agreements to Arbitrate

Unlike the traditional dispute resolution mechanism, this arbitration occurs at the invocation of private persons and can result in binding financial awards that a defaulting state must pay more or less automatically. Giving private persons the right to pursue a claim increases the strength of a promise and solves the agency cost problem that otherwise results because states do not have the same incentives to protect investment-exporting citizens as do the investors themselves. Accordingly, recognition of private standing strengthens the value of the host country's commitment to protect investor rights by increasing the expected cost of sanctions. The ultimate enforceability of an arbitral award through international attachment, as well as the existence of reliable mechanisms in some countries (such as the United States) for independent judicial enforcement, makes the outcome of these arbitrations material and direct.¹¹⁵

Historically, arbitration tribunals represented a post hoc response to a dispute. Once a controversy arose, the countries involved agreed to let injured persons submit their claims and provided some mechanism for domestic enforcement of the tribunal's decisions. An especially powerful example of this mechanism is the United States-Iran Claims Tribunal. This body, still at work more than two decades after its creation as a mechanism for resolving a hostage standoff between the two countries, exercises its authority to dispense awards out of a fund created from Iranian assets located in the United States. The existence of the several billions of dollars of Iranian property under U.S. control, as much as the formal agreement establishing the tribunal, explains that body's successful record as an arbiter of claims by U.S. and Iranian persons.¹¹⁶

¹¹⁵ For a brief discussion of the role of damages in investment treaties, see Andrew T. Guzman, note 8 *supra*, at 45-47. Guzman does not investigate damages as an enforcement mechanism under either European supranational law or through domestic courts.

¹¹⁶ For the background of the tribunal's creation, see *Dames & Moore v. Regan*, 453 U.S. 654 (1981). For a review of its

Since World War II, and particularly in the last two decades, capital-exporting states have entered into agreements that provide ex ante for coercive enforcement of disputes over the treatment of foreign investors. The details of these investor protection agreements vary, and our knowledge of practice under them is incomplete because of a (now waning) tradition of treating both the agreements and outcomes as confidential. In general, however, a host country commits in advance to resolve disputes stemming from a private investment through third-party arbitration, either invoking ICSID auspices or through some other institution. The investor can invoke the procedure without seeking any government's permission, and can translate their victories into monetary payment through the normal judicial process, which at least in some countries operates independently of the government.¹¹⁷

Chapter 11 of the North American Free Trade Agreement (NAFTA) is an especially powerful example of ex ante creation of a coercive third-party mechanism for resolving investment disputes.¹¹⁸ Its provisions codify certain substantive protections for investors and authorize complainants to invoke the auspices of third-party arbitration against the three state parties. To date aggrieved investors have brought 33 claims against Canada, Mexico and the United States.¹¹⁹ In each case, the World Bank's International Center for the Settlement of Investment Disputes (ICSID) established a panel of arbiters empowered to issue a monetary award. Prevailing claimants enforce their awards through the normal judgment collection procedures available in national courts.¹²⁰

practice, see THE IRAN-U.S. CLAIMS TRIBUNAL AND THE PROCESS OF INTERNATIONAL CLAIMS RESTITUTION (David D. Caron & John R. Crook eds. 2000); Charles N. Brower & Jason D. Brueschke, THE IRAN-UNITED STATES CLAIMS TRIBUNAL (1998).

¹¹⁷ For a list of those disputes arbitrated under ICSID auspices that have been made public, along with some opinions, see <http://www.worldbank.org/icsid/cases/cases.htm>. For a more general discussion of bilateral investment treaties, see Andreas F. Lowenfeld, *Investment Agreements and International Law*, 42 COLUM. J. TRANSNAT'L L. 123 (2003); Andrew T. Guzman, *Why LDCs Sign Treaties that Hurt Them: Explaining the Popularity of Bilateral Investment Treaties*, 38 VA. J. INT'L L. 639 (1998).

¹¹⁸ North American Free Trade Agreement, Dec. 17, 1992, Ch. 11, 32 I.L.M. 605, _____. For further discussion of this provision, see Frederick M. Abbott, *The Political Economy of NAFTA Chapter Eleven: Equality Before the Law and the Boundaries of North American Integration*, 23 HASTINGS INT'L & COMP. L. REV. 303 (2000); Charles H. Brower II, *Investor-State Disputes under NAFTA: The Empire Strikes Back*, 40 COLUM. J. TRANSNAT'L L. 43 (2001); Todd Weiler, *NAFTA Article 1105 and the Principles of International Economic Law*, 42 COLUM. J. TRANSNAT'L L. 35 (2003).

¹¹⁹ As of November 15, 2003, 9 claimants have initiated arbitration against Canada, 14 against Mexico, and 10 against the United States. See <http://www.naftaclaims.com>.

¹²⁰ For further discussion, see Roger P. Alford, *Federal Courts, International Tribunals, and the Continuum of Deference*, 43 VA. J. INT'L L. 675 (2003); Guillermo Aguilar Alvarez & William W. Park, *The New Face of Investment Arbitration: NAFTA Chapter 11*, 28 YALE J. INT'L L. 365 (2003); Charles H. Brower, II, *Structure, Legitimacy and NAFTA's Investment Chapter*, 36 VAND. J. TRANSNAT'L L. 37 (2003); Paul B. Stephan, *Redistributive Litigation – Judicial Innovation, Private Expectations and the Shadow of International Law*, 88 VA L. REV. 789 (2002).

NAFTA also contains another, less typical instance of coercive third-party arbitration of an international agreement. Its Chapter 19 allows private persons to challenge government decisions to impose antidumping and countervailing duties. Special arbitration panels substitute for the judicial review of administrative decisionmaking that normally would take place. These panels do not have the authority to impose coercive measures against a state, but each of the NAFTA parties has enacted a domestic law making the government's collection of antidumping and countervailing duties contingent on its successful conclusion of the Chapter 19 process.¹²¹

One might argue that the types of arbitration arrangements we describe here are not substantially different from the state-to-state dispute settlement procedures exemplified by the WTO system. In either case, a state might refuse to honor its obligations. In the instance of WTO dispute resolution, a state might not bring itself into conformity with what the WTO dispute settlement body specifies as its obligations, taking its chances with retaliation by other WTO members. In the case of arbitration, a state might not honor an award, either by directing a court to disregard the arbitral outcome or by enacting a law that blocks payment. This argument, however, assumes that inducing a court to set aside an arbitral award or the legislature to block payment entails only insignificant costs. As to authoritarian regimes operating without effective separation of powers, the assumption may be valid. But for many states, especially those in the rich world, judicial and parliamentary independence are relatively stable and robust features of the domestic government. For regimes managing the foreign policy of these states, a refusal to honor a third-party arbitral award entails domestic as well as external costs. From the perspective of such a regime, the decision of a third-party arbiter may function essentially the same as a command of a domestic court.

Why do states enter into these commitments? The rational-reciprocity approach provides an explanation. First, these agreements have at their core complex, relation-specific investments, where the welfare gains from making credible promises are significant. Absent protection from host government overreaching, investors either will make lesser investments or demand greater compensation from their hosts. Second, the commitments contained in these agreements refer to

¹²¹ See Curtis A. Bradley, note 17 *supra*, at 1773-75.

conditions that are likely to be verifiable and they do not contain implicit reservations resting on nonverifiable conditions. The events that constitute compensable violations of investment protection agreements involve public acts with reasonably determinable economic consequences. And allowing foreign investors to develop an oil field or build a pipeline is much less likely to threaten a regime's security than is losing an arms race or allowing imports to destroy an entrenched local industry. Third, the persons to whom the host state's promises are directed – foreign investors – have an intense and focused interest in having the obligation honored, while their home state, which in form is the promisee, may have a more complicated set of incentives and might sell out the investors in pursuit of some other objective. Coercive enforcement at the behest of the investors, without mediation by the home state, thus solves an agency cost problem. Finally, the prospect of externally coerced compensation responds to the change-of-regime end-game problem that historically has undermined self-enforcing commitments to foreign investors.¹²²

3. International Tribunals and Embedded Dispute Resolution

In terms of substantive obligation, the basic economic commitments that make up the Treaty of Rome, the constitutional basis of the European Community, and the norms embraced in the European Convention for the Protection of Human Rights and Fundamental Freedoms, which extends to the European Community members and twenty of their neighbors, do not differ greatly from those contained in more broadly based multilateral agreements such as the Uruguay Round Agreements and the various UN humanrights conventions.¹²³ Unlike the nearly universal multilateral regimes, however, the two European agreements each establishes an independent court with the authority to hear complaints brought by private persons and to levy fines on states that default on their commitments. Why have the Europeans departed from the self-enforcing mechanisms used elsewhere?

¹²² The existence of this enforcement mechanism suggests it provides some value, but not necessarily that, standing alone, it is optimal. For evidence that domestic judicial enforcement of claims against the state has a salient, and perhaps greater, role in encouraging investment, see authorities cite in note 47 *supra*.

¹²³ See THE EU, THE WTO, AND THE NAFTA: TOWARDS A COMMON LAW OF INTERNATIONAL TRADE (Joseph H.H. Weiler ed. 2000); Joel P. Trachtman, *The Theory of the Firm and the Theory of the International Economic Organization: Toward Comparative Institutional Analysis*, 17 NW. J. INT'L L. & BUS. 470 (1996-97).

We believe the answer involves the particularly embedded nature of the courts within the European commitments to elimination of trade barriers and the protection of fundamental personal rights. In Europe, these policies are not ends in themselves, but rather means of realizing a broader project. The Treaty of Rome and the European Convention and the courts they establish are part of an integrated and ambitious effort to replace the national structures that proved so unsatisfactory for much of the twentieth century with a European-wide federal state.

To appreciate this point, some historical perspective is necessary. The identity of states is contingent and in flux, even if not to the same degree as regimes. In 1783 there existed thirteen independent, albeit cooperating, sovereignties along the Atlantic coast of North America; by 1789 the cooperating states had become one. In the south of Europe, many principalities, kingdoms and the like became the Italian state, and in the north a single German Reich, during the middle of the nineteenth century. We recently have seen state decompositions (the Soviet Union, Yugoslavia and Indonesia). Projects to integrate Latin America, portions of the former Soviet empire, and various parts of the Pacific Rim are under way, although none has gone nearly as far as the European experiment.¹²⁴ The phenomenon is sufficient frequent, however, for us to regard the current experiment in European state building as representative of a potentially broader phenomenon.

Consider the specifics of the European project. In Europe, fifteen (as of 2004, twenty-five) nominally independent states have transferred significant policymaking and bureaucratic discretion to the organs of the European Community (the EC, the principal component of the EU), and have proposed giving the EU a broader role in framing a common foreign and defense policy on their behalf. In particular, the European Court of Justice generates a considerable body of interpretive law, which the courts of the members of the EU regard as having direct effect in their countries. Meanwhile, the 45 members of the Council of Europe (which includes all the present and prospective EU members) have delegated to special purpose institutions substantial authority to police their human rights performance. The Council of Europe's judicial body, the European Court of Human Rights, has the authority to hear suits by private persons and to fine states that have transgressed the European

¹²⁴For one somewhat visionary discussion of regional integration by a respected academic, see Lester Thurow, *HEAD TO HEAD: THE COMING ECONOMIC BATTLE AMONG JAPAN, EUROPE, AND AMERICA* (1992).

Convention on Human Rights; some of the member states also have incorporated this body of law directly into their legal systems. The substantive human rights law applied by the European Court comprises not only dignitary interests, but also economic rights such a protection against arbitrary seizures of property.

Although the process of building European institutions rests ultimately on a series of treaties, it also seems clear that these agreements constitute more than conventional international law.¹²⁵ What interests us about these two bodies of law is the way they enforce the international commitments that lie at the heart of the European project. It seems clear that the edicts of the two courts, as well as the legislation generated by the EU's Council and Commission, have more authority than conventional international arbitration as described above. The fundamental distinction lies in the embeddedness of the EU and Council of Europe judicial institutions. First, they come attached to a complex of commitments, from border administration to regulatory harmonization to criminal justice standards. A state cannot abandon these commitments selectively, but rather must lose the benefits of all if it wants to avoid the costs of any. Second, the commitments fit into an even larger process involving the gradual diminution of European states and the construction of something like a single European state. The outcome of this process remains uncertain and contested, but costs of exclusion from the broader project serve as yet another deterrent to defaulting on the agreements. At the same time, like conventional international arbitration, the two courts have the authority to levy fines and dispose of valuable licenses and other administrative "new property" derived from Community law.¹²⁶

We regard the fact of embeddedness and the latent threat of exclusion from a wide range of existing and potentially broader future institutional arrangements, as important additional coercive elements in both EU and Council of Europe Law. They suggest a blurring of domestic and international

¹²⁵ Some even have proposed creating a new category of "supranational law" to accommodate the European project. We regard categorization as not as important as an appreciation of context. Because the effort to create a European state both has great potential and remains highly contingent, particular instances of the construction of European law (whether EU or European Convention) are fraught with implications for the general enterprise.

¹²⁶ The Community's Court of Justice in particular seems virtually identical to a national court, given that its fines can be paid out of Community funds earmarked for particular member states. Unlike arbitration, no second enforcement process through national courts is necessary.

law, as the potential for state coercion seems practically indistinguishable from national law. For the participants in the process of European unification, it surely is important, indeed central.

The rational-reciprocity approach offers some explanations about the decision to build coercive enforcement into the European project. The ultimate goal of a unified Europe constitutes the investment which demands credible promissory enforcement. The creation of a European state presents the prospect of significant economic, political, and cultural benefits, but gradual surrender of national sovereignty in favor of European organs puts the regimes in each state at risk of opportunism by the others and the European organs themselves. Progress towards unification presents a complex problem of reciprocal performance with many opportunities for confusion between defection and justified retaliation. Expansion of the scope of the project, especially salient in a year when the number of states in the EU will increase by two-thirds, precludes preliminary agreements as a means of building trust.¹²⁷ At the same time, the vesting of lawmaking power in the EC Commission and Council reduces problems of verifiability by providing a political and bureaucratic process to specify challenging issues of performance measurement in terms that judicial bodies can verify and implement.

The arguments extend to the European human rights regime, even though in the abstract human rights commitments normally do not involve reciprocity or investment.¹²⁸ Although no formal linkage exists between the two European regimes, they function as strong complements. In particular, the EU has required adherence to the human rights regime as a precondition of candidacy for membership. The human rights commitments in the European Conventions, especially those directed at building political openness, accountability and cultural tolerance, are seen as disarming some of the historical pathologies that could make economic dislocations more difficult to tolerate. Those commitments directed at judicial independence also expand the scale of independent coercive enforcement, as discussed in the next section. The human rights commitments thus can be seen not just as an end in themselves, but rather as a means of bolstering the investments contemplated by the European state-building project.¹²⁹

¹²⁷ See Francesca Bignami, note 23 *supra*.

¹²⁸ See note 107 *supra*.

¹²⁹ We note that the proposed international bankruptcy court of the International Monetary Fund, if it had come into being, would

4. *Domestic Enforcement*

In recent years, considerable attention has focused on the use of domestic legal institutions, and particularly domestic courts, to enforce international agreements. The European project, for example, has relied not only on international judicial institutions such as the European Court of Justice and the European Court of Human Rights, but on commitments by states to allow their domestic courts to enforce the obligations contained in the Treaty of Rome and the European Convention. In the case of the Treaty, the Court of Justice early on determined both that domestic courts had the obligation of applying Community law even when in conflict with national legislation, and that it has the final say in determining whether domestic courts were carrying out this mission.¹³⁰ In the case of the Convention, states have had greater freedom as to how to use their domestic courts, but the United Kingdom's Human Rights Act instructing domestic courts to treat the decisions of the Human Rights Court as authority is an important example of how extensive domestic enforcement can be.¹³¹

Domestic enforcement is not limited to the Treaty of Rome and the European Human Rights Convention. Traditionally many agreements, especially in private law, assume direct effect. The Warsaw Convention, which governs contracts by international air carriers, overrides local contract and tort law. The Hague Rules, which states have implemented by enacting statutes, similarly governs contracts by international sea carriers.¹³² The International Monetary Fund's Articles of Agreement,

have conformed to our description of an embedded tribunal. For discussion of the proposal, see William W. Bratton & G. Mitu Gulati, *Sovereign Debt Reform and the Best Interest of Creditors* (Georgetown Law and Economics Research Paper No. 387880, 2003); Lee J. Buchheit & G. Mitu Gulati, *Sovereign Bonds and the Collective Will*, 51 EMORY L.J. 1317 (2002); Michael G. Hilgers, *Debtor-States and an International Bankruptcy Court: The IMF Creditor Problem*, 4 CHI. J. INT'L L. 257 (2003). The tribunal would have had power to order stays of domestic legal proceedings and would have been embedded in the IMF, an institution with enormous influence within the international financial system. But, as the proposal seems to have gathered no significant support, we will not discuss it further.

¹³⁰ *Van Gend en Loos v. Nederlandse Administratie der Belastingen* (Case 26/62), [1963] E.C.R. 3 (Treaty of Rome provisions create rights that individuals can directly enforce in domestic courts); *Costa v. ENEL* (Case 6/64), [1964] E.C.R. 1141 (European Court of Justice has final authority to review application of Treaty by domestic court).

¹³¹ See Human Rights Act, 1998, c. 42 (Eng.). For fuller discussion, see Gordon Anthony, *UK PUBLIC LAW & EUROPEAN LAW: THE DYNAMICS OF LEGAL INTEGRATION* (2002).

¹³² For a review of the Warsaw Convention and the Hague Rules, see Paul B. Stephan, *The Futility of Unification and Harmonization in International Commercial Law*, note 74 *supra*.

adopted at the end of World War II, supersedes national law regarding the enforceability of contracts involving an exchange of currencies.¹³³ Other examples abound.

The significance of domestic enforcement of international agreements depends on the quality and independence of domestic courts.¹³⁴ In countries with a powerful judiciary that acts free of government control, the courts function as a significant coercive check on the capacity of the government to renegotiate or repudiate international agreements. In the United States, for example, courts enjoy a tradition of independence, possess the power to enjoin government officials as well as to award damages, and operate in a culture of innovative judicial lawmaking. Their decisions, including their choices about enforcing international agreements, remain subject to legislative override, but even the power of Congress and the President to adopt legislation overturning judicial decisions is limited to instances that, in the eyes of the judiciary, do not violate the Constitution.¹³⁵

¹³³ See Joseph Gold, EXCHANGE RATES IN INTERNATIONAL LAW AND ORGANIZATION (1988); Paul B. Stephan, Don Wallace, Jr. & Julie A. Roin, INTERNATIONAL BUSINESS AND ECONOMICS – LAW AND POLICY 314-28 (2nd ed. 1996).

¹³⁴ See Paul B. Stephan, *Revisiting the Incorporation Debate: The Role of Domestic Political Structure*, 31 VA. J. INT'L L. 417 (1991). For evidence that U.S. courts exercise independence in the interpretation of treaty provisions and do not necessarily defer to claims by the U.S. government, see, e.g., *Washington v. Washington State Commercial Passenger Fishing Vessel Ass'n*, 443 U.S. 658 (1979); *United States v. California*, 436 U.S. 32 (1978); *Menominee Tribe of Indians v. United States*, 391 U.S. 404 (1968); *McCulloch v. Sociedad Nacional de Marineros de Honduras*, 372 U.S. 10 (1963); *Moser v. United States*, 341 U.S. 41 (1951); *Black Diamond S.S. Corp. v. Robert Stewart & Sons*, 336 U.S. 386 (1949); *United States v. Fullard-Leo*, 331 U.S. 256 (1947); *Chew Heong v. United States*, 112 U.S. 536 (1884); *Xerox Corp. v. United States*, 41 F.3rd 647 (Fed. Cir. 1994); *United States v. Palestine Liberation Organization*, 695 F. Supp. 1456 (S.D.N.Y. 1988); cf. *Hampton v. Mow Sun Wong*, 426 U.S. 88 (1976) (preserving power of President to negotiate treaties insufficient justification for statutory default rule excluding noncitizens from civil service); *United States v. Guy W. Capps, Inc.*, 348 U.S. 296 (1955) (acquittal upheld on grounds of insufficient evidence, leaving unresolved issue of validity of U.S.-Canada trade agreement on which prosecution was based)

¹³⁵ For the argument that some legislation limiting the power of U.S. courts to enforce international agreements may violate the Constitution, specifically the Supremacy Clause, see Louis Henkin, *U.S. Ratification of Human Rights Treaties: The Ghost of Senator Bricker*, 89 AM. J. INT'L L. 341, 346 (1995); Carlos Manuel Vásquez, *The Four Doctrines of Self-Executing Treaties*, 89 AM. J. INT'L L. 695, 716-17 (1995). For a response, see Curtis A. Bradley & Jack L. Goldsmith, *Treaties, Human Rights, and Conditional Consent*, 149 U. PA. L. REV. 399, 442-56 (2000). Courts also have the power to give surprising glosses to old and neglected statutes to justify assertions of power. The present authority of the U.S. courts to create and enforce the customary international law of human rights, for example, rests on a 1980 rediscovery of a 1789 statute. The validity of this move currently is before the Supreme Court. *Alvarez-Machain v. United States*, 331 F.3rd 664 (9th Cir.), cert. granted sub nom. *Sosa v. Alvarez-Machain*, 124 S. Ct. 807 (2003).

Some international agreements expressly disavow direct domestic enforcement.¹³⁶ Others expressly call for domestic judicial enforcement, either by their own terms¹³⁷ or through implementing domestic laws.¹³⁸ Many of the agreements that expressly call for direct domestic enforcement involve private commercial transactions where the parties typically are strangers to each other and not likely to engage in repeat play.¹³⁹ Typically the potential losses parties face have a high variance. As a result, self enforcement is unlikely to be optimal.

There remain, however, a significant number of agreements where the intention of the parties is unclear. Those situations force national lawmakers, and in particular domestic courts, to articulate interpretive strategies and to construct default rules to determine the domestic effect of agreements that fail to address the issue¹⁴⁰ Courts and commentators tend to approach defaults from one of two directions. The proactive approach, as we shall term it, presumes the efficacy of domestic judicial enforcement of international agreements and puts the burden of proof on those arguing against intervention. It focuses only on capacity issues such as verifiability, asking whether an agreement contains “sufficiently determinate standards” on which courts can base their actions.¹⁴¹ The bargaining approach, as we shall call it, instead asks whether the agreement bargained for judicial enforcement. Implicit in the latter is a recognition that agreements might contain a mix of verifiable and nonverifiable conditions representing offsetting concessions, and that enforcement of only some might upset the parties’ expectations and skew performance away from observable but nonverifiable conditions.¹⁴²

¹³⁶ See *Portuguese Republic v. Council* (Case C-149/96), [1999] E.C.R. I-8395 (analyzing deliberate decision of Uruguay Round parties not to have agreements directly enforceable).

¹³⁷ Convention for the Unification of Certain Rules Relating to International Transportation by Air, Oct. 12, 1929, 49 Stat. 3000, T.S. No. 876, 137 L.N.T.S. 11; Articles of Agreement of the International Monetary Fund, July 22, 1944, 60 Stat. 1401, 2 U.N.T.S. 39, *as amended*, 20 U.S.T. 2775; 29 U.S.T. 2203, T.I.A.S. NO. 11898.

¹³⁸ Convention for the Unification of Certain Rules of Law Relating to Bills of Lading, Aug. 25, 1924, 51 Stat. 233, 120 L.N.T.S. 155, *implemented by* Carriage of Goods at Sea Act of 1936, 46 U.S.C. app. §§ 1300-1315 (1994).

¹³⁹ For a review of international agreements involving private commercial transactions, see Paul B. Stephan, *The Futility of Unification and Harmonization in International Commercial Law*, note 74 *supra*.

¹⁴⁰ The controversial instances involve human rights treaties, which the United States invariably ratifies subject to a statement that the domestic courts shall not enforce the obligations contained therein. For a debate on whether the treaties permit this limitation, compare Curtis A. Bradley & Jack L. Goldsmith, note 135 *supra*, with Ryan Goodman, *Human Rights Treaties, Invalid Reservations, and State Consent*, 96 AM. J. INT’L L. 531 (2002).

¹⁴¹ Carlos Manuel Vásquez, note 135 *supra*, at 713-15 (citing cases).

¹⁴² See text at notes 49-57 *supra*. See also Andrew T. Guzman, note 8 *supra*, at 41-45.

Two recent cases illustrate how these approaches work. In *Kappus v. Commissioner*,¹⁴³ the court had to decide whether U.S. taxpayer resident in Canada could offset all U.S. income tax liability with Canadian taxes paid on the same income, or instead would have to pay some portion of the U.S. taxes. Article XXIV of the 1980 Tax Treaty between Canada and the United States seemed to require a full credit.¹⁴⁴ But Section 59(a)(2) of the Internal Revenue Code, a provision enacted after the Treaty went into effect, capped the available U.S. credit at 90 percent. To complicate matters further, the United States and Canada had entered into additional protocols to the Treaty in 1995 and 1997, after the enactment of Section 59(a)(2), but neither protocol addressed Article XXIV of the original Treaty.

A court might unravel this conflict by adopting one of four possible defaults: (1) a domestic court could enforce rights under an international agreement unless domestic legislation explicitly repudiates the agreement;¹⁴⁵ (2) a domestic court could treat the last enactment as controlling, but would interpret “enactment” generously so as to treat treaty amendments as ratifications of the entire treaty; (3) as the D.C. Circuit held in *Kappus*, a domestic court could treat the last enactment as binding, and interpret “enactment” as limited to the language in dispute; or (4) a domestic court could never allow a treaty provision to supersede a statute without express legislative authorization to do so.¹⁴⁶ This progression illustrates a transition from the proactive to the bargaining approach, as each successive move reflects an incremental reluctance on the part of the judiciary to hold the government to its treaty commitment.¹⁴⁷

¹⁴³ 337 F.3rd 1053 (D.C. Cir. 2003).

¹⁴⁴ Convention Between Canada and the United States of America with Respect to Taxes on Income and on Capital, Art. XXIV(4)(b), Sept. 26, 1980, U.S.-Can., T.I.A.S. No. 11,087.

¹⁴⁵ A variant of this approach would call on courts to interpret legislation so as not to violate an international agreement. See *Murray v. The Charming Betsy*, 6 U.S. (2 Cranch) 64 (1804). For an extreme example of this interpretive move, see *United States v. Palestine Liberation Organization*, 695 F. Supp. 1456 (S.D.N.Y. 1988) (interpreting statute intended to bar establishment of a PLO observer mission at the United Nations as limited by earlier treaty regarding the UN headquarters); Harold H. Koh, *Transnational Public Law Litigation*, 100 YALE L.J. 2347, 2373-75 (1991) (praising decision as exemplary interpretive strategy).

¹⁴⁶ Cf. INTERNAL REVENUE CODE § 7852(d)(1) (specifying relationship between tax treaties and Internal Revenue Code).

¹⁴⁷ On the analogous issue of the efficacy of the Uniform Commercial Code’s generous contract-creation rules, see Robert E. Scott, *The Rise and Fall of Article 2*, note 16 *supra*.

In *Chubb & Sons, Inc. v. Asiana Airlines*,¹⁴⁸ the court had to decide what law governed a dispute over the misdelivery of goods shipped by international air carrier. The flight originated in South Korea and ended in the United States, presumptively bringing the transaction under the Warsaw Convention. A complicating factor, however, was the existence of several versions of that multilateral Convention, one of which the United States had joined in 1934 and a later version of which Korea had embraced in 1967. The court had to decide whether an agreement existed between the two countries as a result of their acceptance of overlapping but distinct obligations.¹⁴⁹

A proactive approach would invoke a pro-agreement default on the theory that judicial enforcement of some obligations is preferable to no enforcement. One could find a template in § 2-207 of the Uniform Commercial Code, which allows courts to make a judicially enforceable contract in cases where an offer and acceptance do not match.¹⁵⁰ Alternately, as the Ninth Circuit did in *Chubb*, a court might require something closer to the common law's mirror-image rule to limit treaty relations to instances where states had assumed identical obligations.¹⁵¹ This default implements the idea that parties might regard the total bargain as motivating performance and correspondingly consider partial enforcement as an unwanted outcome.

What we find interesting about both cases is the implicit relationship between judicial coercion and the international agreement. In both cases, the degree of domestic enforcement of an international agreement turns on which template the courts will use. *Kappus* implicates the level of domestic lawmaking required to free a court of its responsibility to enforce an international agreement, while *Chubb* considers the level of international interaction required for enforceable obligations to arise.

¹⁴⁸ 214 F.3rd 301(2nd Cir. 2000) *cert. denied*, 533 U.S. 928 (2001).

¹⁴⁹ The existence of treaty relations determined both the subject matter jurisdiction of the federal court that heard the case and the substantive rules for determining liability and damages.

¹⁵⁰ UCC § 2-207(3). The trial court in *Chubb* did precisely this, although it based its decision on an interpretation of the Warsaw Convention and the Vienna Convention on the Law of Treaties, May 23, 1969, 1155 U.N.T.S. 331. *Chubb & Son, Inc. v. Asiana Airlines* 1998 WL 647185 (S.D.N.Y.1998). The United States was (and is) not a party to the latter instrument, although the Department of State maintains that some portions of it have the force of customary international law.

¹⁵¹ The Ninth Circuit based its decision on both the general desirability of such a default and its interpretation of Article 40(5) of the Vienna Convention.

The former involves exit, the latter entry, but both implicate the activity level of an independent and coercive judiciary.

Focusing only on U.S. practice, we can see that decisions about when to use domestic courts to enforce international commitments, and when not to, is broadly consistent with the rational-reciprocity approach. The agreements most likely to generate rights and duties that domestic courts will enforce involve complex, private commercial transactions, where the benefits from enhancing the credibility of reciprocal promises seem greatest and difficulties with verification of salient conditions seem manageable. The agreements that the United States most often has directed its courts not to implement involve open-ended human rights commitments, which, on the one hand, have great moral clarity and, on the other hand, present vexing verifiability issues due to the nonspecific nature of the obligations. As *Kappus* and *Chubb* illustrate, the courts also use secondary doctrines that minimize uncertainty about judicial reinterpretation of the scope of commitments set by the political branches.

At the same time, U.S. doctrine regarding domestic enforcement of obligations contained in international agreements is far from settled, and academic opinion largely favors a greatly expanded role for U.S. courts, a role that the rational-reciprocity approach suggests might well be counterproductive. We do not assert that we have developed a powerful positive theory, at least as applied to U.S. practice. Rather, we argue that the contract theory approach provides an explanation of at least the broad outlines of U.S. approaches to domestic enforcement of international law and may provide a basis for a normative critique of academic calls to change that practice.

C. Summary

In designing international agreements, parties can choose from an array of enforcement techniques, including progressively more coercive measures wielded by independent tribunals. Self-enforcement explains much about these agreements, but it is incorrect to limit analysis to this mechanism. Moreover, the framers of international agreements must take into account not just those aspects of enforcement to which they advert, but also the background assumptions about coercive enforcement that independent institutions, first among them domestic courts, will employ. By adjusting these

assumptions, independent institutions can alter the enforcement structures that attend international agreements.

This section has described the enforcement techniques conventionally applied to international agreements, and in particular has noted the connection between interpretive techniques and enforcement. We have observed some support in international practice for the claim that choices about enforcement, and the interpretive strategies that affect enforcement, reflect the theoretical framework developed in the first part of this paper. In the next part we consider in greater detail the suggestion made in the first part that some enforcement techniques work at cross purposes with others and explore the implications of this suggestion for optimal enforcement strategy.

III. SELF-ENFORCEMENT AND COERCION AS COMPLEMENTS AND RIVALS IN INTERNATIONAL AGREEMENTS: THEORY AND EVIDENCE

The previous part makes clear that international agreements, no less than private contracts, rely on some mixture of self enforcement and public coercion to promote compliance. Does the similarity between international agreements and private contracts run deeper? As we observed in Part I, a substantial literature explores the complex relationship between self-enforcement and coercion in private contracts, and argues that, across a significant range of transactions, coercion diminishes the value generated by the contractual relationship.¹⁵² Here we consider the possibility that coercive enforcement of international agreements may undermine self enforcement, and thus for some set of obligations detract from optimal enforcement.

A. Optimal Enforcement Structure in International Agreements

A naive view of international law might regard more enforcement as always better. Some of the international law literature seems to rest on a view that international agreements are a product that has positive returns to both scale and scope.¹⁵³ A little reflection will suggest the implausibility of this

¹⁵² See notes 59-67 *supra* and accompanying text.

¹⁵³ We do not find bald claims in the literature, but we do note a tendency to attack any limitations on domestic enforcement that

claim. International agreements as much as private contracts involve predictions about future states in the face of uncertainty. When these predictions prove incorrect, rigid commitments make the parties worse off. Knowing this, parties rationally should specify the conditions under which the obligations contained in an agreement should not apply, rather than insist on the necessity of the commitment in all possible future states.¹⁵⁴ Absent complete prescience, states rationally must balance commitment with flexibility by limiting both the scope of their commitments and the extent of the sanctions that a violation will trigger.

We concede that conditions sometime exist where network effects argue for expanding the scope and the scale of international agreements. An important example involves the conventions, customs and practices that underlie the assignment of internet domain names.¹⁵⁵ Our argument is only that most subjects of international agreements do not have this quality, and that a significant range of agreements entail diminishing marginal returns accompanied by increasing marginal enforcement costs. Optimizing the value of such agreements involves finding the point where the benefits of any increase in the scale or scope of the agreement equals the cost associated with the increase.

emerge. For influential examples of this tendency, see, *e.g.*, Jordan J. Paust, *INTERNATIONAL LAW AS LAW OF THE UNITED STATES* (1996); Harold H. Koh, note 145 *supra*; Louis Henkin, note 135 *supra*; Louis Henkin, *International Law as Law in the United States*, 82 *MICH. L. REV.* 1555 (1984); Jules Lobel, *The Limits of Constitutional Power: Conflicts Between Foreign Policy and International Law*, 71 *VA. L. REV.* 1071 (1985).

¹⁵⁴ Recall that the canonical contracting problem is to ensure both efficient *ex ante* investment and efficient trade. To do so, parties must balance the benefits from credible commitments against the benefits of flexibility in adjusting to realized states of the world. To be sure, one method of ensuring flexibility is renegotiation. If the parties anticipate that the uncertainties they face at the time of contracting ultimately will be resolved, they would logically prefer to renegotiate their initial agreement once the future is known. From this perspective, the initial agreement is only a means by which the parties set the entitlements that will be traded *ex post*. If the parties are symmetrically informed, bargaining theory teaches that they will renegotiate to the efficient result. For example, the parties could agree to a fixed price contract that paid the promisor to invest efficiently or conditioned the price on the realized value of the investment. In such a world, neither party can exploit the specific investments of the other strategically. The problem, however, is that if information is asymmetric (that is, costs and valuations are not observable) either party can exploit the sunk cost investment of the other. Under these conditions, in order to insure both efficient investment and efficient trade, parties must be able to precommit not to renegotiate. See Alan Schwartz & Robert E. Scott, *Contract Theory*, note 16 *supra*, at 611-614. In either state of the world, the parties will renegotiate (or not) against the background of the enforcement mechanisms that they design. Only if the relevant values are observable but not verifiable will they have reason to select between self-enforcement and coercive enforcement based upon the prospect of renegotiation.

¹⁵⁵ We note that even here the international community has not created a monolithic legal system. An independent tribunal, working with the consent of the administrators of the internet, has the authority to resolve disputes over the ownership of domain names, but national authorities have no obligation to respect the determinations of these tribunals. The speed and low cost of the international process, in contrast to domestic litigation, ensures that conflicts seldom arise. When they do, however, domestic courts are free to disregard the tribunal's decision. For an example, see *Barcelona.Com, Inc. v. Excelentísimo Ayuntamiento de Barcelona*, 330 F.3rd 617 (4th Cir. 2003).

This straightforward analysis can be extended to enforcement. Rather than specifying that an obligation will cease to exist if certain circumstances arise, an agreement instead may present the parties with a series of options. For any specified future state, a party may choose between compliance and undergoing some sanction. An optimal level of enforcement requires setting the sanction so that the marginal costs of imposing the sanction equal the marginal benefits from compliance. And the determination of a sanction in turn implicates the choice of enforcement mechanism, and in particular the mix of self enforcement and third-party coercion.

B. Complementarity, Rivalry and States

In Part I we reviewed evidence indicating that coercive sanctions might function as a rival to self enforcement. As we noted, the experimental data and the conjectures about their significance relate to individual behavior. We again must ask whether political elites exhibit similar tendencies. Does the ex ante announcement of the prospect of coercive sanctions undermine self enforcement of agreements to which states are parties?

Some similarities seem clear. We expect political elites to care about reputation generally, and in particular about being known as submissive. Indeed, we expect foreign relations decisionmakers to be more averse to this perception than average individuals, both because of qualities reinforced by the process of elite selection and because reputational effects would involve domestic as well as international capabilities.

Membership in a governing group, we surmise, entails surviving a selective process in which perceptions of weakness are not rewarded. Even where advancement depends on convincing superiors of servility, the supplicant still must present a credible image of mastery of his inferiors. We therefore would expect that people who survive the process and become decisionmakers generally would be willing to incur substantial costs to punish those who treat them unfairly.

Moreover, the risks of appearing weak are compounded by their consequences not just for future international interactions, but for place within the domestic power hierarchy. We surmise that

perceptions of weakness are generalized, rather than particular, in that observers usually do not distinguish among contexts. Thus, an international relations decisionmaker plausibly can believe that appearing coerced will not only encourage future demands by other states, but induce domestic rivals to challenge his position. It seems paradoxical to expect that a regime too weak to resist coercion by a third party nonetheless can reliably compel its population to comply with the third party's command. In the case of Europe, for example, governments on occasion have found themselves incapable of squelching civil defiance of the edicts of the Brussels court.¹⁵⁶

We are less sure that institutionalizing coercive sanctions for international agreements induces free riding on the part of the elites of regimes that might otherwise inflict costs on noncompliers.¹⁵⁷ Consider the UN Security Council's supposed monopoly under the UN Charter regarding authorization of the use of force to punish Charter violations. The example is not perfect because the Security Council monopoly itself is only self-enforcing, in the sense that no third-party enforcer exists to punish states that encroach on this monopoly. Yet we can make two observations: Countries do use force to punish violators of international law without Security Council permission, as U.S.-led coalitions did in Serbia and Iraq, and it is unclear whether decisions not to use force in the face of clear Charter violations reflect free riding or a belief that doing so without Security Council approval entails additional costs.

Some ambiguous evidence of free riding comes from the field of investor protection. Before the advent of third-party arbitration, disputes over injuries to foreign investors typically involved state-to-state intervention. In extreme cases, a great power would threaten to use military force to focus the attention of the injuring state on the investors' complaint. The rise of the arbitration mechanism over the last forty years coincides with the end of gunboat diplomacy, and more generally in a reduction of efforts by states to expend resources unilaterally on behalf of injured investors.

¹⁵⁶ See, e.g., *Commission v. Greece* (Case 68/880, [1989] E.C.R. 2965 (complicity of domestic officials in scheme to evade EC customs)); *Commission v. France* (Case C-265/95), [1997] E.C.R. I-6959 (inability of French government to manage civil strife designed to obstruct free movement of agricultural products required under EC law). [other examples from Marshall court]

¹⁵⁷ See notes 84-85 *supra* and accompanying text.

We do not want to make too much of this example, because of two confounding factors. First, over the same forty-year period the International Monetary Fund and the World Bank have acquired increased influence over the policies of many states and have used threats to withhold financing as an incentive to induce a reduction of attacks on foreign investors. To some extent, these institutions have become enforcement agents for the capital-exporting world.

Second, the arbitration mechanism rests on treaties that create formally symmetrical obligations. Because arbiters operate free of state control, to a certain extent they can redefine the scope of protection provided by these treaties. The capital exporters thus have lost the ability to fix the scope of their obligations, and have found themselves faced with claims that their own actions violate these treaties.¹⁵⁸ Now that their arguments can be used against them, they may have become more cautious about supporting claims for protection. We thus concede that the evidence that coercive enforcement of international agreement creates a free riding problem is mixed at best.

Finally, do some states, like some individuals, find coercion inconsistent with a preference for reciprocal fairness? Both general arguments and some anecdotal evidence suggest an affirmative answer. Introducing coercive commitments to international agreements makes some states less willing to honor the agreements or to interpret their obligations liberally. With states as with individuals, coercion can undermine the desire to cooperate rather than reinforce it.

Consider first the observation that some states prefer coercion because they regard international agreements as a precommitment strategy. These countries seek to lock in currently tolerated policies against shifting future domestic majorities.¹⁵⁹ The argument assumes that the effects of an agreement can be controversial and potentially unpopular, but will not generate such strong opposition as to lead a country to withdraw from the coercive system. The critical premise is that regimes do not trust their future selves to honor commitments, and that others should not trust them either.

¹⁵⁸ See Paul B. Stephan, note 120 *supra*, at 830-39.

¹⁵⁹ See Ryan Goodman, note 140 *supra*, at 540-45 (discussing “lock-in” goal in treaties); Andrew Moravcsik, *The Origins of Human Rights Regimes: Democratic Delegation in Postwar Europe*, 54 INT’L ORG. 217 (2000) (documenting effect).

Reciprocal fairness, however, depends fundamentally on trust in the propensity of counterparties to behave similarly. Signals that this trust is unwarranted should compel states to alter their premises, and in particular should induce precautions against defection. These precautions in turn further indicate lack of trust. To the extent, then, that explicit commitments to coercive enforcement rest on doubts about future capacity to comply, they also convey information that undermines self enforcement.

U.S. practice provides direct evidence of the tension between coercion and cooperation in the context of international agreements. One instance involves the UN Charter. After the founding of the United Nations and its issuance of the Universal Declaration of Human Rights, civil rights groups in the United States began invoking these instruments as a ground for attacking the government's role in enforcing racial segregation.¹⁶⁰ A litigation strategy emerged in which opponents of restrictive laws invoked these international commitments to attack the status quo. A few progressive state courts embraced the argument, which rested on the premise that they had the authority to interpret and enforce the UN Charter, particularly with regard to discrimination.¹⁶¹ Greater reliance on international-agreement-based judicial decisions to dismantle American segregation seemed likely.

The United States responded to this pressure by seeking to eliminate the coercive force of its international commitments. During the early 1950s Senator John Bricker introduced versions of a constitutional amendment that would have barred treaties from having any effect on domestic law, absent implementing legislation. The effort won the endorsement of the American Bar Association and at one point came within a single vote of clearing the Senate. The Eisenhower Administration eventually fended off the movement to codify the nonenforceability of U.S. international commitments by representing that it would not enter into any international agreement that had domestic law reform as its purpose.¹⁶² The practice of conditioning U.S. adherence to agreements regarding domestic civil rights on the enactment of further domestic legislation has remained the norm since that time.¹⁶³

¹⁶⁰ See Mary L. Dudziak, *COLD WAR CIVIL RIGHTS: RACE AND THE IMAGE OF AMERICAN DEMOCRACY* 43-64 (2000).

¹⁶¹ *Sei Fujii v. State*, 217 P.2nd 481, 218 P.2nd 595 (Cal. Dist. Ct. App. 1950), *vacated*, 242 P.2nd 617 (Cal. 1952); *Kenji Namba v. McCourt*, 185 Or. 579, 204 P.2nd 569 (Or. 1949); *Perez v. Lippold*, 32 Cal.2nd 711, 198 P.2nd 17 (Cal. 1948).

¹⁶² See 32 DEP'T STATE BULL. 820 (1955) (statement of Secretary of State Dulles that United States would not join multilateral treaties pertaining to human rights); Curtis A. Bradley, *The Treaty Power and American Federalism, Part II*, 99 MICH. L. REV. 98, 122-23 (2000); Curtis A. Bradley & Jack L. Goldsmith, note 135 *supra*, at 413.

¹⁶³ Curtis A. Bradley & Jack L. Goldsmith, note 135 *supra*, at 410-23.

Critics of this episode in U.S. history emphasize its links to the defense of segregation, suggesting that only bigotry explains the opposition to coercive invocation of international law.¹⁶⁴ This connection is part of the story, but fails to do justice to the complex dynamic surrounding the use of international law in the battle against U.S. racial segregation. First, the controversy survived *Brown v. Board of Education*'s definitive attack on segregation: Secretary of State Dulles gave the final assurances that ended the push for constitutional action a year after *Brown* came down, and some version of the proposal remained under consideration in the Senate until 1957.¹⁶⁵ Second, and more importantly, the range of support for the Bricker Amendments far exceeded that for segregation. Some important part of the U.S. legal establishment separated the issues by supporting both the ending of segregation and some quarantine of international law from the U.S. legal system.

In hindsight, it appears that most progress on dismantling the legal basis of segregation (which at some level of abstraction could be described as a cooperative effort to comply with the obligations of the UN Charter) took place after Dulles reassured the Congress that the Administration would not use international law to pursue domestic reform. The Civil Rights Act of 1957, the sending of federal troops to Little Rock, the Kennedy Administrations' mobilization of the Civil Rights Division of the Justice Department, and the civil rights legislation of the mid-1960s all unfolded against a backdrop of no coercive international obligation. The domestic courts acquired an enlarged arsenal to use against racial discrimination, but none involved the UN Charter or any other international agreement.

A similar, if lower-profile episode involves the GATT. President Truman signed this agreement in 1947, relying on the authority of the Trade Agreements Act of 1934. Neither the 1934 Act nor the 1947 Agreement expressly addressed the issue of implementing legislation, but a handful of courts and slightly more commentators argued that a rule of direct effect should apply.¹⁶⁶ Congress responded by

¹⁶⁴ Louis Henkin, note 135 *supra*, at 348-49.

¹⁶⁵ Duane Tananbaum, THE BRICKER AMENDMENT CONTROVERSY—A TEST OF EISENHOWER'S POLITICAL LEADERSHIP 191-215 (1988). In the view of contemporaries, the Supreme Court decision that undermined the Amendment supporters was not *Brown*, but *Reid v. Covert*, 354 U.S. 1 (1957), which assuaged fears that a Senate-approved treaty might indirectly amend the Constitution. Duane Tananbaum, *supra*, at 211-14.

¹⁶⁶ Cases striking down legislation as inconsistent with GATT obligations include *Territory of Hawaii v. Ho*, 41 Haw. 565 (1957); *Baldwin-Lima-Hamilton Corp. v. Superior Court*, 208 Cal. App. 2nd 803, 25 Cal. Rptr. 798 (Dist Ct. App. 1962). For the commentators, see Ronald A. Brand, *The Status of the General Agreement on Tariffs and Trade in United States Domestic Law*, 26 STAN. J. INT'L L. 479 (1990); Robert E. Hudec, *The Legal Status of the GATT in the Domestic Law of the United*

inserting in each new trade bill language to the effect that it should not be seen as either endorsing or rejecting the proposition that the GATT had direct effect in U.S. law.¹⁶⁷ During the time that this remained an open question, the GATT process focused mostly on tariff reduction and the elimination of transparently discriminatory barriers to imports. The issue of direct effect of GATT for the most part remained an academic question of little practical effect.

Beginning with the Tokyo Round Agreements of 1979, however, the GATT regime began to extend to more subjects and to threaten a greater range of domestic policies. The Tokyo Agreements represented a turn towards combating disguised discrimination, which typically involves the appropriation of a valid regulatory objective as a (more or less) pretextual basis for barring imports. Efforts to discourage this practice present a greater risk of interfering with legitimate public policies. In response, some of the more problematic commitments contained express language barring coercive enforcement.¹⁶⁸

The further broadening of the scope of international economic agreements induced even greater commitments against coercive domestic enforcement. The 1993 NAFTA, which took shape during the Uruguay Round and in significant respects reflected it, provided the first occasion for Congress to state explicitly that the domestic courts could not engage in independent enforcement of the agreement.¹⁶⁹ The following year Congress applied exactly the same conditions to the Uruguay Round Agreements, which superseded as well as extended the 1947 GATT.¹⁷⁰

States, in *THE EUROPEAN COMMUNITY AND GATT* 187 (Meinhard Hilf, Francis G. Jacobs & Ernst-Ulrich Petersmann eds. 1986); John H. Jackson, *The General Agreement on Tariffs and Trade in United States Domestic Law*, 66 MICH. L. REV. 260 (1967); Note, *The United States Participation in the General Agreement on Tariffs and Trade*, 61 COLUM. L. REV. 505 (1961).

¹⁶⁷ Act of June 16, 1951, § 10, 65 Stat. 75, *codified at* 19 U.S.C. § 1366; Act of Aug. 7, 1953, § 103, 67 Stat. 472; Act of July 1, 1954, § 3, 68 Stat. 360; Trade Agreements Extension Act of 1955, 69 Stat. 162, *codified at* 19 U.S.C. § 1351(a)(1)(A); Act of Aug. 20, 1958, P.L. 85-686, § 10, 72 Stat. 680.

¹⁶⁸ See *Trojan Technologies, Inc. v. Commonwealth of Pennsylvania*, 916 F.2nd 903, 908 (3rd Cir. 1990), *cert. denied*, 501 U.S. 1212 (1991) (discussing 1979 Agreement on Government Procurement).

¹⁶⁹ North American Free Trade Agreement Implementation Act, § 102, *codified at* 29 U.S.C. § 3312. The Act does carve out a small role for judicial enforcement of the Agreement, but only in cases where the U.S. government (and not private parties or other governmental bodies) seeks to enforce the Agreement against itself or state or local governments. *Id.* § 102(b)(2), (c)(2).

¹⁷⁰ Uruguay Round Agreements Act of 1994 § 102, *codified at* 29 U.S.C. § 3512.

Neither the human rights nor the GATT case study proves that the United States displayed a more cooperative attitude toward potential partners in multilateral agreements because it was free of the threat of coercive enforcement by independent domestic courts. We recognize that other factors might explain both the periods of cooperation and the earlier conservatism about international commitments. We still find it interesting that, in these two disparate areas and at two different times, U.S. efforts to exhibit cooperative compliance with its international obligations coincided with strong measures to discourage coercive enforcement.

C. Creeping Coercion

Up to this point our analysis has assumed that coercive enforcement attached to an international agreement operates in an *ex ante* fashion on the basis of deliberate choices by the framers of the agreement. But, as U.S. experience with the UN Charter and the GATT illustrates, the prospect of coercive enforcement can arise after the fact or function as an uncertain background risk at the time of entering into an agreement. In the case of arbitration commitments, the scope of the obligation may be indeterminate, both as to what constitutes a protected investment and as to what constitutes an compensable encroachment.¹⁷¹ Both the economic obligations embedded in the EU and the human rights commitments embedded in the Council of Europe have changed in surprising ways over time

¹⁷¹ Consider some of the issues that have led to NAFTA Chapter 11 arbitration: *The Loewen Group, Inc. v. United States*, Final Award (Jun. 26, 2003) (egregious misconduct in civil trial leading to enormous damages manifestly a denial of justice subject to Chapter 11; no relief available because victim failed to seek appellate review); *S.D. Myers, Inc. v. Canada*, Final Award on Merits (Nov. 13, 2000) (government order banning export of PCBs from Canada for waste processing in United States violates obligations under Chapter 11); *Metalclad Corp. v. United Mexican States*, Final Award (Sep. 2, 2000) (denial of permit for hazardous waste landfill violates obligations under Chapter 11); *Ethyl Corp. v. Canada*, Award on Jurisdiction (Jun. 24, 1998) (statute banning fuel additive destroyed business and raises issue of Chapter 11 violation). For fuller discussion, see Charles H. Brower II, *Structure, Legitimacy, and NAFTA's Investment Chapter*, 36 *VAND. J. TRANSNAT'L L.* 37 (2003).

as a result of judicial creativity.¹⁷² And the process of internalization of international rules by domestic courts, based as it is on highly flexible doctrine and nonspecific criteria, undermines predictability.¹⁷³

If foreknowledge of coercive enforcement can interfere with self enforcement of an agreement, does it follow that substantial uncertainty also can reduce investments in self enforcement? We believe that the answer is yes. First, we observe that, by hypothesis, asymmetrical uncertainty is unlikely. Those characteristics of third-party enforcers that affect the likelihood of their future actions normally are public knowledge, or at least are known to specialists whose services are available symmetrically. As a result, parties ought to know that doubts about coercion will affect the behavior of their counterparties as well as their own. Both their own preferences for reciprocity and expectations about those of their counterparties should be shaped by the knowledge that at some future date an obligation may become subject to third-party enforcement.

Second, the logic of the arguments for the rivalrous relationship of self enforcement and coercive enforcement of international agreements extends to the case where coercion is possible rather than certain. The reputational effect of compliance will reflect doubts about whether the behavior will be seen as cooperative or submissive. The possibility of coercive enforcement, not just its certainty, should affect decisions whether to free ride rather than absorb the costs involved in sanctioning noncompliers. Finally, uncertainty about whether compliant behavior reflects a preference for cooperation or a fear of coercion will reduce the desire of other parties to reciprocate.

The two episodes discussed above – the United States’s resistance to judicial enforcement of the UN Charter and to direct application of the GATT by the courts – both involved uncertainty over third-

¹⁷² See, e.g., *Van Gend en Loos v. Nederlandse Administratie der Belastingen* (Case 26/62), [1963] E.C.R. 3 (Treaty of Rome provisions create rights that individuals can directly enforce in domestic courts); *Costa v. ENEL* (Case 6/64), [1964] E.C.R. 1141 (European Court of Justice has final authority to review application of Treaty by domestic court); *Lopez Ostra v. Spain*, 20 Eur. H.R. Rep. 277 (1994) (solid waste treatment plant located near home violates right to privacy); *Maria Guerra v. Italy*, 26 Eur. H.R. Rep. 357 (1998) (serious environmental pollution violates right to privacy); *Jordan v. United Kingdom*, 37 Eur. H.R. Rep. (2001) (deficiencies in police investigation of homicide constitute a violation of European Convention’s right to life); *Pretty v. United Kingdom*, 35 Eur. H.R. Rep. 1 (2002) (right to die); *I v. United Kingdom*, 36 Eur. H.R. Rep. 53 (2002) (failure to give legal recognition to sex change violates right to privacy); *E. v. United Kingdom*, 36 Eur. H.R. Rep. 31 (2002) (failure by social services to exercise due diligence in supervising children endangered by home environment violates European Convention).

¹⁷³ Carlos Manuel Vásquez, note 135 *supra*, at 695 (subject is the “most confounding” in United States law).

party coercion, not a reaction to a clear decision to provide coercive enforcement. In both cases, a few progressive state courts had endorsed direct enforcement, but no federal court had done so and nothing like a consensus for that result existed. The Dulles reassurances of 1955 and the Uruguay Round Agreements Act of 1994 represented steps to foreclose possibilities, rather than decisions to reverse any particular outcome.

We emphasize this point because most instances of coercive enforcement of international agreements involve uncertainty rather than clear decisions incorporated into the agreement. Unless framers make an explicit commitment against coercion, they face a positive, and perhaps substantial, risk that third parties at some future date will take it upon themselves to compel compliance. To the extent third-party enforcement impedes self enforcement, this risk lowers the value of international agreements. The logical next question, then, is what meta-rules parties to international agreements might apply to minimize the risk of unwanted future assertions of third-party enforceability.

D. The Limits of Coercion

To recapitulate our argument, self enforcement plays an important, and perhaps underappreciated, role in the enforcement of international agreements. This role, however, is not limitless. Coercive sanctions imposed by third parties can increase the value of commitments contained in agreements, if (1) the commitments rest on verifiable information and (2) repeat play, reputation, and a preference for reciprocal fairness are, for any number of reasons, ineffectual under the circumstances. The challenge is to define the conditions for the application of third-party coercion with sufficient clarity to avoid compromising the value that self enforcement can generate. And because third-party coercion sometimes can be desirable and sometimes is not, the clearest meta-rules – an all-or-nothing approach of an absolute ban on third-party enforcement or of a rule of automatic third-party enforcement – will not work.

An obvious complication is that complete elimination of uncertainty about the availability of third-party coercion would be unrealistic. Neither international agreements themselves, nor any hypothetical meta-rules for interpreting and applying those agreements, can anticipate all future states and specify

for each whether third-party or self enforcement will apply. Instead, decisionmakers confront the unavoidable tradeoff between accuracy and clarity, the former approach implemented by open-ended rules that delegate to the third party considerable discretion as to whether to apply coercive sanctions, the latter by bright-line rules that may produce both Type I and Type II errors.¹⁷⁴ The choice of any particular meta-rule necessarily involves accepting one or the other of these costs.

Consider as a limiting case a formalistic approach that incorporates four distinct but interrelated interpretive defaults: (1) the third party (say, a domestic court) applies the same interpretive strategy to all international agreements, regardless of subject or object; (2) the third party requires an explicit reference to coercive third-party enforcement as a condition of enforceability; (3) the third party divines obligations within an agreement only where clearly stated; and (4) the third party will allow only those persons clearly embraced by the agreement to invoke its provisions. Applying this approach, a court would intervene to enforce the agreement only if the agreement invoked this mode of enforcement,¹⁷⁵ only if the agreement clearly created an obligation relevant to the dispute before the third party,¹⁷⁶ and only if the agreement applied to the person seeking to enforce it.¹⁷⁷ In making these decisions, the court would not consider the subject or object of the agreement, and in particular would not distinguish between agreements intended to benefit or protect particular classes of individuals.¹⁷⁸

Together these defaults implement a consistent strategy, in the sense that they are necessary for a comprehensively restrictive approach to third-party coercion. As a matter of logic, however, a third-party enforcer could depart from one or more of them. The issue is why. One might justify a different

¹⁷⁴ For fuller discussion of the accuracy/precision tradeoff, see Isaac Ehrlich & Richard A. Posner, *An Economic Analysis of Legal Rulemaking*, 3 J. LEG. STUD. 257 (1974).

¹⁷⁵ See, e.g., *Argentine Republic v. Amerasia Shipping Corp.*, 488 U.S. 428, 441-43 (1989) (no judicial enforcement of Geneva Convention on the High Seas or Pan American Maritime Neutrality Convention).

¹⁷⁶ See, e.g., *Sale v. Haitian Centers, Inc.*, 509 U.S. 155 (1993) (interpreting Convention Relating to the Status of Refugees as not creating an obligation to provide a hearing on refugee status before returning detained persons to state of emigration).

¹⁷⁷ See, e.g., *Chubb & Son, Inc. v. Asiana Airlines*, 214 F.3d 301 (2nd Cir. 2000) (finding that U.S. obligations under Warsaw Convention did not apply to Korean parties), *cert. denied*, 533 U.S. 928 (2001).

¹⁷⁸ See, e.g., Hersh Lauterpacht, *Restrictive Interpretation and the Principle of Effectiveness in the Interpretation of Treaties*, 26 BRIT. Y.B. INT'L L. 48 (1949) (discussing general presumption in international law of restrictive interpretation of international agreements so as to minimize derogations of national sovereignty).

default for substantive reasons unconnected to contract theory, or because of strategies suggested by contract-based arguments for optimal defaults.

Consider first conventional arguments that disregard contract theory. Among specialists in international law, it has become popular to distinguish between traditional international-law obligations and the new body of human rights law that has emerged in the years since World War II. An early decision of the International Court of Justice, for example, noted that the postwar multilateral human rights treaties do not aspire to “individual advantages or disadvantages to States, or . . . the maintenance of a perfect contractual balance between rights and duties.”¹⁷⁹ Because they involve a general humane purpose rather than bargained exchange, these instruments, the argument goes, require distinct and generous interpretive strategies to ensure the fulfilment of their benign objectives.¹⁸⁰

It is possible to view arguments of this sort as a variation of the claim that international law produces economies of scale and scope, albeit limited to the particular category of human rights law.¹⁸¹ In essence, it asserts that because human-rights-law obligations are so fundamentally important, any increase in enforcement resources devoted to them would produce a positive return. Domestic judicial enforcement is simply one way of augmenting these resources.

If the premise of this argument is correct, of course, this argument is unanswerable. And perhaps human rights rest on such incommensurable values that no cost-benefit analysis can be applied. Scholarly debate then would have to move to second-order issues, such as whether the power to determine which values demand super-enforcement should rest with judges or instead with more politically accountable actors. That debate is interesting, but not the subject of this paper.

¹⁷⁹ Reservations to the Convention on the Prevention and Punishment of the Crime of Genocide, 1951 I.C.J. 15, 23.

¹⁸⁰ See, e.g., Martin A. Rogoff, *Interpretation of International Agreements by Domestic Courts and the Politics of International Treaty Relations: Reflections on Some Recent Decisions of the United States Supreme Court*, 11 AM. U.J. INT'L L. & POL'Y 559, 632-41 (1996) (linking argument for new interpretive approaches to changing nature of international community and corresponding obligations); Harold H. Koh, note 145 *supra*, at 2398-2402 (arguing for new interpretive strategies to take account of a “new international legal process”).

¹⁸¹ See note 153 *supra* and accompanying text.

The best we can do is express skepticism that rationalist, instrumentalist analysis is so limited. We suspect that enforcement of even dignitary values and decency come at some cost. Moreover, if coercive enforcement crowds out self-enforcement, then more enforcement is not only costly but also reduces the effective level of compliance with the agreement.¹⁸² Enforcement thus can become counterproductive on two separate dimensions. Perhaps the episode involving international law and the civil rights revolution supports this point, although we understand that this inference is controversial. In any case, rational reciprocity presents a challenge to those who would not recognize any limits on the enforcement of this class of rights.

Consider next what contract theory might suggest about optimal defaults. On the one hand, the literature argues that a default should represent the contract term that the parties would prefer regarding the issue in question, were bargaining resources unlimited – what conventionally is called a majoritarian default.¹⁸³ An alternative approach asks whether a default can correct information asymmetries between the parties by inducing the disclosure of information that would augment the joint value of the parties bargain – what is called an information-forcing or penalty default.¹⁸⁴ Consistent with this analysis, we should speculate as to whether parties to international agreements generally prefer coercive third-party enforcement or not, and whether any information-forcing argument exists for choosing an alternative enforcement regime.¹⁸⁵

¹⁸² See notes 59-67 *supra* and accompanying text.

¹⁸³ See, e.g., Charles J. Goetz & Robert E. Scott, *The Mitigation Principle: Toward a General Theory of Contractual Obligation*, 69 VA. L. REV. 967 (1983).

¹⁸⁴ See, e.g., Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87 (1989). Information forcing defaults are also designed to mimic majoritarian preferences because parties can always create such rules on their own. Thus, the default is, in effect, replicating what the parties would want *ex ante* when it creates incentives for a party to disclose information to its contracting partner. See Robert E. Scott, *Rethinking the Default Rule Project*, 6 VA. J. 84, 85-86 (2003).

¹⁸⁵ We recognize that we are eliding a number of difficult questions raised in the contracts literature, including whether to determine a majoritarian default by the number of prefers or the intensity of preference, how to address offsetting asymmetries in information, *etc.* For further discussion of these problems, see Barry E. Adler, *The Questionable Ascent of Hadley v. Baxendale*, 51 STAN. L. REV. 1547 (1999); Clayton P. Gillette & Paul B. Stephan, *Richardson v. McKnight and the Scope of Immunity after Privatization*, 8 SUP. CT. ECON. REV. 103, 130-32 (2000); Charles J. Goetz & Robert E. Scott, *The Limits of Expanded Choice: An Analysis of the Interactions Between Express and Implied Contract Terms*, 73 CALIF. L. REV. 261, 309-11 (1985); Jason Scott Johnston, *Strategic Bargaining and the Economic Theory of Contract Default Rules*, 100 YALE L.J. 615 (1990)

We have not done the empirical work that might provide confident answers to these questions. Instead, we will focus on a single issue that relates to the optimal default literature. All things being equal, do parties to an international agreement typically wish to start from a premise of reciprocal fairness, or from a premise of chiseling and noncompliance? Should defaults, in other words, put the onus on parties who distrust one another to articulate their suspicions, or should defaults force bargaining states to make a case for pure self enforcement without external coercion?

The experimental evidence discussed in Part I of this paper suggests a response to this question. Introducing the subject of failure and opportunism tends to have an unhappy effect on the resulting agreement. People who know that they may be punished perform less well than people who know that they may be rewarded. Forcing states to rebut the premise that they will defect may trigger the same dynamic. Beginning with a default of no coercive third-party enforcement and then requiring the bargaining states to stipulate deviations from this rule allows negotiations to proceed without undermining a premise of reciprocal fairness.

Moreover, if the prospect of coercive enforcement makes some international bargains unattainable, then uncertainty about that prospect also should impose a burden on valuable transactions into which states might enter. Defaults that reduced this uncertainty at a reasonable cost thus would be optimal.

This analysis does not require U.S. courts to rethink their approach to the enforcement of international agreements. The articulated doctrine may be unsatisfactory, but the general pattern of U.S. practice conforms to a set of defaults that restricts domestic enforcement to those instances where the international agreement calls for it in express terms.¹⁸⁶ Our immediate, tentative observation is simply that international practice generally, and U.S. practice in particular, suggests that an optimal default of no coercive enforcement already operates.

IV. CONCLUSION

¹⁸⁶ See notes 175-178 *supra* and accompanying text.

This paper represents an initial attempt to develop a general theory of the enforceability of international agreements. We focus attention on the experimental evidence of a general, though not universal, preference for reciprocal fairness and the possibility that reciprocity-based enforcement mechanisms can be crowded out by coercive enforcement. Both the evidence for these phenomena in individuals and the extension of the experimental hypotheses to international behavior remains preliminary and tentative. We nevertheless find the implications of these hypotheses sufficiently important to justify our analysis.

First, we need a theory that can better explain the relationship between individual preferences and the behavior of states. International relations theory makes a first step in this direction with its distinction between regimes and national interest. But how might a preference for reciprocal fairness manifest itself in regime preferences? Under what circumstances would a regime expose itself to the risk of external coercion? How do preferences vary among types of regimes, and what explains these variations? We have made a start at answering these questions, knowing that further analysis lies ahead.

Second, more empirical work is needed. States obviously are not amenable to the kinds of laboratory experiments that we have described, but other techniques exist to test hypotheses about state behavior. We have offered here a typology of enforcement mechanisms in international agreements and made some anecdotal observations, but we do not pretend to have made a systematic study of the field. That work also lies before us. We are confident, however, that our speculations point in the right direction and suggest a research agenda for the field.

However tentative our conclusions may be about the nature of self enforcement and its relation to coercive enforcement, we wish to stake out a strong methodological claim. Contract theory offers rich and important insights for the field of international law. It has made robust contributions to our understanding of private behavior, and we believe its extension to international bargaining and agreements is overdue. A focus on rational behavior under conditions of limited information, which is so important in enabling us to explain and evaluate private contracting, also can enrich our understanding of international relations and international law. Whether rational reciprocity ultimately

carries the day in providing a robust account of international relations and lawmaking, as we hope it will, we are convinced that future work in international law will owe a substantial intellectual debt to contract theory and scholarship.