The Economics of Public International Law

Alan O. Sykes

Follow this and additional works at: https://chicagounbound.uchicago.edu/law_and_economics

Part of the Law Commons

Recommended Citation

This Working Paper is brought to you for free and open access by the Coase-Sandor Institute for Law and Economics at Chicago Unbound. It has been accepted for inclusion in Coase-Sandor Working Paper Series in Law and Economics by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
The Economics of Public International Law

Alan O. Sykes

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

July 2004

This paper can be downloaded without charge at:
The Economics of Public International Law

Alan O. Sykes*

Abstract: This paper is a preliminary draft for eventual inclusion in the Handbook of Law and Economics, A. Mitchell Polinsky & Steven Shavell editors. It reviews and synthesizes the work of economists and law and economics scholars in the field of public international law. The bulk of that work has been in the area of international trade, but many of the ideas in the trade literature have implications for other subfields. Recent years have seen a significant increase in research on other topics as well. The paper begins with a general framework for thinking about the positive and normative economics of public international law, and then proceeds to a treatment of specific topics including customary law, strategic alliances and the laws of war, international trade, international investment, international antitrust, human rights law, conflicts of law, and the international commons (fisheries).

1. Introduction

International law has been recognized as a distinct field of study within the legal academy for well over a century, but economically-oriented scholars have paid it relatively little attention. Dunoff and Trachtman (1999). By far the bulk of the law and economics research in international law pertains to the law of international trade. Systematic work on other topics is limited at best, although research in the field generally is accelerating and the subject can properly be considered a growth area.

The relative dearth of prior work, especially formal work, and its concentration on international trade issues, poses a number of challenges for a survey of this sort. An excessive emphasis on international trade will mask the richness and diversity of the field, and obscure rather than illuminate the potential research agenda. Moreover, large segments of the economic literature pertaining to international trade law already receive attention in the three volume Handbook of International Economics series, especially in Staiger (1995a), and in a number of other extant and forthcoming volumes focused on WTO issues including Bagwell and Staiger (2002) and Grossman and Helpman (2002). Accordingly, I have limited the treatment of trade issues in this chapter, emphasizing topics that illustrate broader themes for the economic analysis of international law. Much of the chapter will instead be devoted to a general framework for thinking about international law, along with brief discussions of a number of topics outside the realm of international trade.

* Frank & Bernice Greenberg Professor of Law, University of Chicago. I have received thoughtful comments on an earlier draft from Kyle Bagwell, Eric Posner, Steven Shavell, and participants in the Handbook of Law and Economics conference at Stanford Law School.
trade, with the understanding that an informal treatment is all that the existing literature has to offer on many of them.

Sections 2 and 3 of the chapter provide legal background on the field of international law, followed by a discussion of general economic considerations that cut across a range of possible topics. The analysis will encompass the various possible functions of international law, the challenges involved in realizing gains from international cooperation, the design and function of mechanisms for its enforcement, and the interface between domestic and international law. Sections 4-10 of the chapter consider particular topics, including strategic alliances and laws of war, international trade, international investment, international antitrust, human rights law, conflicts of law, and the international commons (fisheries).

As a final, preliminary disclaimer, I have not undertaken to survey and incorporate the vast political science literature on public international law. Much of that literature is excellent, and the “rational choice” literature in particular is often quite close in both spirit and method to the work of economists. I omit attention to it not because of any negative judgment about its quality, but to make the task at hand a manageable one. Readers seeking a window into the rational choice perspective on international relations and institutions might wish to consult Snidal (1996) and (2002). Carlsnaes, Risse and Simmons (2002) provide a broader introduction to modern international relations work in political science.

2. Legal Background

The field of international law is conventionally divided into two subfields: “public” international law and “private” international law. “Public” international law refers to the body of law that governs relations between states or countries. “Private” international law refers to the body of law that governs international relations between private citizens or companies. Most private international law relates to international business transactions, and may be subsumed for analytic purposes under other topics such as contract law, corporate law and tax law that are the subject of other chapters in the Handbook of Law and Economics. Accordingly, my focus in this chapter is almost exclusively on public international law.

The genesis of public international law necessarily differs from that of domestic law. No international legislature exists to pass the equivalent of domestic statutes, and no international court exists with the power to create a general international common law. Instead, public international law arises only by agreement among states.
Often, agreement is manifest in an instrument known as a *treaty*. A treaty is an agreement executed by duly authorized officials of signatory states, evincing an intention to make it a binding legal obligation. Treaty obligations are themselves governed by an over-arching treaty known as the Vienna Convention on Treaties, which supplies rules for their interpretation and enforcement. Note that the concept of a “treaty” is not necessarily the same in international and domestic law. The U.S. President, for example, has the authority in many areas to bind the United States internationally through “Executive Agreements,” sometimes accompanied by formal Congressional approval and sometimes not. These agreements have the same status as treaties under international law, even though they are not treaties under domestic law (which provides that “treaties” must be approved by a two-thirds vote of the U.S. Senate).

Agreement may also become manifest as *customary international law*, which is defined as a “general and consistent practice of states followed by them from a sense of legal obligation.”\(^1\) The traditional test for the emergence of customary law thus requires a high degree of consistency in state practice, and a belief that the practice has become a legal obligation. Both of these requirements are imprecise, and scholars often disagree about what practices have achieved the status of customary law. Some rules of customary international law are uncontroversial, however, such as those relating to aspects of diplomatic immunity. It is generally said that states may avoid an obligation to obey customary international law by “opting out” at an early stage of its evolution, but once they have manifested agreement with it through conforming behavior, any subsequent deviation is illegal.

International legal scholars also make reference to the concept of “soft law.” Soft law encompasses a range of things, including formal agreements that are understood not to be “binding” under international law, as well as agreements that may be “binding” but that are essentially hortatory or aspirational. Examples of each abound—the Cuban Missile Crisis was settled by an informal agreement, for example, while numerous provisions in WTO treaty text encourage but do not require special trade treatment for developing countries.

The enforcement of international law, to the extent that it is successful, occurs in a variety of ways. The closest analog to the coercive enforcement powers often exercised by domestic courts is found in the United Nations. A serious breach of U.N. obligations may result in the authorization of substantial sanctions by the Security Council, or in

extreme cases in a resolution authorizing the use of military force against the violator state. Much of international law falls outside the purview of U.N. obligations, however, and thus outside its enforcement mechanism. Some international legal regimes have their own tribunals with the power to adjudicate violations (such as NAFTA and the WTO), The International Court of Justice also has jurisdiction to hear a broad range of disputes. The power to adjudicate disputes may or may not be accompanied by the power to authorize or impose sanctions, however, and the nature of any sanctions may be tightly circumscribed. Formal sanctions for the violation of WTO obligations, for example, are limited to the withdrawal of benefits under WTO agreements. If international law is incorporated into domestic law, as quite often occurs, then the powers of domestic courts can be brought to bear on certain types of violations. This mechanism too has its limitations, as many international legal obligations are never incorporated into domestic law. Further, domestic courts are often limited in their jurisdiction to enforce international obligations that are so incorporated—principles of foreign sovereign immunity, for example, often insulate states from actions against them in foreign courts. Finally, many international legal obligations exist as to which there is no formal enforcement or sanctioning mechanism at all.

Where some enforcement mechanism exists, a further issue arises as to who has standing to invoke it. Public international law governs relations among states and, generally speaking, only states have standing to enforce it. Private citizens have no right to pursue most claims under international law even if they have suffered substantial injury due to a violation. An important exception of sorts exists, however, if nations incorporate international law into their domestic legal systems. Private actors may then be able rely on their access to domestic courts to enforce what originates as an international legal obligation. Finally, private citizens occasionally have standing to pursue claims before international tribunals, as in the case of the NAFTA investor rights provisions which allow investors access to NAFTA arbitration.

3. Economic Aspects of International Law

Public international law represents a number of distinct phenomena. Some “law” may be no more than a behavioral regularity in the practices of states, while other law may represent rules coercively imposed on less powerful states by more powerful states. Still other types of law may arise to promote the domestic objectives of participating officials. But many of the more interesting and important pockets of international law may be seen as efforts to coordinate the behavior of states to address externalities. These
externalities may be nonpecuniary, the sort that produce inefficiency in competitive markets, or pecuniary, creating inefficiency due to an absence of competitive conditions.

In this section, I begin with some broad observations about the economic perspective on public international law. Because systematic discussions of international law often imagine that states behave as if they have “preferences,” the first topic concerns the conceptualization of states as rational actors. The analysis proceeds to a general discussion of customary international law, to a discussion of the economics of treaties, and finally to consideration of the interface between domestic and national law.

A. States as Rational Actors

Positive economic analysis of international legal regimes conventionally proceeds from an assumption that states behave as if they are rational maximizers over some set of preferences regarding the outcomes of their interaction. The specific assumptions that may be made in this regard are myriad. States may be assumed to behave as economic welfare maximizers, or to maximize a social welfare function that weighs the welfare of certain constituencies more heavily than others. The preferences of the “state” may be assumed to be those of its political leaders, who may maximize votes, campaign contributions, or their personal welfare. Innumerable other variations can be imagined depending on the context.

Whatever precise assumption is made about the nature of preferences, it is common to embody a further assumption that states act as if they “care” primarily or exclusively about their own welfare or interests, and less or not at all about the welfare or interests of other states or their political leaders. A divergence will then arise between the national maximand and the global maximand.

The assumption that states have preference orderings and act as rational maximizers is surely somewhat simplistic. States represent an aggregation of many different actors, whose preferences may well be at odds. The actor with the power to choose among alternatives may change over time, and the constraints imposed on actors with the power to make choices can change over time (in the United States, think of the President as the actor with the power to make choices on international matters, subject to constraints imposed by Congress). Even when it is plausible to assume that a pertinent decisionmaker has a preference ordering over the available alternatives at a point in time, therefore, the notion that the “preferences” of the “state” are stable over time, or that they obey potentially important regularity assumptions, may be quite problematic.

Although one must acknowledge this problem, there is often little to be done about it in a tractable modeling framework beyond remaining attentive to its possible
implications for each subject area. Such a framework proceeds in the tradition of other areas of economic analysis, which embrace their own simple assumptions about the objective functions of corporations, bureaucracies, and other large institutions. Here, as in those other areas, the test is not whether the assumptions are fully descriptive of behavior, but whether they yield useful insights with empirical purchase.

Economic analysis of international law also has its normative side, of course, which rests on assumptions about what states ought to be maximizing. Once again, a variety of possible objective functions might be assumed, although the conventional measure of economic welfare is often employed.

B. The Economics of Customary International Law and “Soft” Law

A great deal of work has been done on the economics of “custom” in various contexts. Commentators have written about the use of custom evidence to prove negligence in tort actions, the use of customary business practices as a basis for default rules in contract law, the efficiency of social norms, and the general phenomenon of “order without law” in primitive or frontier societies. Such topics receive significant attention in other chapters in this Handbook.

Despite the attention to custom in other contexts, very little has been written about customary international law from an economic perspective. The most notable exception is Goldsmith & Posner (1999 & 2004).

Recall the standard characterization of customary international law: it emerges when there is a high degree of convergence in the practice of states, and a belief that adherence to the practice has become a legal obligation. The latter requirement is known as opinio juris and is central to the existence of customary law according to traditional doctrine. Mere regularities in state behavior, without opinio juris, are not law.

Goldsmith and Posner contend that this description of customary international law is largely incoherent. Their alternative theory begins by offering a positive theory of convergence is state practice, which they suggest may result from four distinct phenomena. The first is simple coincidence of interest, whereby all states behave the same way because it is in their unilateral interest regardless of the choices made by other states. They offer “ambassadorial immunity” as a possible example (although this subject is governed by treaty as well as custom in modern times)—states may protect the ambassadors of other states, even in times of conflict with them, because the ambassadors perform a valuable function in facilitating communication with other governments. A second explanation for convergence of practice is pure coercion. Here, they suggest that
the custom of “free ships, free goods,” whereby all property on neutral ships is immune from seizure (including enemy property), is at times illustrative—powerful states may respect the principle because the seizure of neutral ships to capture enemy property is not worth the bother, while weaker states may respect the principle for fear of retaliation by powerful states. The third possible reason for convergence arises when a common practice represents the solution to an iterated Prisoner’s Dilemma, which can be sustained over time by states with open-ended time horizons and sufficiently low discount rates. They again offer ambassadorial immunity as a possible illustration, suggesting that an exchange of ambassadors amounts to an exchange of hostages, and that the prospect of retaliatory acts against one’s own ambassador can dissuade any temptation to interfere with the ambassadors of others. Finally, convergence may arise in the face of a pure coordination problem or otherwise where a “focal point” is useful. They suggest that the convergence on a three-mile limit for territorial waters is an example here. For a variety of reasons including security, nations have an interest in claiming dominion over waters along their coast, but the exact limits of territorial waters is to a degree a matter of indifference—a three mile limit supplies a focal point that all nations can accept.

In short, Goldsmith and Posner argue that convergence in state practice occurs for reasons of pure national self interest, albeit not the same reason every time. They further suggest that continued adherence to customary practice happens because the self-interested reasons for convergence remain in place, not because of any independent sense of legal obligation. Opinio juris, they suggest, is a fiction, and what legal scholars refer to as customary “law” is really no more than a descriptive account of certain regularities in the behavior of states.

To bolster this latter claim, Goldsmith and Posner document how ostensible rules of customary law are frequently violated when states have an interest in deviating. They further illustrate how rogue states, which they suggest have shorter time horizons and higher discount rates, are more likely to deviate than others. Because historical violations and breakdowns of custom can be linked to self-interested reasons for them, Goldsmith and Posner find anecdotal empirical support for the claim that customary practices are mere regularities of self-interest, and that customary law per se exerts no tug on state behavior.

The proposition that “customary international law” generally emerges from the self-interested interaction of states, and that it promotes with their mutual interest for one reason or another, seems rather unremarkable. It would indeed be odd if a customary practice emerged on a large scale that made its adherents worse off over an extended period of time. While this aspect of Goldsmith and Posner’s analysis seems compelling, a
skeptic might argue that they have not fully made their case on the nonexistence of *opinio juris*. Even if customary international law had some force of its own quite apart from the narrow self-interest of a state regarding a particular custom, one might still observe the same anecdotal bits of evidence that Goldsmith and Posner catalog. Nations might still deviate when their self-interested reasons were strong enough, for example, and rogue states might still be the most likely to deviate. All that would be required is that the behavioral force of *opinio juris* be limited, so that counter-incentives of sufficient strength could override it. Thus, although Goldsmith and Posner are surely right that the traditional scholars cannot prove the existence of *opinio juris* by pointing to conformity with custom, neither can the detractors of the traditional view prove its nonexistence merely by pointing to self-interested deviations from custom.

If the empirical evidence is inconclusive, it remains to ask whether *opinio juris* can be given any theoretical content. Why would states feel any obligation to observe a custom that is no longer in their self interest? Traditional international law scholars suggest that once a practice becomes “law,” it infuses the morality of national bureaucrats, who then feel a sense of obligation to obey it. One might restate the proposition as a suggestion that “law” has expressive force and alters the preferences of pertinent national actors, leading them to prefer to obey it (or that they simply have an exogenous preference to obey all “law.”) The difficulty with this account, as even noneconomic scholars have noted, is its circularity. Law exists only after the sense of legal obligation arises according to the definition of customary law, yet the sense of legal obligation is said on this account to follow after the emergence of “law.”

A possible alternative account is suggested by Guzman (2002), who relies on the idea that violations of international law may damage a state’s reputation. He suggests that international strategic interaction on narrow issues is generally embedded within a larger games, and that players’ willingness to cooperate with other players on current issues may then turn on whether a player has developed a reputation for cooperation in the past. Under the usual assumptions that prevent backwards unraveling of cooperation (an infinite or open-ended time horizon) and that limit the short terms gains from defection (such as a low discount rate), Guzman argues that reputational considerations create the possibility of an equilibrium in which mutual cooperation is sustained over time in what might otherwise appear to be a one-shot game with defection as the Nash outcome.

The addition of reputation to the analysis suggests a possible economic interpretation of *opinio juris*. One might define it simply as a tendency to obey customary law due to the damage that defection does to a state’s reputation as a cooperator, costs
that are incurred not in the simple game in which defection is contemplated but in all other games where reputation affects the strategies played by other states.

Traditional international law scholars will find little solace in the reputational interpretation of *opinio juris*, however, because concern for reputation is no less self-interested than concern for payoffs in a narrower strategic interaction. Further, reputational considerations may be of minimal significance as a practical matter in many settings as both Goldsmith and Posner and Guzman argue. This general issue receives further attention below.

Aside from its examination of *opinio juris*, the law and economics literature makes a number of other useful points about the role of customary law. The commentators seem to agree that the ability of customary international law to orchestrate cooperation is limited to narrow circumstances. Problems that require complicated solutions are unlikely to be solved by implicit cooperation—express negotiation and communication will probably be necessary. Further, problems that require the simultaneous cooperation of large numbers of nations will also be difficult to solve because of free rider problems in the enforcement mechanism. Even when a practice appears “customary” on a global scale, therefore, and is thought to represent mutual cooperation, the suggestion is that it is usually no more than a recurring regularity of bilateral interaction.

Guzman makes the further point that if reputation is what creates some “force of law,” then there is no reason to limit our conception of “law” to customary international law and treaties. Reputational concerns may be quite important to a world leader who gives her word to another, whether or not it is done in any formal fashion and whether or not it concerns some practice that is widespread in the international community. The traditional line between “hard law” (binding treaties and customary law) on the one hand and “soft law” (such as informal agreements and statements of intention) on the other may thus be quite misleading. Depending on context, states may be considerably more likely to comply with soft law than with hard law, and there is no reason to think that hard law is always preferable for orchestrating cooperation.

In short, economic thinking about customary international law calls into question the very meaning of the concept. It suggests that practices termed “law” in various quarters are no more that behavioral regularities that emerge from self-interested interaction between states facing similar problems. The codification of customary law merely serves to publicize focal points, and to write down the rules of any game to facilitate future adherence to them. The capacity of customary “law” to solve important problems that require cooperation or coordination is quite limited, and will tend to be
restricted to issues that admit of a simple solution that can be sustained through small numbers strategic interaction.

Formal modeling bearing on these issues is in its infancy. Fon and Parisi (2004) offer a simple model of custom formation that supports the intuition that customary law is more useful when the preferences of states are relatively more homogeneous. They also consider the role of what they term the “persistent objector” and “subsequent objector” doctrines that allow states to obtain an exemption from customary rules.

C. The Economics of Treaties and Other International Agreements

Virtually all of the economic writing on treaties focuses on particular subject areas, with the notable exceptions of Goldsmith & Posner (2004) and Guzman (2002) cited earlier. In this section I draw to a limited extent on those two sources, but also on ideas developed in more specialized contexts to suggest some general points about the economics of treaties.

In contrast to customary international law, which can emerge through convergence of practice without much communication across states, treaties always involve direct communication, negotiation, and the embodiment of the results in a document. This process is costly, and the reasons for the creation of treaties are narrower or at least different from the reasons given earlier for convergence of state practice on custom. The coincidence of interest explanation for some customary practices, for example, cannot explain why states would incur the costs of creating a treaty. The exercise of pure coercion does not require a treaty either, although to be sure a treaty may be used to orchestrate an end to coercion. Treaties are likely to be valuable instead when state action creates externalities for other states, and when purely decentralized cooperation without formal communication is inadequate to address them (although a few treaties may have other functions, as discussed in later sections).

The mere fact that cooperation is better orchestrated through a process of direct communication, of course, is not sufficient to justify a treaty. Much communication between states occurs without any resulting agreement, and international agreements can arise in the course of communication that are informal and never rise to the level of a treaty. Goldsmith and Posner thus consider the question of why states resort to “legalization” by formally executing a treaty and making it “binding” as a matter of international law in preference to reliance on less formal, nonbinding agreements. They suggest that the legalization of an agreement may reveal information about a state’s commitment to the agreement—in their terms, it shows that the state is “serious” about
the agreement. A signal of “seriousness” will only be needed when “seriousness” is private information, and will only be credible if reputational penalties are greater for the violation of a “binding” agreement than for violation of an informal agreement. Hence, this explanation requires that reputation be important to state actors. A second consideration affecting the choice to legalize is the fact that informal agreements may bypass domestic constitutional constraints on the creation of treaties. In the United States, for example, the President has the capacity to conclude and execute informal agreements without Congressional oversight in many areas. Formal treaties (or Executive Agreements that must be approved by Congress) give the legislature greater opportunities to participate and may then constrain the President to less preferred options. But legislative participation may also give the agreement greater durability against changes in administrations, as well as greater force in domestic law. The President will choose between the two options depending on the balance of competing considerations in each case. A final consideration is that legalization subjects the treaty to the interpretive default rules of the Vienna Convention on Treaties. Informal agreements may be chosen out of a desire to opt out of those rules.

Leaving aside for now the choice between formal and informal agreements, it is perhaps useful at this point to set forth the basic formal framework for modeling the externality problem and for identifying the potential gains from international cooperation. Variations of this basic approach pervade the literature on individual topics. Imagine two states, denoted A and B, each of which have control over a vector of policy instruments, \( \alpha \) and \( \beta \), respectively. [Nothing changes importantly (beyond the algebra) if the analysis is generalized to \( N \) states.] The respective welfare functions for the two states are \( W^A(\alpha, \beta) \) and \( W^B(\alpha, \beta) \). Assume that each state’s welfare is increasing and concave in its own policy choices. The vectors \( \alpha \) and \( \beta \) can represent a myriad of policy areas—tariffs, tax rates and rules, immigration restrictions, emissions controls, and so on. In the absence of communication and agreement, each state maximizes its welfare taking the actions of the other as given, selecting \( \alpha \) and \( \beta \) such that \( \partial W^A(\alpha, \beta) / \partial \alpha = 0 \) and \( \partial W^B(\alpha, \beta) / \partial \beta = 0 \). Equilibrium (Nash) arises when both conditions hold, given the other state’s choice of policies. Will the equilibrium be efficient? The answer is plainly no in general: A point on the Pareto frontier may be derived by choosing \( \alpha \) and \( \beta \) simultaneously to maximize the welfare of one state, subject to the constraint that the welfare of the other achieve some fixed, attainable value (a standard technique for deriving conditions for any optimal contract). The first order conditions for this problem require that \( \partial W^A(\alpha, \beta) / \partial \alpha + \lambda \partial W^B(\alpha, \beta) / \partial \alpha = 0 \), and \( \lambda \partial W^B(\alpha, \beta) / \partial \beta + \partial W^A(\alpha, \beta) / \partial \beta = 0 \), where \( \lambda \) is a Lagrange multiplier. With the welfare constraint binding and thus \( \lambda > 0 \), it is
clear that the conditions for Pareto optimality cannot correspond to the earlier conditions for Nash equilibrium unless \( \partial W^B(\alpha, \beta)/\partial \alpha = 0 \) and \( \partial W^A(\alpha, \beta)/\partial \beta = 0 \). In words, the equilibrium without international cooperation will achieve the Pareto frontier only in the absence of externalities, and the function of international agreements in the presence of externalities is to enable states to commit to behavior that will move them closer to the Pareto frontier.

\[ i. \text{Factors that Facilitate or Impede International Agreement} \]

If externalities suggest an opportunity for international agreements to improve on the equilibrium without them, it does not follow that states will succeed in achieving agreement. Casual empiricism suggests that useful agreements have been reached in a number of areas (e.g., trade), but that many areas laden with apparent externalities have not been successfully addressed through international agreements (e.g., immigration). In other areas, some issues have been addressed through agreement but not others (e.g., investment and environment). What explains these various “successes” and “failures”? Perhaps the starting point for analysis is the Coase Theorem. One would expect international agreements to exhaust potential joint gains from solving externality problems only to the degree that those gains remain after all transaction costs have been accounted for in the calculus (which must be understood broadly to include not only monetary costs of achieving agreement but all factors that affect the political acceptability of agreement). To understand the universe of international agreements (and the areas in which they are absent), one must therefore ask not simply whether externalities arise, but also whether the transaction costs of international agreements to address them are low enough to permit agreements to go forward. A variety of considerations will affect the magnitude of transaction costs.

Trivially, agreement can only arise if some agreement lies in the core of the bargaining game—each state that becomes party to an agreement must perceive itself better off than by refusing to participate (or by breaking off with others into a smaller numbers agreement in the multilateral case). Yet it is easy to imagine settings in which international externalities lead to an equilibrium off the Pareto frontier in the absence of cooperation, but the bargaining options are too limited to admit of a Pareto improvement in the core. For example, imagine two states, one of which contains a monopoly producer and exporter of widgets, and the other of which is a consumer of widgets with no monopoly power over any tradable good or service. The monopolist exploits its market power with the familiar deadweight losses, although much of its monopoly profit comes
as a transfer from consumers abroad; joint gains would arise if each state pursued a sensible anti-monopoly policy. But if the two states try to strike a bilateral agreement that merely requires each of them to adopt an anti-monopoly policy, the state with the monopolist may well object that such an agreement leaves it worse off than without it.

Two obvious and related solutions to the problem suggest themselves. The first is monetary sidepayments from one state to the other. This device is sometimes employed in practice, usually in the form of a promise of monetary aid to a state that cooperates on some issue (the recent aid to Pakistan associated with its quiet assistance in the invasion of Afghanistan is illustrative). Analytically equivalent, and probably more common in practice, the scope of bargaining can be expanded to include other issues (or perhaps other states, a point explored later). The state that enjoys a widget monopoly may be willing to forego monopoly rents in exchange for valued concessions on security issues, environmental matters, and so on. The general lesson is that issue linkage in international negotiations can greatly expand the scope of possible agreements. A likely recent example is the Agreement on Trade Related Aspects of Intellectual Property (TRIPs) in the WTO. Developing nations, which are primarily consumers rather than producers of intellectual property, were induced to agree to strengthen their intellectual property laws in ways that would confer considerable rents on foreign rights holders in exchange for concessions on other trade issues such as textiles and agriculture.2

It is too optimistic to imagine that monetary sidepayments and issue linkage will always eliminate the problem of an empty core. Take the monopoly hypothetical above and consider the monetary sidepayment option—if an anti-monopoly agreement is to lie in the core, the consuming state must make a monetary payment to the state with the monopolist that induces it to give up its monopoly rents. If the state with the monopolist maximizes national economic welfare, such a payment must be equal to the monopoly profit rectangle plus some share of the deadweight loss triangle. That may leave precious little gain to the consuming nation in relation to the transaction costs of negotiation. As for issue linkage, an expansion in the scope of negotiations inevitably increases their cost, and draws in a greater number of domestic political constituencies with an interest in the outcome. The resulting increase in the costs of international negotiation and of domestic political deliberation may be great enough to undermine the process. Nevertheless, many international agreements display a great deal of issue linkage in their

---

2Although I use TRIPs as an illustration of issue linkage here, I do not mean to imply that TRIPs standing alone necessarily enhanced global welfare conventionally defined. Here, as in many other settings, one must distinguish carefully between political optima that maximize the interests of parties to negotiations, and conventional welfare optima.
formation, as prominently illustrated by the WTO and NAFTA. The costs and benefits of issue linkage bear importantly on the wisdom of drawing other subjects into their ambit (such as competition policy).

Moving beyond the problem of ensuring that an agreement lies in the core, the transaction costs of reaching an international agreement turn importantly on the complexity of the issues to be addressed. Some agreements entail reasonably simple commitments (such as the ban on nuclear testing in the Nonproliferation Treaties). Others require highly complex commitments spanning a wide array of issues. The WTO is again illustrative. To make its central commitments on tariff reductions valuable, signatories must disable themselves from turning to substitute instruments of protection. The result is hundreds of pages of treaty text addressing quotes, import licensing restrictions, balance of payments policy, domestic tax and regulatory policy, state trading and monopolies, and many other subjects. Complexity can be said to increase not only as the detail involved in specifying the proper behavior of a state increases, but also as the optimal behavior across states becomes more variable. Both dimensions of complexity raise the costs of agreement not only because desired behavior becomes more costly to describe and memorialize, but also because enforcement may become more difficult as the number and variety of possible violations multiplies.

A third important factor affecting the transaction costs of reaching and enforcing an international agreement is the number of countries involved in it. Even holding constant the complexity of the commitments (which of course may increase as more states participate), greater numbers of states add to the negotiation costs of reaching agreement. Logistical costs increase because of the larger numbers, and delays may develop as states hold back on what they offer to achieve agreement hoping to free ride on inducements offered by other states. A free rider problem may also manifest itself in the enforcement mechanism. Imagine that breach of agreement requires some action by nonbreaching parties to punish the party in breach. If these actions are costly, each state may prefer that others do the punishing, and unless some coordination mechanism can be designed to overcome the problem the threat of punishment may lose credibility. This observation suggests an important distinction between situations in which punishment requires actions that the punishers view as costly to themselves, and situations in which the opportunity to punish another state is welcomed. An example of the first may be military force. A state that employs military force risks the lives of its troops and consumes monetary resources, and most would no doubt be happy to have others undertake the task if they can perform it as well. By contrast, imagine an environmental agreement limiting emissions of some pollutant in each state, enforced by a threat of
mutual defection. Here, should another state defect, a punishing state might well benefit economically from the opportunity to defect itself and there may be no free rider problem in enforcement.

These last observations suggest some further principles. If externality problems can be handled adequately on a bilateral basis, such arrangements will tend to be preferred to multilateral arrangements. The case for agreements involving larger numbers of states maps directly onto the case for international agreements in the first instance—multilateral cooperation is potentially desirable mainly when bilateral cooperation itself creates externalities for third states. The caveat is that if a multitude of bilateral agreements would all look about the same, it may be more economical to create one multilateral agreement that sets out the same principles in a single document.

A related point is that when bilateral or small numbers agreements create externalities, an expansion of the number of states participating in an agreement will not simply increase the costs of agreement, but also the benefits. Hence, the international community may move toward large scale multilateral agreements despite their higher cost.

Finally, not only does the number of states involved affect the costs of achieving and enforcing agreement, but the size distribution of states may also matter. In particular, the presence of a few large states may help at least on the enforcement problem. If punishment is seen as costly to the punishers, the free rider problem will be less acute when some states are large enough to capture a considerable portion of the joint gains from enforcing the agreement. It may then be in their private interest to act as enforcers even if smaller states will free ride. This possibility suggests a constructive role in some contexts for what political scientists term a “hegemon,” a powerful state that enforces its will in an environment populated by other smaller and less powerful states.

**ii. Treaties as Contracts**

As noted, one may model international agreements as an effort by states to reach their Pareto frontier in the presence of externalities. They do so subject to a participation constraint for each state and perhaps other constraints depending on the context.

Readers familiar with the literature on optimal contracting will find this type of problem quite familiar. Indeed, from an economic standpoint, a treaty is a contract albeit one between states rather than individuals. As long as one is prepared to reduce the “preferences” of the state to an ordering akin to that of an individual rational actor, there is no difference in a general sense between the problem of designing an optimal contract
and the problem of designing an optimal treaty. I briefly note here some of the issues that have received some attention in the international literature, and that are directly related to the literature on optimal contracts.

**Treaties as Incomplete Contracts.** Some treaties address simple matters about which optimal behavior changes little over time. But many treaties address complex matters in an environment subject to uncertainty. For familiar reasons, treaties in the latter group are likely to be incomplete as to certain behaviors and contingencies. States may then benefit from various devices to fill the gaps in their agreement. Specialized tribunals may supply useful default rules in some cases, as may overarching treaties on interpretation such as the Vienna Convention on Treaties. As in the literature on default rules for private contracts, one can ask what the guiding principle should be in designing them—should they replicate what the parties would have negotiated expressly if they had addressed the matter, penalize a party that has withheld pertinent information, or something else?

**The Contractability of Desired Behavior.** It is a commonplace in the literature on optimal contracting to suppose that some behavior is noncontractable, usually because the behavior in question cannot be observed by other parties, or at least cannot be verified by a court. The same class of problems arises under many international agreements regarding everything from hidden nontariff trade barriers to cheating on nonproliferation commitments. Treaties may then employ the familiar types of (generally second-best) solutions. Treaty obligations may be conditioned on verifiable behavior that is correlated with the unverifiable behavior, as in the classic principal-agent models where payoffs must be conditioned on observable outcomes rather than unobservable effort. Alternatively, provisions may be designed to encourage decisionmakers to internalize the externalities from their choices, and thereby to eliminate the divergence between privately and socially optimal behavior.

**Renegotiation, Modification and Efficient Breach.** Treaties are often negotiated under conditions of uncertainty. A variety of shocks may cause particular commitments to become inefficient, or may leave some signatory worse off than it would be by exiting. Some treaties will address the problem simply by providing for the possibility of withdrawal, perhaps after a period of notice (the SALT treaty, for example). But many treaties address a broad range of issues, and changed circumstances may justify only a modification of the bargain, not a complete end to it. Accordingly, treaties may contain provisions providing for the renegotiation of parts of the bargain, or may specify contingencies under which states may deviate from their prior commitments. A treaty
may provide that a party can breach its obligations at a price, which if set correctly can facilitate “efficient breach.”

The design of such provisions can raise a number of interesting questions. In a multilateral treaty, for example, how does a renegotiation avoid the holdout problem? Does that problem argue for allowing breach at a “compensatory price?” How is the price for breach set, and when is it preferable to employ stiffer sanctions that move the regime toward a “property rule” and away from a “liability rule?” These latter questions presuppose that the legal regime has the capacity to employ meaningful sanctions for breach and to calibrate them—a subject that will receive further attention in a moment.

In addition to the problem of facilitating efficient adjustments to changed circumstances, treaties must confront the fact that efforts to modify the bargain may be opportunistic. In the event that a party deviates from its commitments and then offers to renegotiate, will it be in the interest of the other parties at that point in time to hold it to the original bargain and will they have the ability to do so, or will they capitulate to new terms that extract some of their surplus? In many important contexts, therefore, a treaty designer must worry whether the agreement is renegotiation proof. This problem is usually related to the presence of sunk costs. After states have made sunk investments in reliance on an agreement, the returns to those investments may be vulnerable to expropriation during opportunistic renegotiation.

This is but a partial listing of the issues that international agreements must confront, and that have direct parallels in optimal contracting problems. Economically-oriented scholars will find many fruitful applications to international law of the ideas developed in the contract literature.

iii. Enforcement and Dispute Resolution

States contemplating an international agreement often confront another set of problems that is absent for parties to conventional contracts. Private actors commonly rely on state enforcers to hold them to their commitments. An award of damages in contract can be enforced through a seizure of the promisor’s assets or an injunction backed by threat of imprisonment. By contrast, although the use of military force and the seizure of assets is surely seen in international relations, such devices are not part of the enforcement arsenal for many international agreements. Indeed, the observation that much of international law is not backed by a credible threat of military force or other strong coercive measures has led many commentators to question whether international law is “law” at all.
Such skepticism is misplaced in my view for at least two reasons. First, states that join international agreements can and do create international regimes where strong coercive measures are possible—witness the deployment of troops after U.N. authorization through the years, or the various episodes of U.N. sanctions that have affected the economic vitality of rogue states. International law should not be viewed as hopelessly weak because of the limited enforcement regime that is imposed on it exogenously. Instead, one must recognize that the choice of enforcement regime is endogenous. There is much that states can do, if they wish, to make their commitments more credible. One must therefore ask why some international legal commitments carry much greater punishments for breach than others, and whether weaker regimes of enforcement are inadequate to their task or are instead chosen precisely because nothing more is necessary.

Second, although parties to private contracts can often appeal to state enforcers, in many settings they cannot. Litigation costs may swamp the benefits of an appeal to the courts when the monetary stakes are modest. Transnational contracts can confront difficult issues of securing personal jurisdiction in the event of a dispute, enforcing awards, and securing unbiased decisionmakers. Such problems have led economic scholars to recognize that there are valuable mechanisms for making contractual commitments credible that do not rely on the existence of a third-party enforcer with coercive powers. (See, e.g., Telser 1980) Even when these mechanisms cannot achieve the “first-best,” they can often accomplish a great deal. So too it is at the international level.

I thus offer a brief catalog of the mechanisms that are available for the enforcement of international agreements that do not rely on the coercive powers of third party enforcers. Some have been mentioned in passing already but deserve more detailed comment.

Hostage Exchange and Bond Posting. The paradigm example of the hostage exchange involves a hypothetical peace treaty between warring kings. Each king sends a son to live in the other kingdom, with the understanding that if he attacks that kingdom his son will be killed. Both kings value their sons’ lives more than any possible gains from aggression, and so peace is sustained.

The exchange of human hostages threatened with death is rather unseemly by modern standards, of course, and we do not observe it in international agreements. But analogous situations may arise, as when foreign nationals or their assets travel and become subject to the jurisdiction of other states. When such movements are reciprocal a situation akin to a hostage exchange may arise implicitly—recall the suggestion by
Goldsmith and Posner that the rules of ambassadorial immunity may be sustained through such a mechanism. A possible difficulty with the hostage exchange mechanism, of course, is that the reciprocal threats to the hostages may not be credible. It may not be in the interest of the king who is attacked to kill the attacker’s son, for example, if he knows that his own son will then be killed.

A related mechanism is bond posting. Each party posts a bond that is large enough to exceed its potential gains from cheating on the agreement. Should cheating occur, the bond is forfeit. Compliance with the agreement then becomes an equilibrium as long as cheating will be detected with sufficient probability in relation to the size of the bond. Bond posting differs from hostage exchange in that it generally relies on a third-party arbiter to determine whether breach of agreement exists, but the arbiter need not possess any coercive powers beyond the capacity to declare that one party has forfeited its bond to another. If the arbiter is trusted by all parties, and they cannot interfere with the exercise of the arbiter’s authority, then a bond posting mechanism can do quite well in encouraging compliance. Interestingly, formal bond posting arrangements seem quite rare in international law, although the reason why is by no means obvious.

**Mutual Threats of Defection.** International agreements that address externalities often have many of the qualities of a Repeated Prisoner’s Dilemma. Each party makes commitments that it would prefer not to make, other things being equal, in exchange for valuable commitments from others that result in a Pareto improvement. States will be tempted to defect if they think they can do so without retaliation.

Repeated Prisoner’s Dilemmas have been studied extensively, both theoretically and empirically. On the theoretical side, the well known Folk Theorem holds that long term cooperation is a possible equilibrium, enforced by mutual threats to defect from cooperation should another party defect. Cooperation requires that the game has no fixed ending (defection would become a dominant strategy in the last period and cooperation would unravel from there), and that the parties have low enough discount rates that the current gains from defection do not loom too large in relation to the long term gains from cooperation. Rasmusen (1989). A requirement that the equilibrium be renegotiation proof makes the equilibrium rather complicated, but such equilibria do exist. (Fudenberg & Tirole 1991). Empirical research suggests that the “tit for tat” strategy, whereby a period of defection by one party is met by a subsequent period of defection by others, can be a successful enforcement device, even if the theoretical literature notes that this strategy is not subgame perfect. Axelrod (1984). Although much of the discussion in the literature is
of 2-person repeated Prisoner’s Dilemmas, many international agreements represent the N-person variant. Here too, cooperation is a possible equilibrium but hardly the only one.

Regardless of the number of players, threats of mutual defection are more likely to sustain cooperation when defection by one party can be detected readily by others. It is also important that parties be able to agree on what constitutes defection, and hence that the rules of the game be clear. Otherwise, what one party claims to be a justifiable punishment for defection by another party may itself be viewed by others as opportunistic defection.

One important factor that aids cooperation in international agreements is absent from classic models of the Prisoner’s Dilemma. Under the usual assumptions about the way the game is played, the parties cannot communicate with each other during a “period” of play, and can only discover what the other party has done after each has made a simultaneous move. Parties to international agreements, by contrast, can remain in close communication with each other at all times. Depending on the context, any movement toward defection may be easy to detect. Indeed, any plans for defection may be public knowledge and even the subject of public debate long before defection actually occurs. Consequently, defection may become punishable more quickly. From an analytic standpoint, improved communication shortens the “periods” during which the repeated game is played, and has a tendency to reduce the short term gains from defection.

*International Sanctions (Unilateral or Multilateral).* Any punishment for breach of an agreement may be termed a sanction, but here I use the term more narrowly. Define “sanction” as a costly measure, other than retaliatory defection, taken by a state aggrieved by breach of an agreement. For example, if state A violates its obligations under a Nonproliferation Treaty, state B might impose a sanction in the form of a suspension of trade relations. Sanctions may be undertaken unilaterally, or in coordinated fashion by a number of states.

So defined, a key feature of sanctions is that they impose costs on the states that use them, viewed in isolation. The strategic setting thus differs importantly from that of a Repeated Prisoner’s Dilemma, in which retaliatory defection can benefit a state after another has already defected (at least in the 2-person case). The question whether to carry out a sanction under these conditions is sometimes termed the “Punisher’s Dilemma.”

Sanctions too have received considerable study, both theoretical and empirical. Eaton and Engers (1992) model sanctions in a two-country framework with one country as the “target” of sanctions and one as the “sender.” The sender makes a demand on the target, which can then balk or comply. If the target balks, the sender can impose a sanction that is costly both to itself and to the target. They show that subgame perfect
equilibria exist in which the threat to impose the sanction succeeds and the target complies as long as discount rates are not too high. The sender’s threat to impose the costly sanction is credible, they argue, because a reputation for toughness is valuable to it in the future. The equilibria that sustain compliance are not renegotiation proof, however, as following an incident of balking it would be in the two countries interest to forget about past transgressions, so that in the simple case the only renegotiation proof equilibrium is unpunished balking.

They then consider a more complex case in which compliance is a matter of degree. Renegotiation proof equilibria then do emerge in which the threat of sanction induces some level of compliance. The degree of compliance will be greater, other things being equal, the less costly the sanction is to the sender. A high degree of compliance can be exacted if, for example, the sender can extract reparations from the target to cover the cost of the sanction. Further complications arise when there are more than one “sender” country. Depending on the distribution of costs and benefits from acting, some countries may free ride on the sanctions efforts of others, and it is possible that no country will take action.

Chang (1995) considers a somewhat different set of issues—the choice between sanctions and bribes, or “sticks” and “carrots.” He develops his analysis with particular reference to global environmental issues, and to the ongoing debate over whether developing nations should be compensated for measures that create positive externalities or that avert negative externalities (such as preserving rain forest or protecting sea turtles). The analysis has broader applicability, however, and bears on any situation where nations are choosing between bribes and sanctions to induce cooperation. Chang’s essential concern is that in any regime where “carrots” are offered, nations may react strategically to extract larger “carrots.” Drawing from the game-theoretic literature on predatory pricing, Chang develops a signaling model in which states that offer bribes are imperfectly informed about the potential recipients’ private optima in the absence of a bribe. There are two “types” of potential recipients, those whose private optima in the absence of a bribe are relatively close to the desired cooperative behavior (good types), and those whose private optima in the absence of a bribe are relatively far from the desired cooperative behavior (bad types). They all signal their type by engaging in some level of the behavior in question (cutting down rain forest, for example). A possible outcome is a pooling equilibrium in which the good types mimic the bad types to extract larger bribes. Such a development is potentially unfortunate, of course, because it raises the costs of securing cooperation and may lead to perverse signaling behavior that exacerbates the underlying problem.
Turning from theory to empirics, the historical efficacy of sanctions has also received considerable study. Any such study must come with an important caveat—we cannot observe sanctions that were never imposed because the mere threat of them achieved desired compliance all along, and we cannot observe cases in which sanctions were not imposed because they were perceived to be futile. Cases where they are actually employed, therefore, must represent intermediate instances where states were uncertain about their efficacy or were simply building a reputation for toughness. A notable study is that of Hufbauer, Schott and Elliott (1990), who reviewed the use of sanctions in 116 cases since World War I. They found that sanctions succeeded about one-third of the time. They were more likely to succeed, \textit{inter alia}, when the policy change sought by sanctions was relatively modest, when the sanctions were more costly to the target, and when the sanctions were less costly to the sender. A more thorough review of the theoretical and empirical literature on sanctions may be found in Eaton and Sykes (1998).

\textit{Reputation.} Analytic accounts of reputation rely on the notion that actors have different propensities to cooperate with others, and that their “type” in this respect is private information. Actors prefer to enter cooperative arrangements with other actors who will honor them. An instance of defection from cooperation by an actor thus conveys useful information to other actors about their likely type; those who defect will be judged less reliable in the future, and will lose opportunities for cooperation. When future gains from cooperation are important and discount rates are low enough, the desire to maintain a reputation for cooperating can lead actors to do so when they would otherwise be tempted to defect. Baird, Gertner and Picker (1994).

The role of reputation in policing international agreements is controversial. Most commentators acknowledge that it likely plays some role, although some imagine it to have a central role while others doubt that it amounts to much most of the time. In general, it is unclear to what extent reputation crosses over from one set of issues (say, trade) to another (say, military security). Even within an issue area, the notion that the propensity of states to cooperate is “private information” seems questionable, particularly in the more transparent Western societies. It is also unclear how well reputation survives through time, especially as political leaders change, raising the possibility of “end game” issues. These are the points on which the skeptics focus, and ultimately only empirical research can determine whether they are telling. I am unaware of any empirical work that isolates the force of reputational considerations from other factors.

\textit{Information Revelation Mechanisms.} All of the enforcement devices noted above rely on the ability of states to detect violations of international agreements. Sometimes violations are obvious, as when a nuclear test occurs within the territory of a signatory to
a Nonproliferation Treaty. But sometimes violations are surreptitious, and in other cases disagreement may arise over what constitutes a violation. In response, a number of devices may be employed to improve on the information possessed by parties to an international agreement.

If violations are surreptitious, a need may arise for investigators to examine suspected violations or to conduct routine inspections for compliance. Parties may place greater confidence in neutral investigators who are not linked to disputants, and a role for an institution with independent investigators may emerge. U.N. weapons inspectors are perhaps a useful illustration of this possibility.

Even in cases where pertinent behavior is observable by other states without investigators, a mechanism may be needed to resolve factual and legal disputes regarding alleged violations. None of the mechanisms outlined above can work well unless states know whether observed conduct is compliant or deviant. Here, a role emerges for a formal tribunal to conduct factual and legal analysis. Related, it may be valuable for violations to be publicized extensively. Tribunals and their decisions, if public, can also serve this role. Requirements for centralized reporting of violations, even in the absence of a tribunal, can serve a valuable publicity function.

D. The Interface between Domestic and International Law

A number of interesting issues arise regarding the relationship between domestic law and international law. Here, I will focus on two areas: the problem of making negotiators credible in the face of domestic political constraints on their power, and the process of conforming domestic law to international obligations.

Putting aside customary law, international law emerges only after a process of state to state negotiation. To reduce transaction costs, it is useful for each state to speak with “one voice” in the process. But it is not uncommon for power over the issues in question to be shared among many domestic actors. The United States is a good example, and I will use the trade area for concreteness. Although the President and his trade officials are the natural parties to negotiate on behalf of the United States, the constitutional separation of powers requires that virtually all changes in trade policy be approved by both Houses of Congress. Indeed, all tariff bills must originate in the House Ways and Means Committee. As a result, certain key legislators and committees have a great deal of power over trade policy, enough to block or bottle up trade agreements negotiated by the President if they wish.

Because of this problem, it is conventional wisdom that the President should not embark on trade negotiations without benefit of the so-called “fast track” authority. The
fact track amounts to a procedural rule, adopted by both Houses, promising that any proposed trade agreement submitted to Congress will be voted on within a fixed timetable (powerful committees cannot delay the vote), and will be approved or disapproved in its entirety with no opportunity to offer amendments. The last aspect of the rule ensures that Congress cannot condition its approval on some item that would require renegotiation.

The economics of the fast track are interesting. A possible explanation is that legislators expect trade deals to benefit themselves ex ante, but also know that ex post some powerful figures will likely wish to block approval. Because of this prospect, they know that agreements are unlikely to be approved without fast track, and indeed other countries may not bother to negotiate. They agree on fast track “behind a veil of ignorance,” uncertain as to who will end up regretting it ex post but secure in the knowledge that the expected benefits to those who vote for it are positive ex ante.

Although plausible, this explanation is not without difficulties. It relies heavily on the notion that uncertainty prevails ex ante as to which constituencies will suffer under a trade agreement. Yet, the negotiating agenda of other trading nations is usually quite well known before negotiations begin, and the industries likely to be imperiled by further trade liberalization are not terribly difficult to identify. The mystery then becomes why the fast track procedure does not simply cause the political bottleneck to emerge at the time of the fast track vote, when the powerful figures who would stand in the way of a final agreement instead make their stand against fast track authority. A possible retort is that the political costs to legislators who refrain from blocking fast track authority are somehow lower than the costs that they would face should they retain the power to block an agreement ex post and fail to exercise it. This explanation requires constituents to be ill-informed to a degree, and unable to discern that their plight ex post was highly predictable ex ante. Hence, it is also a somewhat uncomfortable explanation. Leebron (1995) offers a thorough discussion of fast track, accompanied by skepticism that it really makes much difference.

The fast track example is but one illustration of a broader class of problems that arises with respect to many types of international agreements—how to make the actors who negotiate on behalf of each nation credible so that their representations may be relied on, while ensuring that they act as faithful agents for their principals. This class of problems has received very little study by economically oriented scholars.

Another interesting aspect of the interface between domestic and international law concerns the status of international law in domestic legal systems. With respect to treaties, it is understood that they may be “self-executing” or “nonself-executing.” A self
executing treaty is intended by the signatories to be incorporated directly into their respective legal systems. As a result, the treaty is designed to become immediately enforceable in domestic courts—it may thereby create private rights of action, or trump prior inconsistent domestic law. A nonself-executing treaty is a binding international obligation but has no domestic legal effect \textit{per se}. A party to such a treaty can choose to give it effect in domestic law or not through further domestic legislation or legal action (of course running the risk that it may violate international law if its domestic law is not properly conformed).

A closely related distinction exists between “monist” and “dualist” states. A “monist” state takes the position that all binding international legal obligations are incorporated into domestic law automatically. In effect, international law is always “self executing,” even when other states do not treat it as such. A “dualist” state takes the position that international law has no direct domestic effect, except in particular cases where a mutual agreement exists that it should be self-executing. Absent such a case, the dualist state retains the right to implement the international obligation into its domestic legal system as it sees fit, again subject to the risks and consequences of violating international law. The United States, for example, is a dualist state. Except for rare self-executing treaties, international law becomes effective domestically only through additional (usually Federal) legislative enactments. By contrast, Costa Rica is a monist state, and even goes so far as to elevate international law above its own constitution.

At first blush, dualism seems to create a great danger of opportunism. Many international agreements require extensive changes in domestic law before states can bring their behavior into compliance. The dualist state may promise to make the necessary changes, but the task of monitoring compliance at all levels of government may be a daunting one for other states. The dualist state may then fail to behave as promised without detection for a long time, and one might think it better for international law to have direct domestic effect to eliminate the problem. This view of dualism is too harsh, however, for it neglects the fact that conforming domestic laws and practices is costly, whether accomplished through domestic legislation or through the application of new international law. And on many issues, strict compliance with international law may be of little moment for other states—U.S. trading partners may care little about whether Alaska conforms its building codes to the WTO Technical Barriers Agreement, for example, if the affected trading volume is nil or de minimis. Dualism allows states to incur the costs of converting their domestic laws only as they have reason to believe that other states care about the conversion, and may thus afford Paretian gains to the participating states. Whether these gains offset the problem of opportunism in any given
setting is an empirical question, and where the opportunism problem is obviously dominant the dualist states can avoid it case-by-case through the occasional self-executing treaty.

The harder question concerns the existence of some monist states. Why would any state subject itself to the direct effect of international law unless other parties to an agreement promise to do the same? The answer here may also lie with transaction costs. The costs of enacting domestic legislation to implement international obligations may be partially fixed, so that such costs represent a larger proportional tax on a small economy than on a large economy. At some point, the costs of domestic lawmaking may exceed the benefits of the selective deviation from commitments that dualism permits. The smaller state may then prefer to embrace the international law wholesale rather than to incur the costs of determining how it can benefit from a dualist approach. On the general phenomenon of monism and dualism, and the variety of intermediate approaches that emerge in various nations, see the collection of papers in Jackson and Sykes (1997).

4. Security Issues in International Law

This section addresses the international law of military conflict and strategic alliances. These subjects have been the subject of intensive study by political scientists, but relatively little work has been done by law and economics scholars until recently. The exception concerns work on collective defense as an application of the theory of public goods.

A. International Alliances

Olson and Zeckhauser (1966) offer an early formal treatment of “alliances” that supply defense services to their members as a public good. Various scholars have extended their work, primarily by treating defense as a mixed public/private good. Sandler and Cauley (1975). A recent survey of the literature both theoretical and empirical may be found in Sandler and Hartley (2001).

The essential issues may be captured in a two country model, which generalizes to N countries quite easily. Imagine two allied countries, one and two, each of which must decide how much of its resources to allocate to defense measures and how much to allocate to a numeraire good. Denote units of defense measures in each country by $d_i$. For simplicity, let the price of defense measures be $p$ in both countries. Each country’s consumption of the numeraire good is $y_i$, and the total income is $I_i$. The national budget constraint is thus $y_i + pd_i = I_i$. 


Defense measures by each country may benefit the other—a nuclear deterrent, for example, may discourage attacks on both nations. But defense measures may also have a rival component to them, as with measures to defend one’s own border against terrorist intruders. In this sense, defense measures yield a joint product. Let the rival or “private” component of defense be denoted $x(d_i)$, and the nonrival or “public” component be denoted $z(d_1 + d_2)$. The welfare functions of country one and country two are $U(\cdot)$ and $V(\cdot)$, respectively. Each welfare function has four arguments: consumption of the numeraire good $y$, the private component of defense $x(d_i)$, the public component of defense $z(d_1 + d_2)$, and a shift parameter $\Omega_i$, which may be interpreted as the magnitude of the perceived external threat.

Consider first the choices that each country will make, taking the choice of the other as given. It will suffice to examine country one’s problem, as country two’s problem is completely analogous. Country one will choose $y_1$ and $d_1$ to maximize $U[y_1, x(d_1), z(d_1 + d_2), \Omega_1]$, subject to its budget constraint. The first order conditions imply:

$$x'U_x/U_y + z'(U_z/U_y + V_z/V_y) = p$$

The first term on the left hand side is the marginal rate of substitution between the private component of defense and the numeraire good, multiplied by the marginal effect of greater defense measures on the private component of defense. It can be interpreted as the marginal money value of the additional private component of defense generated by an increase in country one’s defense measures. The second term may similarly be interpreted as the marginal money value to country one of the additional public component of defense generated by an increase in its defense measures. Country one will equate the sum of these values to the price of a unit of defense measures. Country two’s optimum is identical, restated using derivatives of its own welfare function $V(\cdot)$. Nash equilibrium occurs when each country’s choice of defense measures is optimal given the choice of the other.

The Nash equilibrium is not socially optimal, of course, because each country neglects the positive externality generated by its defense measures through the public component of defense. A social optimum may be found by choosing the $y_i$ and $d_i$ to maximize the welfare of country one, subject to the constraint that the welfare of country two achieve a fixed value, and to the joint budget constraint $y_1 + y_2 + p(d_1 + d_2) = I_1 + I_2$. The first order condition for a socially optimal choice of $d_1$ may be written:

$$x'U_x/U_y + z'(U_z/U_y + V_z/V_y) = p$$
This expression includes country two’s marginal rate of substitution between the public element of security and the numeraire good, and thus captures its willingness to pay at the margin for additional defense measures by country one. The socially optimal choice of $d_2$ must satisfy an analogous condition. The appearance of the sum of the marginal rates of substitution across countries in the conditions for social optimality is a standard type of result in the theory of public goods. Indeed, if one were to eliminate from the welfare functions the arguments $x(\bullet)$ and thereby eliminate the private component of defense, the model would be a trivial variant of a standard pure public goods model. Conversely, if one were to eliminate the $z(\bullet)$ arguments, the externality would vanish and Nash behavior would be socially optimal.

This type of model may be used to generate a number of implications. For example, if we consider the pure public good case for simplicity and make the conventional assumption that marginal welfare is diminishing in each argument, smaller defense expenditures by one country will induce larger expenditures by the other in Nash equilibrium and vice-versa. An increase in the threat parameter to country one, therefore, which increases its marginal welfare from defense, will lead it to increase its defense measures and lead country two to reduce them (the matter becomes more complicated when security has both a private and public element). Defense measures are also a function of income, of course, and changes in the relative income levels of the two countries will have similar implications. With sufficiently large income disparities, the countries’ reaction functions may not cross at all and the wealthier country may shoulder all of the joint defense burden while the poorer country free rides completely. Models of this sort have been offered to explain a number of empirical observations, such as why so much of the cost of nuclear deterrence was borne by the United States during the Cold War. The other key implications arise when we allow the private component of defense to become more important. The greater the extent to which defense measures yield rival instead of nonrival benefits, the lesser the tendency of allies to undersupply defense.

The reader will no doubt notice that while this model comes directly from the literature on “alliances,” it is in fact a model of equilibrium behavior in the absence of an alliance (aside from the fact the two countries are “allied” against some of the same threats). The role for a formal alliance in this context is to overcome the collective action problem that arises when defense measures have important positive externalities. Thus, having identified the externality problem, one must move on to ask whether defense policy in practice is importantly affected by it. If so, one must next ask what mechanisms can best abate it, and whether existing alliances have addressed it satisfactorily.
These are difficult questions. As to the first, it is hardly enough to observe that some types of defense measures generate positive externalities for allies because potentially offsetting factors exist. At least since the concept of the “military-industrial complex” was introduced by Eisenhower, a public choice perspective on defense questions whether interest group politics may bias defense expenditures upward. It is thus difficult to say whether the level of defense spending is too high or too low in many contexts, either on an individual basis or across some specific alliance. The judgments of individual commentators about the adequacy of defense often depend on varying and highly uncertain estimates of the magnitude of external threats.

Assuming that members of an existing or potential alliance can agree that their collective spending is too low, however, the opportunity for them to negotiate over their respective contributions may bear fruit. Each nation can offer to contribute more in return for greater contributions by others. Although the incentive to free ride and hold out during negotiations will no doubt present itself, it is possible that mutual commitments to increase contributions will arise and be respected. A useful illustration may be drawn from the history of funding for U.N. peacekeeping. In the early days, the organization would solicit voluntary contributions, which were predictably small and few in number. As peacekeeping activities increased through the years, the General Assembly in 1975 passed a resolution that assessed each member an amount for peacekeeping based on a formula that considers national income, Security Council membership, and other factors. Some nations (including the United States) have neglected to pay their assessments at times, but the system has generated considerably more revenue than one could have expected from the system that it replaced. And because the “optimal” contribution by each nation is likely unknowable, such crude, formulaic mechanisms may do about as well as can be done.

To be sure, such an approach may also fail badly. To the degree that promises to contribute more are enforced by mutual threats to withhold contributions in the event of defection, alliance members may doubt their credibility. For example, members may recognize that in the event of a funding shortfall, those with a substantial private interest in the outcome of a particular conflict may be led to undertake collective defense unilaterally or in a smaller coalition. It may even be possible for would-be free riders to avoid defecting on their promises simply by blocking the alliance from acting in a given conflict, forcing members with high stakes to strike out on their own.

The notion that smaller states will free ride on the efforts of larger states, and the suggestion that U.S. allies in particular free ride on U.S. defense efforts, has long been a theme in strands of the literature on alliances. Convincing tests of the hypothesis are
again complicated by the fact that each country’s proper share of collective defense in the absence of free riding is difficult to determine at best. Olson and Zeckhauser (1966) posited that military expenditures as a percentage of GDP ought to be approximately constant in the absence of free riding. They then presented evidence of a significant positive correlation between GDP and military expenditures as a percentage of GDP in their dataset, supporting the free riding hypothesis in their view. More recent evidence using the same approach suggests that free riding has diminished with time, especially following the Cold War, as might be expected—nuclear deterrence provided substantial nonrival benefits to allied states, while the more modern concern for small scale conflicts and terrorism has led states to shift their spending toward conventional measures to protect their own territories and nationals.

B. The Laws of War

The international law of war may be divided roughly into two components: rules on the use of force (jus ad bellum) which govern the initiation of hostilities and the response to them, and rules governing the conduct of war (jus in bello) such as the customary law and conventions on the treatment of prisoners of war. As with many subjects in this area, very little has been written on either set of rules by economically oriented scholars.

Goldsmith and Posner (2004) address jus in bello. They posit that war can be conceptualized as a two-stage game. At the first stage, states are uncertain whether they will go to war. They recognize that war is destructive and that they would be better off if they could reach peaceful settlements that reflect the outcome of future war without the need to bear its costs. But like unions and management, they realize that some destructive conflict may occur. Accordingly, they design rules for the conduct of war with the goal of minimizing the joint losses should it happen. Whatever the outcome of possible war, both sides expect to gain if the number of casualties is held to a minimum along with the damage to their economies. To this end, they adopt rules such as those requiring prisoners to be treated humanely rather than killed, and those prohibiting unnecessarily destructive technologies like chemical and biological weapons.

The rules are enforced by mutual threats of retaliation should one side deviate from them during a war. In some instances, this enforcement structure works—the fact that all sides generally refrained from the use of poison gas during World War II may be a good illustration. Any party could have readily manufactured gas and retaliated in the event of deviation. But it may also break down, as illustrated by the highly variable
treatment of prisoners of war in various conflicts through the years. When the number of prisoners is large and the costs of adhering to rules of humane treatment become great, for example, Goldsmith and Posner suggest that deviation from the rules is common. Likewise, smaller powers with limited financial resources are considerably less likely to comply with international rules for the treatment of prisoners.

Posner and Sykes (2004) consider the modern rules on *jus ad bellum*. Under the U.N. Charter, to which virtually all important states are signatories, the use of force is allowed only pursuant to a resolution of the Security Council that authorizes force, or in self-defense. Acceptable cases of self-defense are limited to situations involving an actual or “imminent” attack. In apparent contravention of these principles, the United States has indicated that it will consider the use of preemptive force in self defense under some circumstances. Posner and Sykes focus on the welfare implications of preemptive attacks in self defense. They offer a simple two-period model in which a potential aggressor state has private information about its future plans to attack another more powerful state, and cannot reveal it credibly. In period one, the threatened state can choose to attack preemptively or wait. A preemptive attack will eliminate the threat from the potential aggressor, but may do so unnecessarily if the potential aggressor would never have attacked anyway. Absent a preemptive attack, the potential aggressor state will reveal its intentions by attacking or not in period two, and if attacked the threatened state will respond with force. The analysis assumes that the threat of retaliation in period two is inadequate to deter attack, either because its magnitude is insufficient to deter or because the potential aggressor does not care about retaliation (perhaps its leaders expect to escape, or are undeterable zealots). The model suggests that if the costs of conflict in period two are sufficiently greater than in period one (as, for example, when the potential aggressor state develops powerful new weapons for period two), and if the probability that the potential aggressor will turn into an actual aggressor is sufficiently high, then preemptive attack may be both privately and socially optimal—the expected discounted costs of conflict computed under either welfare criterion may be lower with preemptive attack. The threatened state will only use preemptive force in accordance with its private interest, however, which may lead to too many or too few preemptive attacks. The latter possibility arises because collateral damage to foreign innocents can be much larger if the threatened state waits until period two to take action. An extension of the model to three periods illustrates how delay in period one has further option value, but does not change the essential points above. In this simple framework, the traditional rule against preemptive measures appears questionable.
Several objections to this conclusion may be offered. First, nothing in the model distinguishes “good guys” from “bad guys.” The analytics are the same whether the more powerful state is viewed as the United States and the potential aggressor state is in league with Al Qaeda, or the threatened state is Stalinist Russia and the potential aggressor is a repressed ethnic state forced into the Stalinist orbit. Indeed, repressive regimes may face growing threats against them more often than liberal regimes. A rule permitting the use of preemptive force may then countenance behavior that is considered undesirable on average, for reasons “outside the model.” Related, if the danger to the more powerful state is not imminent but a preemptive attack is genuinely desirable, perhaps the threatened state has the time and should be able to persuade the Security Council to authorize preemptive force. The vote of the Council then serves the function of identifying the “good guys” and allowing them to proceed. The difficulty with this argument is that members of the Security Council may vote (or exercise their veto) on the basis of their private interest in the potential conflict rather than the social interest, and it is an empirical question whether the repeat play nature of Security Council interaction can induce cooperation on the pursuit of social optima. The asymmetry of threats faced by its members may offer reason to doubt this hopeful scenario. A final possible defense of the traditional imminence requirement is that the conditions under which preemptive force is optimal are difficult to verify, and if nations have a right to use it they may invoke that right opportunistically. If preemptive attack is usually undesirable from a social standpoint, therefore, a bright line rule against it may be the best option.

5. International Trade Law

I now turn to the subject that has received by far the most attention in the existing law and economics literature. As noted previously, my emphasis will be on topics and ideas that suggest broader lessons for the study of international law, although some familiarity with the traditional economic literature on trade is essential as background to what follows.

A. Trade Policy, Trade Externalities and Trade Agreements

The normative economics of international trade suggests that government intervention in trade flows generally reduces global welfare—one of the few propositions, according to Paul Samuelson, on which almost every economist will agree. But government intervention in trade has been extensive throughout modern history,
raising a puzzle as to why and motivating a vast economic literature on the positive economics of trade policy. Early work focused on the incentives facing “large” countries that act as national welfare maximizers, understanding “large” to mean simply that the demand for the nation’s exports, or for the supply of its imports, is less than perfectly elastic. Such countries have monopoly power over their exports, and/or monopsony power over their imports. Competitive export industries and small consumers of imports cannot organize to exploit that power, but the government can exploit it through taxation. At appropriate import and export tax rates, national welfare will rise because some of the tax revenue is extracted from foreign exporters and consumers. Tariffs set on this basis are sometimes termed “optimal tariffs,” and their contribution to national welfare is conventionally said to arise through their effects on the “terms of trade”—the ratio of the price of a nation’s exports to the price of its imports. Optimal export taxes improve the terms of trade by increasing export prices (inclusive of taxes) relative to import prices; optimal tariffs improve the terms of trade by lowering import prices (net of tariffs) relative to export prices. A classic early treatment is that of Johnson (1953).

Economists quickly realized that these simple models of national income maximization did not explain the pattern of trade intervention very well in practice. Among other things, the pattern of tariffs across industries did not seem closely related to the market power of the importing nation. And on the export side, nations seemed more likely to engage in policies such as export subsidization than export taxation, a behavior that cannot reflect national welfare maximization in competitive markets. Economists began to examine different models of what governments maximize, generally drawing on ideas from the contemporaneously burgeoning literature on public choice. A wide range of modeling strategies was employed, all capturing in one way or another the idea that governments behave as though they care about the distribution of national income as well as its magnitude, usually because some interest groups are better organized than others and will reward their leaders more generously for pursuing policies that benefit them. A good illustration of this approach is the model developed by Grossman and Helpman (1995a). They posit that some domestic industries are organized and some not, and that the organized industries commit to schedules of campaign contributions which are a function of trade policy choices by their government. The government maximand is a weighted average of campaign contributions and per capita economic welfare. The model

---

3Under conditions of imperfect competition, however, export subsidies may emerge from the pursuit of national welfare by individual states, as suggested by the considerable literature on “strategic trade policy.” For a survey see Brander (1995).
has a wide range of implications, and makes progress with respect to the apparent deficiencies of the national welfare maximization models. For example, when governments behave as Nash actors, equilibrium trade policy involves higher tariffs for well-organized import competing industries than national income maximization would predict. It also involves lower export taxes on organized export industries, even to the point that export subsidies may arise. Rodrik (1995) provides a valuable survey of the work in this area.

Although the primary focus in much of the literature is on domestic policy formulation, many of the models have immediate implications for the role of international cooperation. The general equilibrium tradition in trade theory necessitates the presence of at least two countries in most models, and the noncooperative equilibrium of the models generally reveals externalities. The externality almost always arises because of the terms of trade effects noted earlier—part of the cost of trade intervention is borne by foreign producers and consumers. The failure of governments acting unilaterally to account for the external costs and benefits of changes in the terms of trade leads to equilibrium trade policies that are off the Pareto frontier. As in most other areas of international law, therefore, the role of international agreements is to overcome the externality problem.4

Some commentators are skeptical of the notion that terms of trade effects are the engine of international cooperation on trade policy. This skepticism arises in part because trade policy officials rarely discuss them explicitly, and may not even understand them as a reflection of each state’s market power. Bagwell and Staiger (1999 and 2002) argue forcefully, however, that terms of trade externalities can and should be interpreted as the basis for the “market access” demands that pervade and motivate modern trade negotiations. Exporters seek negotiated reductions in barriers to trade precisely because those barriers affect their net prices on world markets.

Having identified the likely source of externalities, the next question is how trade agreements will address them. If governments are national welfare maximizers, then an

4I note briefly two strands of literature in which the terms of trade externality is not the dominant reason for trade agreements. Ethier (2000 and 2004) asks the question whether countries can benefit from trade agreements due to some “political externality” that is distinct from any terms of trade effects. A somewhat more substantial strand of literature suggests that trade agreements are not motivated by international externalities at all, but by dynamic consistency problems in domestic politics. In these models, governments are imagined to have some preferences over investments in various sectors, but producers know that once investments are sunk, government may change policy in a way that undermines the returns to investment. In anticipation of this behavior, producers will invest in such a way as to minimize their exposure to it, distorting investment away from the preferred pattern. Government may thus benefit from commitment devices, and it is conceivable that a commitment to free (or freer) trade might be valuable given appropriate governmental “preferences.” See, e.g., Maggi and Rodriguez (1998).
optimal trade agreement should maximize global welfare conventionally defined. Under competitive conditions, free trade will do so, although free trade may not lie in the core and hence some transfer mechanism may be necessary. The allocative effects of free trade can be achieved with transfers in the background if the nations that are offering transfers subsidize their exports and the nations that are receiving transfers impose an equal and offsetting tax on them.

Of course, actual trade agreements do not achieve completely free trade (or its allocative equivalent with transfers). An adequate theory of trade agreements must thus explain how they deviate from that benchmark and why. Grossman and Helpman (2002), Bagwell and Staiger (2002), and numerous others offer political economy models in which governments care about the distributional consequences of trade policy, and thus in which the equilibrium trade agreement does not achieve free trade. The common thread is that cooperation will seek to eliminate the terms of trade externalities across nations, but will at the same time allow governments to respect their preferences over distribution. Well organized import competing industries will retain protection in these models, for example, especially if the foreign exporters with whom they compete are poorly organized.

In sum, mainstream international economics, including many of its most prominent practitioners, has developed an impressive set of tools for modeling the externalities that arise when nations act noncooperatively. These models offer a number of predictions that are broadly consistent with actual trade agreements. They suggest that terms of trade externalities lead tariffs to be excessive relative to those on the Pareto frontier, and that cooperation will thus tend to reduce tariffs in general. They suggest that governments motivated by distributional considerations as well as aggregate welfare will not achieve free trade through trade agreements, but instead will tend toward agreements that protect and/or subsidize the interest groups that are better organized (or whose welfare counts more in the view of their governments, for whatever reason). They are also capable of explaining in broad brush terms the evolution of trade agreements over time, particularly the gradual reduction of trade barriers through a series of negotiating rounds (intuitively, organized resistance to further liberalization diminishes after each round of liberalization as industry specific capital in import-competing industries depreciates without replacement). See Staiger (1995b). I will not dwell on these general points any further, nor will I include formal treatments of the underlying models given the readily accessible surveys elsewhere (especially in the Handbook of International Economics.) Instead, I turn to work on some of the particular legal issues that arise in the trade area and that bear on some of the more general themes developed in Section 3.
B. The Legal Architecture of World Trade and Its Lessons for International Law

Formal trade agreements arose in the 19th century, but I focus here on the modern WTO/GATT system. GATT began in 1947, and was limited to trade in goods. The centerpiece of the GATT was its reciprocal commitments to lower tariffs embodied in Article II, along with a general commitment to nondiscrimination among trading partners contained in Article I. The remainder of GATT served three primary functions: it constrained policy instruments that were substitutes for tariffs; it provided for various adjustments to the bargain including tariff renegotiation, amendments, waiver, accessions and preferential trade exemptions to the MFN obligation; and it provided a dispute resolution system.

GATT subsequently evolved through a series of negotiating “rounds,” which initially involved little more than further reciprocal tariff reductions. By the time of the Tokyo Round in the 1970’s, however, GATT signatories had become increasingly concerned about the growth of various nontariff barriers that were inadequately disciplined by the original GATT. The result was several plurilateral “codes” on such matters as subsidies, antidumping measures, product standards, and government procurement. The next major stage in the evolution of GATT came in the Uruguay Round, which lead to the creation of the WTO in 1995. The WTO replaced the GATT (although its treaty text was incorporated into WTO law as “GATT 1994”): it elaborated and tightened obligations with respect to nontariff barriers and made them binding on all members; it created the General Agreement on Trade in Services (GATS), for the first time bringing services trade under multilateral discipline; it created the Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPs), which required all members to create and enforce intellectual property rights in a number of areas; and it overhauled the dispute resolution system.

It is, of course, difficult to know what would have happened if the GATT system had never evolved, but conventional wisdom holds that it has been quite successful. Tariffs have declined steadily from an average rate of around 40% when GATT was founded to an average rate of about 4% presently. And although the history of the WTO/GATT system is not free of tension, it has so far avoided the problem of large scale trade wars that plagued the global economy before it, such as the widespread tariff increases during the 1930’s initiated by the Smoot-Hawley tariff in the United States. Member states also appear to comply with the rules of the new dispute resolution system. In a clear majority of cases where a measure has been held to violate WTO law, members
have abolished it. And although members have declined to abolish a few measures that have been ruled illegal, they have accepted the measured retaliation authorized by the system as an alternative.

However its success is measured, the WTO/GATT system is surely among the most intricate and complex international legal arrangements. Careful study of the system is important not only for trade scholars, but for anyone interested in the mechanisms of international cooperation.

i. Coping with Complexity: Identifying and Structuring Trade Protection

Tariffs were the principal instrument for the protection of import-competing industries historically, but many other instruments may be used for that purpose. Quantitative restrictions on imports are an obvious alternative to tariffs, and other available devices include discriminatory domestic taxation, subsidies, regulatory policies that burden imports disproportionately, state-franchised monopolies that disfavor imports in their purchasing decisions, and discriminatory government procurement policies. Effective agreements to reduce trade protection must attend to this array of protectionist options. The challenge of formulating appropriate constraints on all of these instruments is complicated by the fact that many of them also address “legitimate,” nonprotectionist objectives. Subsidies may correct other externality problems; regulations may address important issues of safety and quality; procurement preferences may affect national security; and so on.

These observations suggest that an agreement such as GATT must pursue a number of related objectives concurrently. First, signatories must ensure that commitments to reduce protection associated with one policy instrument are not undermined by the substitution of protection using some alternative instrument. Second, and closely related, signatories must manage the challenge of negotiating reductions in protection given all of the available instruments of protection, and will benefit from strategies that reduce the costs of negotiation. Third, to the extent that some instruments of protection produce greater deadweight costs than others, signatories will benefit if they can channel whatever protection remains under their agreement into the instruments that do the least damage. Finally, signatories must design agreements that constrain trade protection appropriately on the one hand, while leaving themselves free to pursue their domestic regulatory agendas on the other.

Regarding the first three of these objectives, the central strategy in GATT has been a process of “tariffication”—the conversion of most protectionist measures into
tariffs. This process is accomplished through a number of legal measures, such as the general prohibition on quantitative restrictions, import licensing schemes and the like in Article XI of GATT 1994. Bagwell and Sykes (2004) address the benefits of tariffication. They argue that it lowers the transactions costs of reciprocal trade negotiations directly by reducing the number of protectionist instruments that are part of the negotiation. It may afford additional benefits by reducing uncertainty about the degree of market access that trade concessions afford, an important consideration if trade negotiators are risk averse. Tariffs also have the virtue that when applied in nondiscriminatory fashion, they avoid the loss of joint surplus that results from trade diversion (discussed further below). Other instruments, such as quotas, may result in greater trade diversion depending on how they are administered or in other kinds of deadweight loss (see the discussion of regulatory protection below). Finally, because tariffs tend to be more transparent, cheating on tariff commitments is easier to detect than cheating with respect to commitments on other protectionist instruments. As long as the general commitment to tariffication is itself enforceable, therefore, tariffication discourages cheating and facilitates cooperation.

To say that it is valuable to channel protection into tariffs, however, begs the question of what measures constitute protection. Perhaps nowhere is this problem more acute than in the case of domestic regulatory policies—for example, is a prohibition on hormone-raised beef imports, ostensibly for health reasons, really a pretense for measures to protect the domestic beef industry?

The modeling framework developed by Bagwell and Staiger (2002) suggests a clever solution to such issues, based on the idea that trade agreements should induce nations to “internalize the externality.” The externality arises because choices by individual governments affect the prices confronting foreign buyers and sellers on world markets. If nations commit through trade agreements to maintain a particular level of market access—a particular terms of trade with other nations—they can then be allowed freedom to vary their domestic regulatory policies as they wish as long as any effects on the terms of trade are offset by countervailing changes in their tariff and subsidy policies. Through such a mechanism, other nations are insulated from the terms of trade consequences of domestic regulatory policies, the externality is “internalized,” and nations will then behave efficiently in relation to their internal welfare judgments. Bagwell and Staiger model this approach explicitly for labor and environmental standards and competition policy, but the strategy has general applicability as long as the only international externality is transmitted through the terms of trade.
In practice, however, trade agreements do not undertake to specify the terms of trade. Such undertakings may seem realistic in the static models of trade theory, but much less so in an environment of volatile exchange rates and innumerable other macro and microeconomic shocks to world prices. The WTO/GATT system does provide compensation for trade injury caused by certain measures, such as the withdrawal of tariff concessions and the introduction of unanticipated subsidy programs (each discussed further below), but it does not grant compensation for many other policies that affect the terms of trade. Instead, a key strategy of the system has always been to constrain protection \textit{per se}, while leaving signatories free to pursue nonprotectionist objectives such as health and safety regulation, environmental protection, and national security. But this strategy requires an additional set of rules to distinguish protectionist measures from nonprotectionist measures.

The stakes in drawing this line can be considerable. Sykes (1995 and 1999) compares regulatory protection to tariffs. Suppose, for concreteness, that an importing nation wishes to protect its grain industry, and that it has a target for the domestic grain price. Assume that this target price can be achieved with a tariff of $\tau$ per bushel of grain, which will limit grain imports to the quantity $B$, and result in tariff revenue to the importing nation of $\tau B$. Alternatively, the importing nation can enact a regulation that requires foreign grain producers to incur additional costs of production (for example, the cost of testing all export shipments for the presence of some ostensibly dangerous pest or chemical). Let the additional cost per bushel of grain exports be $\tau$ under this regulation, and assume that no tariff will be imposed in this alternate scenario. Assume further that the ostensible health or safety basis for the regulation is pretense, and that the sole motivation for the regulation is to achieve protection. From the standpoint of the grain industry in the importing nation, both approaches are equivalent in their ability to achieve the target domestic price. Likewise, both measures will have the same effect on the terms of trade, extracting surplus from grain exporters if the import supply curve is less than perfectly elastic. But the two approaches are not equivalent in welfare terms—under the regulatory approach to protection, the tariff revenue “rectangle” $\tau B$ is transformed into pure deadweight loss (given the absence of any bona fide health or safety justification). Wasteful regulations of this sort may hold no appeal to importing nations that are unconstrained in their tariff policies, for they would prefer to achieve the target level of protection \textit{and} to enjoy the tariff revenue. But if tariffs have been constrained through reciprocal trade negotiations, the temptation to engage in wasteful regulatory protection can emerge. Of course, if the importing nation cares somewhat about its national income, regulatory protection will be less attractive than tariff protection, and one would expect
only a partial substitution between regulatory measures and tariff measures. The important point is that if trading nations can substitute regulatory protection without penalty under the law, they will be tempted to do so to some extent. And the Nash equilibrium of the game in which nations make these substitutions can entail considerable welfare costs. It is in the mutual interest of trading nations to prohibit such measures—it is easy to see that by comparison to any situation with regulatory protection, affected importing and exporting nations could eliminate it, adjust their tariffs in some fashion, and make themselves all better off.

With particular reference to domestic regulatory policy, the initial response of GATT to this problem was to prohibit discriminatory regulations through a “national treatment” obligation in Article III. No regulatory burden could be imposed on imports unless it was also imposed on domestic producers of “like products.” GATT also provided a number of exceptions to this principle in Articles XX and XXI, which cover things such as measures “necessary” to protect the environment or measures “considered essential” to national security. To the degree that the national treatment obligation was limited to facially discriminatory regulations, however, it soon proved inadequate. Consider a “nondiscriminatory” regulation that requires all products to embody a certain design or technology, when domestic producers have a cost advantage in producing or obtaining that design or technology, and when some other design or technology available more cheaply abroad will achieve the nonprotectionist regulatory goal just as effectively. The cost difference between the two technologies then represents pure regulatory protection. The need for additional disciplines to deal with and related issues culminated in two important new agreements during the Uruguay Round, the Agreement on Technical Barriers to Trade and the Agreement on the Application of Sanitary and Phytosanitary Measures (pertaining mainly to pests, disease-causing organisms and contaminants in foodstuffs). These agreements impose numerous additional constraints on domestic regulators, such as general least restrictive means requirements and a requirement that product regulations specify performance goals rather than design limitations. Somewhat more controversially, they also introduce requirements that certain health and safety regulations have adequate scientific basis. To return to the example of a prohibition on hormone-fed beef imports, a European regulation to that effect was found to violate WTO law because it lacked scientific foundation. For further discussion see Trebilcock and Soloway (2002).

Where domestic regulations appear to have bona fide, nonprotectionist goals, however, the primary constraint on them remains a simple national treatment requirement. Such a rule tolerates changes in domestic regulatory policy even if they
have significant trade externalities, but it is not clear that any alternative rule is better. Battigalli and Maggi (2003) develop a model of international agreement on product standards, in which negotiators cannot anticipate proper standards for future products. The model suggests that only a national treatment requirement is joint welfare enhancing \textit{ex ante} for arbitrary distributions of future products.

Another important body of WTO law concerns the use of subsidies. Most of the work that has been done on this subject is normative, asking whether WTO rules on subsidies promote global economic welfare conventionally defined. Early writers observed that subsidies to domestic industries can protect them just as effectively as tariffs, yet subsidies may also be justified for reasons other than protectionism. A question then arises whether rules can be developed to distinguish “good” subsidies from “bad” subsidies, avoiding the welfare loss from the former while preserving the welfare gains from the latter. Schwartz and Harper (1972) are deeply skeptical of this enterprise, arguing that workable rules for the identification of welfare-reducing subsidies are impractical, particularly if one allows that “legitimate” subsidies may include those where the citizenry is willing to pay to preserve certain forms of inefficient enterprise (such as cultural industries or family farms).

Nevertheless, the WTO system has evolved increasingly detailed rules governing subsidies. In goods markets, the governing principles are three: (1) New and unanticipated subsidy programs that upset market access expectations are illegal; (2) export subsidies (subsidies that favor exportation over domestic production) are illegal (putting aside some exceptions in the agriculture sector); and (3) domestic subsidies, which are identified by the criterion that they must be targeted to a single industry or narrow group of industries, may be illegal if they cause “injury” to foreign exporters or import competing industries, and may also be a basis for unilateral countervailing duties (duties to offset the value of the subsidy) by importing nations whose industries are injured by subsidized imports. Sykes (2004) evaluates these rules from the traditional welfare economics perspective. Putting aside the difficult question of what constitutes a “subsidy,” the first principle seems defensible as a way to protect the value of the bargain associated with tariff concessions against opportunistically erosive. The second principle also makes sense from the perspective of traditional welfare economics, as it is difficult to imagine any constructive role for export subsidies as a first-best policy. But the third principle is problematic in that it defines “subsidies” using a dubious criterion—the targeting criterion may well condemn useful subsidies, and is quite underinclusive as to economically wasteful subsidies or as to subsidies that may have harmful external effects. Further, regardless of the degree of targeting, it is not clear that coherent criteria can be
developed for determining the existence of domestic subsidies at all given the vast range of tax/subsidy and regulatory policies that affect the competitive position of firms. Finally, as developed at length in Sykes (1989), the unilateral use of countervailing duties by importing nations is exceedingly difficult to defend if the policy objective is to maximize either national or global economic welfare.

Moving from the normative to the positive, WTO rules on subsidies raise a number of puzzles to which satisfactory answers do not yet exist in the literature. Bagwell and Staiger (2002) focus on export subsidies. Following the political economy literature on trade policy, they posit that governments may give considerable weight to the welfare of exporting industries. Export subsidies may then become optimal from a political standpoint. They consider a three-country model with two exporting nations and one importing nation. Absent cooperation, the two exporting nations may find themselves in a “subsidies war”—in Nash equilibrium their export subsidies impose a negative externality on each other (diverting exports from one to the other) and the private benefits of subsidization exceed the joint benefits. If the two exporting nations then cooperate, they will strike a deal that reduces the degree of subsidization and, depending on parameters, it is possible that such a deal would eliminate export subsidization altogether. Now consider adding the importing nation to the bargain. In the Bagwell/Staiger framework, the importing nation will benefit from export subsidies, ceteris paribus, which improve its terms of trade. Thus, unless the importing nation attaches substantial weight to the welfare of the import-competing industry, an agreement involving all three nations might well lead to greater export subsidization than would be observed if only the exporting nations were to cooperate. It is by no means obvious, therefore, why a multilateral agreement should be hostile to export subsidies across the board. More generally, in the absence of international cooperation, the general tendency is for unilateral trade policy to result in too little trade. Because export subsidies increase the volume of trade, it is something of a puzzle as to why multilateral cooperation systematically condemns them.

I conclude with a note about services trade, which has been subject to WTO rules (GATS) only since 1995. Conceptually, the externality in the services area is the same as in the goods area—restrictions on services imports can extract surplus from foreign service providers. The mechanism that gives rise to this externality will often be different in the services area, however, as will be the approach to ameliorating it. The reason lies in the fact that tariffs are generally not possible with services trade because services imports rarely cross the border in a manner that allows them to be taxed. Instead, most of the market access issues in the services area involve efforts by foreign service
providers—banks, insurance companies, law firms—to establish a physical presence in the importing nation. The obstacles to the entry of foreign service providers are then regulatory, in the form of licensing requirements, prudential requirements, residency requirements, immigration restrictions, and the like. Some such regulations are applied in a facially discriminatory fashion, and in other instances regulation simply imposes a disproportionate burden on foreign service suppliers. See Trebilcock and Howse (1995). The liberalization of trade in services thus poses difficult new challenges. The process of “tariffication” that characterizes the approach of GATT to goods markets simply cannot work here, and other strategies must be employed to reduce the number of protectionist instruments. Further, tight nondiscrimination requirements such as those seen in GATT might eliminate the ability of GATS signatories to protect their service industries at all, a result that is unlikely to be politically palatable. Hence, the strategy of GATS is to allow discriminatory policies to be “scheduled.” In each service sector, signatories make a determination whether to afford national treatment or not to foreign service providers. If they do so, they can nevertheless reserve the right to discriminate in particular ways by listing the discriminatory policy in their schedule of service commitments. Although this approach enables nations to retain the protection that they regard as essential and makes their trade barriers more transparent, it does little by itself to channel protection into the least wasteful policy instruments. On that front, GATS has much more to accomplish. See Sykes (2001).

**ii. Multilateralism, Bilateralism, Regionalism and Trade Discrimination**

The world trading system exhibits a multiplicity of trade agreements, some bilateral (such as the U.S.-Israel Free Trade Agreement), some regional (such as NAFTA and the EU) and some global (the WTO). What explains this pattern? The answers to this question are not fully satisfactory even though the economic literature in the area has grown enormously in the last fifteen years. In this section I can only hope to identify some highlights.

Perhaps the first question to ask is why do nations go beyond bilateral agreements if cooperation is generally harder to orchestrate as the size of an agreement increases? The answer, not surprisingly, is that bilateral agreements create important externalities. Imagine three countries, A, B, and C, each of which trades with the others. Imagine further that tariffs on imports into each country are initially nondiscriminatory. Suppose that A and B contemplate a trade agreement in which each lowers its tariffs on imports from the other, and suppose for concreteness that A lowers its tariff on imports of grain
from B. The effect on C will depend on whether it is an importer or exporter of grain. If C initially exports grain to A, then the tariff reduction on A’s imports from B will cause imports into A from B to rise at the expense of imports from C. The weakening in demand for C’s exports causes C’s terms of trade to worsen. And to the degree that the tariff preference for B causes consumers in A to shift purchases from C to B, the phenomenon of *trade diversion* arises. In general, trade diversion occurs whenever discriminatory trade policies cause consumers in an importing nation to make purchases from abroad that would not be made but for the discrimination. It yields a deadweight loss in global welfare, *ceteris paribus*, because imports do not originate from the lowest cost supplier. Viner (1950) is a classic reference. Another type of externality arises if C is instead an importer of grain. The preference for grain imports into A from B will then tend to increase the cost of grain to C. C’s terms of trade worsen because the price of its imports increases.

Of course, an agreement to liberalize trade within a subset of trading nations also has its benefits. Most importantly, if an agreement leads to an expansion of trade between nations that are efficient suppliers of each others’ markets, the phenomenon of *trade creation* arises. It is entirely possible that the benefits of trade creation will exceed the losses from trade diversion, so that a bilateral or regional agreement can on balance enhance global welfare as conventionally defined, and can assuredly enhance the welfare of its members.

Because of the terms of trade externalities for nonmembers, however, any agreement involving only a subset of trading nations has the potential to injure nonparties to the agreement. Formal models of 3 or more country trading networks thus tend to suggest that agreements among any subset of countries will lead to inefficiencies in Nash equilibrium unless some constraint is imposed on the terms of those agreements. In that regard, Kemp and Wan (1976) derive an important result about the formation of agreements known as *customs unions*, in which each every party to the agreement harmonizes its external tariff. They show that such an entity can adjust its common external tariff so that the terms of trade with nonmembers remain the same as before the formation of the union. The welfare of nonmembers is then unaffected, and the gains to the member states are equal to the global gains. Their result is closely related to the concept of reciprocity developed in Bagwell and Staiger (2002).

In light of the mischief that may result from discrimination in international trade, however, one wonders whether a prohibition on discrimination may be valuable. Precisely such a general prohibition is contained in Articles I and XIII of GATT, which prohibit discriminatory tariffs and quota regimes (when the latter are allowable at all).
This obligation is termed the most-favored-nation (MFN) obligation. By joining a multilateral agreement such as GATT and committing themselves to respect the MFN obligation (subject to a very important exception to be discussed in a moment), signatories avoid the loss of joint surplus associated with trade diversion that could arise in a smaller agreement. And in addition to ensuring that goods are imported from the lowest cost foreign supplier, the MFN obligation may facilitate trade negotiations in an important way. Imagine once again a trading system consisting of countries A, B and C, and suppose that A and B contemplate a trade agreement with each other. Negotiators for A and B may worry that after concluding an agreement, their partner may later negotiate an agreement with C, offering C a more attractive arrangement. Whatever benefits they expect to get from the agreement may be undermined, and they may be reluctant to conclude any agreement at all. But if they mutually promise each other that no third nation will be offered better terms later through an MFN commitment, such worries can be put to rest. Schwartz and Sykes (1996).

To be sure, the MFN obligation may also carry a cost. If benefits extended by A to B must be extended automatically to C because C is entitled to MFN treatment, then C may become a free rider on trade negotiations between A and B. To solve the free rider problem, it may be necessary for A and B to draw C into the negotiations. And as the number of countries drawn into negotiations rises, the costs of negotiation rise and the problem of holdouts surfaces.

Thus, the role of the MFN obligation in trade agreements raises cross-cutting and complex issues. A considerable literature now exists on the subject, which confirms that the welfare implications of the MFN obligation are—complicated to say the least. A valuable recent survey is that of Horn and Mavroidis (2001).

The reader may note, however, that even if an MFN obligation is on balance desirable in the trading system, that observation by itself does not necessitate multilateral agreements. One could imagine in the abstract a web of bilateral agreements, each containing an MFN clause. We then return to the question posed at the outset—what is the value of multilateral agreements in this context?

A trivial answer is that economies of scale may exist to the degree that the optimal bilateral agreements would all contain many of the same terms. It may then be easier to write a single multilateral agreement. Further, given the free rider problem that arises if bilateral agreements contain an MFN obligation, it may make sense for all trading nations to meet and negotiate at once, each holding off any final deals until all offers are on the table. Bagwell and Staiger (2002) also show that even when an MFN obligation precludes the possibility of discriminatory tariffs, bilateral negotiations
nevertheless create terms of trade externalities for third countries. For example, if nation A offers a tariff concession on some product to nation B and extends it to all other exporters of the product pursuant to a most favored nation obligation, a worsening of the terms of trade will still occur for any third nation that imports the product. Such externalities provide further justification for multilateral negotiations.

Maggi (1999) identifies another potentially important consideration favoring multilateral cooperation. Imagine three countries once again, A, B and C, and suppose further that A runs a large trade surplus with B, B runs a large surplus with C, and C runs a large surplus with A (although aggregate trade for each country is balanced). In this scenario, an imbalance of power may be said to exist in each bilateral relationship. Because of that imbalance, the amount of self-enforcing cooperation that is sustainable in bilateral tariff agreements is limited—if B defects from an agreement with A, for example, A loses much more than B loses if A defects from the agreement. This scenario suggests a role for an agreement with a multilateral enforcement mechanism, whereby defection by any party can be punished by all of the others. As Maggi acknowledges, however, the extent to which such coordinated enforcement occurs in practice within the WTO is unclear.

If trade discrimination is often problematic, and if important externalities argue for multilateral trade negotiations, an important puzzle remains. GATT Article XXIV permits signatories to form customs unions and free trade areas (the latter differing from a customs union in that the members do not harmonize their external tariffs). Both customs unions and free trade areas permit member nations to deviate from the most favored nation obligation, as long as they eliminate barriers on “substantially all” trade between them. Why should GATT include this gaping exception to the most favored nation obligation, allowing discrimination as long as it occurs on a grand scale (“substantially all trade”)? Related, is the proliferation of preferential trading arrangements such as the EU, NAFTA, and Mercosur a positive or unfortunate development for the international trading system?

Note that nothing in Article XXIV ensures compensation to nations that are injured in their terms of trade by the formation of a preferential trading arrangement. 5 For

---

5If the arrangement is a customs union, Article XXIV does require compensation if the new common external tariff violates a tariff commitment previously made by a member state (e.g., if after Turkey joins the EU, it raises its tariff on widgets above the level it had previously promised to some nonEU trading partner, the EU must compensate that partner). But this compensation obligation is incomplete (it does nothing to compensate for the increase in market power enjoyed by the trading block, for example) and its efficacy in practice has proven questionable. Moreover, when the preferential arrangement is a free trade area, members do not alter their tariffs on imports from nonmembers at all, and no compensation is required.
this reason, an enormous theoretical and empirical literature has developed on preferential trading arrangements, and I will make no attempt at a survey here. In general, theoretical work tends to be skeptical of preferential trade for the reasons suggested by the discussion above. Valuable windows into the literature include Bagwell and Staiger (2002), Bhagwati and Panagariya (1996) and Grossman and Helpman (1995b). Empirical work perhaps tends to be more agnostic, with some support for the idea that the beneficial trade creation under existing agreements may exceed the harmful trade diversion. But on the ultimate question whether preferential arrangements are “building blocks” or “stumbling blocks” to multilateral cooperation in the words of Lawrence (1996), the issue remains unsettled.

iii. Enforcing Trade Agreements

I begin with a review of the WTO/GATT dispute resolution system, a review that will highlight a number of puzzles. Shortly after GATT was formed, the organization gravitated to a “consensus” rule for many decisions. Under the consensus rule, an investigation into an alleged violation of the agreement (undertaken by a dispute “panel”) could only be authorized by consensus (unanimity). Even if the panel was constituted and ruled in favor of the complaining nation, the ruling still had no force of law unless it was “adopted” by the membership, again requiring consensus. Finally, sanctions for violations following the adoption of a ruling could only be authorized by consensus. Thus, the violator nation could always block investigation, adoption, and sanctions. Nations often allowed investigations to go forward nevertheless, and also in many cases permitted adverse rulings to be adopted. But sanctions were only authorized once during the history of GATT (prior to the creation of the WTO).

Under this system, cheating or alleged cheating could only be punished through unilateral “self-help.” Nations such as the United States would bring cases to the GATT, and retaliate unilaterally if the target country refused to comply with an adverse ruling or blocked the process from going forward. Studies of the efficacy of unilateral retaliation during this phase of GATT indicate that its results were mixed, although retaliation often induced nations to alter their policies, and many nations were observed to comply with adverse panel rulings without the need for any retaliation. See Sykes (1992), Bayard and Elliott (1994).

In 1989, GATT signatories agreed to end the ability of accused nations to block an investigation, although they could still block the adoption of panel reports and block sanctions. Finally, with the creation of the WTO in 1995, the system put an end to blocking altogether. Rulings by a dispute panel (or the new Appellate Body) are
automatically adopted unless a consensus exists against adoption, and complaining nations are entitled to use sanctions as a matter of right against a party that has been adjudicated to be in violation of WTO law and that refuses to bring its behavior into compliance. Retaliation is limited to the withdrawal of “substantially equivalent” trade concessions, and the level of retaliation under this criterion is subject to arbitration.

The system has one other unusual feature. If a nation violates WTO law, it is not sanctioned for the violation per se. Rather, a sanction is permissible only after the nation in question has been adjudicated to be in violation, and even then only after it has been given a “reasonable time” to comply with the adverse ruling. A nation that complies within a “reasonable time” thus faces no sanction at all. Further, although the treaty text is not entirely clear on the matter, the general view is that any sanctions should be “substantially equivalent” to the prospective harm from the refusal of the violator nation to comply going forward—there is no sanction for the trade injury caused by the violation until the “reasonable time” for compliance has elapsed.

How well do the available economic models explain the history and practice of WTO/GATT dispute resolution? Standard models of trade agreements suggest that such agreements are self-enforcing, with cooperation supported by mutual threats of defection. Bagwell and Staiger (2002) provide a good illustration. They imagine that two nations agree to cut tariffs in an infinitely repeated game. Each nation is reasonably patient, and plays the “grim” strategy—if either nation cheats in one period, the other reverts to its Nash equilibrium tariff choice in the next and all future periods. Recurring cooperation is then a subgame perfect equilibrium as long as the cooperative tariff does not create too much incentive to cheat in the short term. This latter requirement places a limit on the amount of tariff cooperation that is self-enforcing, and it is possible that the politically first-best tariff level is unsustainable.

In such models, trade retaliation (as distinguished from compensation for an adjustment to the bargain, discussed below) is an out of equilibrium behavior and will not be observed in practice. Subgame perfection likewise rules out strategies such as “tit for tat,” which allow retaliation to end once cooperation has been restored. There is also no role in such models for a formal dispute resolution system. They thus provide quite an incomplete account of WTO/GATT enforcement, where retaliation is in fact observed, is transitory in almost all cases, and is now (generally) seen only after an adverse ruling by a formal dispute settlement proceeding.

Some features of the system may perhaps be explained by the considerations developed above in Section 3. The formal dispute resolution system, for example, serves as an information revelation mechanism. WTO members may be unable to discern on
their own whether another member has cheated, and a formal investigation may be helpful to identify cheaters and trigger reputational sanctions. In addition, the dispute resolution system may serve as a source of default rules for an incomplete bargain, approximating what WTO members would have negotiated had they addressed the issues directly.

Schwartz and Sykes (2002) address the question of why the old GATT system with “blocking” was replaced by the new WTO system with automatic sanctions. Some commentators suggest that the new system was designed to toughen sanctions and discourage cheating. Schwartz and Sykes argue, however, that the new system was designed to reduce rather than increase the penalty for cheating. They observe that the penalty for cheating in both the old and the new systems is essentially the same—aggrieved nations retaliate with trade sanctions. What has changed is that the aggrieved nations can no longer set the level of sanction themselves without central oversight. Instead, all sanctions are subject to arbitration to determine whether they are “substantially equivalent” to the harm caused by the ongoing violation. The new system thus reins in the magnitude of unilateral retaliation. This is valuable if excessive retaliation can destabilize the system by triggering a trade war, or if it is important to calibrate retaliation to facilitate “efficient breach” (discussed further below).

I am unaware of any satisfactory explanation for the fact that sanctions are prospective, and limited to situations in which violator nations refuse to comply with rulings after a “reasonable time.” Perhaps litigation in the WTO has large positive externalities in clarifying the terms of the bargain, and the limited prospect of sanctions encourages disputants to litigate to conclusion. Perhaps reputational concerns alone are enough to discourage most blatant cheating, so that most cases involve good faith disputes over legal obligations. Given the fact that trade sanctions are costly (creating deadweight costs and perhaps political costs as well), the membership may not wish to sanction behavior that arises in good faith. Sanctions may then be used only as a last resort against recalcitrant cheaters. But these suggestions are little more than conjectures, and a fully convincing explanation remains to be developed.

One last feature of the system that has drawn some attention is the problem of power asymmetries. It is often suggested that smaller nations have no effective power to retaliate for trade violations—because they have little market power, trade sanctions simply raise prices to their consumers with little impact on the targets of sanctions. Mavroidis (2000) provides some evidence that smaller nations are indeed less effective in securing redress through the WTO dispute resolution system. In response to this problem, Mexico has proposed within the WTO that retaliation rights be tradable. Bagwell,
Mavroidis and Staiger (2003) formalize this suggestion in an auction model. If the violator nation is allowed to participate in the auction, their model suggests that it will win the auction and retire the retaliation rights. This outcome has the interesting feature that trade retaliation never actually occurs, and is replaced by a monetary payment to the nation that initially has the right to retaliate. Smaller nations thus gain some leverage, and trade retaliation with its deadweight costs is replaced by a monetary sanction that is a pure transfer. It remains to see whether such proposals have any “legs” within the system.

Some related issues about the dispute settlement system remain largely unexplored. For example, if it is indeed the case that trade sanctions are perceived as costly by WTO members (and it is not obvious that they should be in a political sense), why does the WTO not embrace an alternative sanctioning regime, such as one involving money damages that are a pure transfer? Further, why is standing limited to WTO members, even though violations of WTO law plainly impose substantial costs on private interest groups? Many features of the enforcement mechanism clearly warrant further study.

iv. Adjusting the Bargain

Like most long term agreements, WTO commitments are negotiated under conditions of considerable uncertainty about the future. The treaty text thus contains many provisions for adjusting the bargain over time. Here I will emphasize three that have received attention in the literature: provisions for renegotiation; an “escape clause”; and a mechanism for the facilitation of efficient breach.

In the WTO as in many commercial contexts, it is no doubt too costly to write a complete contingent contract. Commitments entered at one point in time may then prove undesirable (at least from the political standpoint of government officials) at some future time. One solution, of course, is simply to renegotiate when circumstances change, recognizing that renegotiation is a potential double-edged sword if sunk costs can be exploited.

Ordinarily, agreements do not require express provisions providing for renegotiation, in as much as renegotiation is always a possibility without them. But GATT included a provision for the renegotiation of tariff commitments, Article XXVIII, in its original 1947 text. Article XXVIII provides that a nation wishing to withdraw a tariff concession should attempt to obtain the permission of affected nations by offering to substitute some other concession. If such negotiations fail to reach agreement, however, a nation can unilaterally withdraw its tariff concession. Affected nations may then retaliate by withdrawing “substantially equivalent” concessions of their own. As
Bagwell and Staiger (2002) argue, this structure can be seen as maintaining reciprocity—roughly speaking, a nation that proposes to withdraw a concession must restore the terms of trade for affected nations, either by substituting an alternative concession or by accepting a withdrawal of concessions on its exports.

But why is it necessary to memorialize this mechanism in Article XXVIII, instead of simply recognizing that GATT members can renegotiate the agreement when they wish? Schwartz and Sykes (2002) offer the suggestion that the members desired to make clear that tariff concessions could be withdrawn if necessary without securing the permission of affected nations beforehand. In a rough sense, they wished to create a “liability rule” rather than a “property rule.” The reason for preferring a liability rule here is the holdout problem. If the permission of all affected GATT members had to be obtained before a politically uncomfortable concession could be modified, the negotiation process could drag on indefinitely and the opportunity to adjust the bargain to changing circumstances would be diminished.

Of course, renegotiation is but one option for avoiding the performance of obligations that become inefficient due to changing circumstances. Drawing on the analogy to commercial contracts, such agreements routinely contain provisions that excuse performance under various contingencies—force majeure and Act of God clauses are illustrative. In the GATT system, a similar state-contingent device for excusing performance is the Article XIX “escape clause.” It provides for the temporary suspension of tariff concessions when, due to unforeseen developments, increased quantities of imports cause or threaten to cause “serious injury” to an import-competing industry. The escape clause thus permits temporary tariff increases, under specified conditions, in response to import surges.

Formal models of the role of the escape clause include Bagwell and Staiger (1990) and Sykes (1991). Bagwell and Staiger suggest that import surges risk destabilizing cooperation because they enhance the incentive to cheat. Following an import surge, the short term incentive to cheat to exploit terms of trade gains increases, as may the political incentive to cheat if the import-competing industry is well organized. They suggest that to avoid unraveling of the agreement that might occur if tariff increases under these circumstances were defined as cheating, the GATT agreement instead allows temporary tariff increases during the import surge. Sykes takes a slightly different approach, modeling the escape clause as a state-contingent rule that permits parties to an agreement to withdraw a concession whenever it is no longer jointly optimal for the concession to be honored—essentially, when the cost of performance to one party exceeds the benefit to the other party (in political terms evaluated at *ex ante* shadow
prices). Such a rule increases the expected utility of trade concessions and thereby allows more concessions to be negotiated in the first instance. The model thus lends support to the suggestion by Dam (1970) that the function of the escape clause is to encourage trade concessions *ex ante*.

Considerable controversy has arisen in recent years, however, as to the circumstances under which WTO members may resort to the escape clause. Much of the problem relates to the confusing treaty text which requires, among other things, that increased quantities of imports must cause or threaten serious injury to an import-competing industry. How is this to be interpreted, given that the quantity of imports is an endogenous rather than a causal variable? For a survey of the current controversies and some suggestions by economically oriented scholars regarding their possible resolution, see Grossman and Mavroidis (2004).

A final mechanism for adjusting the bargain is suggested by scholarly work on contract damages by Shavell and others. It is now well known that a rule of “expectation damages,” which requires a party who breaches a contract to compensate the other for its losses, induces the breaching party to internalize the joint costs of breach and thus to breach only when it is “efficient.” Schwartz and Sykes (2002) suggest that the rules for retaliation under WTO law, which allow aggrieved parties to withdraw “substantially equivalent” concessions and subjects retaliatory withdrawals to arbitration under this criterion, may be roughly equivalent to an expectation damages regime, allowing aggrieved WTO signatories to retaliate to a degree that approximately restores their prebreach political welfare but no more. The opportunity for signatories to “buy out” their obligations under this rule is particularly valuable in circumstances where the other options for adjusting the bargain, such as tariff renegotiations and the escape clause, do not address the underlying political problem. The beef hormone controversy between the United States and Europe is perhaps illustrative—if one assumes that European officials are under intense pressure to prohibit hormone-raised beef from entering the domestic market, they cannot achieve this limited goal by adjusting most-favored-nation tariffs applicable to all beef in an Article XXVIII renegotiation, or by invoking Article XIX on temporary import surges. The efficient breach hypothesis regarding the role of calibrated retaliation in the WTO receives further exploration and critique in Lawrence (2003).

*v. International Investment Law*

Just as national policies toward trade in goods and services cause externalities, so too can national policies toward factor flows. As in the trade area, the policy choices of a
“small’ country will not affect factor suppliers or purchasers abroad, where “smallness” simply means that the nation has no power to affect factor prices or returns on world markets. But in many practical settings, national policies to restrict factor flows can and do affect prices. If a “large” country imposes a tax on foreign direct investment within its territory, for example, it can extract some of the rents on inframarginal investments.

Interestingly, multilateral agreements regarding factor flows are of minimal significance. Immigration policy has long been viewed as the province of national governments, and has been little affected by multilateral agreements (save for limited commitments within GATS to permit the temporary movement of persons to establish service providers). The WTO has its Agreement on Trade-Related Investment Measures (TRIMs), but that agreement is no more than a restatement of GATT principles regarding trade in goods (for example, TRIMs prohibits restrictions on foreign investors that require their operations to buy domestic rather than imported input products). The OECD put forth a proposal in the 1990’s for a multilateral agreement on investment, but it failed in the face of anti-globalization activism.

The subject of investment has received extensive treatment in bilateral agreements, however, dating back to early Friendship, Commerce and Navigation Treaties. More recent years have seen a rapid proliferation of so-called Bilateral Investment Treaties (BITs), and there are now well over a thousand of them. BITs are negotiated between a developed country and a developing country, and their primary purpose is usually said to be the protection of foreign investors against expropriation by the government of the developing country (although their legal commitments formally run in both directions). They provide investors with a right to prompt compensation in the event of expropriation, and also typically provide that disputes may be referred to neutral arbitration. BITs are something of an anomaly in international law in that they give private actors the right to bring an action against a government alleged to have violated the treaty (private “standing”), and further require a government adjudged to have breached its obligations to pay money damages to the aggrieved investor. Investor rights provisions along these lines are also to be found in NAFTA.

Prevailing international law on investment thus gives foreign investors no assurance against discriminatory treatment in general—it is perfectly legal for a host country to condition the right to invest on an obligation to bring in domestic partners, for example, or on other conditions that do not run afoul of GATT or GATS rules. But the law does protect against expropriation, suggesting that investors are most concerned with their vulnerability after costs are sunk. This pattern may suggest that few countries are “large” given the mobility of capital, or that any “large” countries unilaterally refrain
from policies that systematically tax foreign investors. Instead, the concern is for the unpredictable instances in which investment returns are threatened by host-country policies after investment is in place.

Guzman (1998) reviews the history of BITs, and focuses on a historical puzzle regarding their rising popularity. He notes that the protection of investor rights in developing nations has waxed and waned over the past century. After the emergence of customary international law requiring prompt and adequate compensation for any expropriation of the property of an alien, a movement within the United Nations led to the passage of a General Assembly resolution that weakened the right to compensation and restored national “sovereignty” over domestic resources. Developing nations overwhelmingly supported this resolution. Not long afterward, however, BITs became popular and many developing countries flocked to them, agreeing under BITs to waive their “sovereignty.”

Guzman argues that the apparent contradiction in the behavior of developing countries may be explained by the fact that they were acting collectively in the United Nations, but were forced into unilateral action by the push for BITs. He suggests that the opportunity to selectively expropriate investors amounts to a tax on investment that benefits developing countries because of their collective market power over investment opportunities. Accordingly, they sought to strengthen that power when given the opportunity to do so collectively within the United Nations. But when they were approached on a bilateral basis and asked to sign a BIT, they faced a Prisoner’s Dilemma. Each nation could attract more investment by committing not to expropriate investors, but the private gain from such a commitment was largely a transfer from other developing nations. The growth of BITs thus represents a cascade of defection from the collective interest of developing nations. Guzman thus concludes that although BITs may be globally welfare enhancing because they represent a retreat from the exercise of market power by developing nations, they may well have lowered the welfare of the developing world. An alternative hypothesis, of course, is that developing countries came to realize that their strategy in the United nations was a mistake, and that retaining an opportunity to expropriate raised their cost of capital by more than the gains from any expected expropriation (perhaps because of the risk aversion of investors). Although Guzman’s thesis is surely an intriguing one, it likely does not represent the final word on the matter.

Been and Beauvais (2003) address another important issue that has arisen with respect to investor rights treaties—the definition of “expropriation.” Recent arbitral decisions pursuant to the NAFTA investor rights provisions have suggested that
“expropriation” under NAFTA rules may include what U.S. legal scholars sometimes refer to as regulatory takings. In one case, for example, an American waste disposal company won a multimillion dollar award against Mexico after operating permits for a disposal facility that it was constructing in Mexico were denied on environmental grounds. The arbitrators concluded that the Mexican government had misled the investor and that it had reasonably relied on assurances that the facility would be allowed to operate.

Been and Beauvais criticize this and other decisions in so far as they represent a trend toward incorporating regulatory takings into the protection against expropriation. They review the literature on takings generally, which contains two economic arguments for compensating those injured by a regulatory taking. One strand, often associated with Epstein (1985), argues that compensation is necessary so that governments will internalize the costs of their regulatory decisions and make them efficiently. Been and Beauvais echo some familiar criticisms of this argument, making the central point that the argument for cost internalization by government agencies is weakened by the fact that those agencies do not reap the benefits of their regulatory actions. To require them to bear the costs may then distort rather than correct incentives. Another strand of literature, often associated with Blume and Rubinfeld (1984), argues for compensation as an insurance mechanism. Been and Beauvais suggest that public compensation for regulatory takings is difficult to justify on this basis given the moral hazard problem that it creates. Economically justifiable insurance, they suggest, will generally be supplied by the market and should be procured there (they note the existence of market-based political risk insurance as an example). Finally, they question any suggestion that regulatory takings compensation will benefit nations by lowering their cost of capital enough to exceed the costs of compensation—were that so, nations would regularly offer clear assurances of compensation for regulatory takings and we simply do not observe them. Accordingly, they conclude that the broadening of the concept of “expropriation” to encompass regulatory takings is unwise within NAFTA and in other international settings (such as BITs).

vi. International Antitrust

The last decade has witnessed an explosion of commentary on international antitrust issues, spurred by an initiative on the part of some WTO member states to bring competition policy under the WTO umbrella. Much ink has been spilled over the wisdom of a global competition policy agreement, and over the question whether the WTO is an
appropriate place to locate it. To date, however, only limited international cooperation on antitrust issues has emerged, mainly through bilateral agreements on the exchange of information.

Proponents of an international competition policy agreement often make their case by pointing to policy externalities, the existence of which has long been recognized. See Ordover and Sykes (1988). For example, a single price monopolist is a source of deadweight loss in a closed economy. But if the monopolist is an exporter, at least some of its monopoly profit represents a transfer from foreign consumers. From the national welfare perspective of the exporting country, therefore, the existence of the monopolist may be welfare enhancing even though it assuredly reduces global welfare in the aggregate. If the antitrust enforcement authorities in the exporting country give more weight to domestic welfare than to global welfare, a seemingly plausible scenario, they may then decline to act against the monopolist even if they might do so in a closed economy situation. The same may be said about the decision to act against a domestic cartel, or against domestic members of an international cartel. Anecdotal evidence suggests that governments indeed seem to favor national welfare over global welfare at times—the Webb Pomerene Act in the United States, for instance, exempts export cartels from antitrust prosecution. Then, if the nations that are injured by anticompetitive practices are for some reason unable to do anything about it—perhaps because of difficulties in securing jurisdiction or in mounting a credible threat to sanction the firms in question—behavior that reduces global welfare may persist.

The reverse problem can also arise. Imagine a merger between two firms in a concentrated industry, and assume that the merger increases global welfare because it produces efficiencies that exceed any additional deadweight costs due to the higher oligopoly margins that may result from increased industry concentration. From the perspective of a nation that imports the goods produced by that industry, however, the merger may appear undesirable because the importing nation sees only the possibility of higher prices, while the efficiency gains may be captured primarily by the foreign shareholders of the merging firms. If the importing nation has the ability to block the merger, therefore, it may choose to do so even though the merger enhances global welfare. The general lesson is that the effects of imperfect competition on the welfare of individual nations turns critically on the national identity of the consumers and shareholders in an industry and on the division of surplus among them. It follows immediately that antitrust policy, which can alter conditions of competition in imperfectly competitive industries, has potentially important external effects in industries with significant international trade.
Commentators such as Fox (1999) and Guzman (1998) proceed from such observations to argue that some degree of international cooperation on antitrust is desirable. Even accepting this premise, however, a possible difficulty emphasized by Guzman is the problem of an empty core. He doubts that all nations would gain from a competition policy agreement that imposed sound principles of antitrust policy. He thus argues for issue linkage as a possible solution, which would facilitate sidepayments, and thus for embedding competition policy within the broader WTO umbrella. Opponents of extensive WTO involvement, including Fox, express concern about the ability of its dispute resolution mechanism to address antitrust issues adequately, and fear that the WTO tradition of accommodating protectionist pressures may undermine the pursuit of efficient antitrust policy. Still other commentators, such as Wood (1999), oppose any form of multilateral antitrust agreement. They reason that no consensus exists on the proper principles of antitrust law. If Chicago school thinking has come to pervade American policy, it has by no means swept the debate in venues such as Europe and Canada. The compromises necessary to reach an international agreement on key principles, therefore, might undermine the core values of national antitrust enforcers.

The evident lack of consensus on many antitrust principles has led some observers to propose more modest agreements, perhaps imposing a basic national treatment obligation (an obligation to treat foreign firms no less favorably than domestic firms), and an agreement to prohibit certain hard core cartel practices. But even these seemingly narrow commitments are not without their problems. Antitrust enforcement often involves difficult and rather subjective judgments—will a merger lead to substantial efficiencies, and will the increase in concentration cause prices to rise importantly? If national authorities were accused of shading these judgments to favor their domestic firms over foreign interests, it is not at all clear how an arbitrator could confidently determine the truth of the allegation. Likewise, although a consensus may exist that a hard core cartel practice such as price fixing is undesirable, disagreement may well arise over the question of when price fixing is present. The long-standing controversy in the United States over the line between price fixing and mere “conscious parallelism,” and the “plus factors” that are required to prove the former, illustrates the problem. At least in part because of such difficulties, the prospects for substantially greater international cooperation on antitrust issues appear rather bleak at this writing.
Despite its importance within the legal academy, virtually nothing has been written from a theoretical perspective by economically oriented scholars on international human rights law. Only Goldsmith and Posner (2004) address it briefly.

The subject matter of human rights law is primarily the treatment of domestic nationals by their own governments. Both treaties and customary rules create a number of “rights” in this regard, ranging from prohibitions on genocide and torture to rules requiring nondiscrimination among on the basis of race or gender. It is not immediately obvious why such matters, which are seemingly internal to each state, should become a subject of international law at all. The best theory of an “externality” is perhaps altruism—citizens of foreign states care about the welfare of oppressed people abroad to a degree, and are willing to expend resources to help them. Certainly the recent history of humanitarian interventions in various settings offers some support for the existence of this externality. Yet, unlike the externalities discussed in earlier sections of this chapter, there is little reason to think of this externality as reciprocal: citizens of the United States may care about the fate of repressed minorities in Serbia, Iraq or Afghanistan, but there is little reason to think that citizens of those countries simultaneously worry about human rights violations in the United States. Thus, human rights agreements do not fit well within the standard model of international agreements under which each signatory state gains from the elimination of a reciprocal externality, and it is something of a puzzle as to why nations with weak human rights regimes would sign them (absent sidepayments, which we do not seem to observe). Likewise, international human rights treaties generally lack a formal enforcement mechanism, and it is difficult to see how such agreements can be self-enforcing—threats of mutual defection are useless if a violator state does not care about defection by another.

Hence, a positive theory of human rights agreements must look to other considerations. Some of the explanations that Goldsmith and Posner offer for customary international law may have purchase. Human rights agreements may simply memorialize a coincidence of interest among most signatories, stating principles to which their domestic legal systems already conform. Repressive states may sign such treaties as well thinking that violation is essentially costless, and that a public commitment to human rights may have some modest political benefit. Coercion is another possible explanation. Liberal states may create human rights treaties for the purpose of altering the global political dynamic when they contemplate the use of force or economic sanctions—support for coercive actions against repressive states may be enhanced if the repressive
states can be said to be in violation of international law. Related, human rights violations may become the predicate for war crimes trials after a successful military intervention, and may then have some value for the deterrence of rogue leaders. Finally, political scientists have noted that where human rights treaties are binding under domestic law, such as certain intra-European agreements, they may be used by leaders of newly democratic states to tie the hands of future leaders and discourage a return to authoritarianism.

In light of these considerations, it is perhaps unsurprising that violations of human rights agreements are rampant, as Goldsmith and Posner document. The finding in the empirical political science literature that ratification of human rights agreements seems to have little effect on the behavior of ratifying states is also unsurprising. See Hathaway (2002). Human rights treaties in the main commit liberal states to behave as they would anyway. Whenever a treaty would require a liberal state to change its policies, they typically sign only with a reservation (such as certain reservations taken by the United States to preserve its right to use capital punishment). Rogue states and other repressive regimes may or may not sign human agreements, but when they do so they generally have no intention of altering their behavior to comply.

viii. Conflicts of Law

Conflicts of law arise when aspects of a legal dispute may be governed by two or more different legal rules. The issue usually arises when a dispute involves parties who are citizens of different states. Although the subject is usually studied with reference to domestic conflicts of law, it has an obvious international counterpart. Consider, for example, the events some years ago in Bhopal, India involving an explosion at a chemical plant owned by an American company (Dupont). Suppose that injured parties wish to file a lawsuit seeking compensation—should the suit be governed by Indian law (the law of the jurisdiction where the accident occurred), U.S. law (the law of the jurisdiction in which the defendant is incorporated), or the law of some other jurisdiction? The question is an important one because the choice of governing law may significantly affect the plaintiff’s chances for victory, or the damages that the plaintiff can collect. It may also have other allocative consequences as discussed below.

To the extent that local laws are tailored to reflect local conditions (such as the residents’ willingness to pay for safety), a simple argument may be made for applying the law of the jurisdiction where the harm arose. But such a rule may lead to strategic incentives that create important inter-jurisdictional externalities. The work that has been done on the subject in the domestic context focuses on the question whether states will
manipulate their choice of law rules to extract surplus from other states, and whether that in turn will adversely affect other rules of substantive law. The two leading papers explore these questions with particular reference to state products liability law. McConnell (1988) considers the consequences of the traditional U.S. rule for choosing the law to govern an accident—the rule that the court should apply the law of the jurisdiction in which the accident occurred (the “territorial rule”). He suggests that under the territorial rule, states will have an incentive to adopt substantive rules of accident law that inefficiently extract surplus from other states: Injured parties will tend to sue in their home state, McConnell suggests, so that most suits will involve in-state plaintiffs, yet many defendants will be out of state companies. On average, therefore, an inefficiently pro-plaintiff law can transfer wealth into the state as long as sellers cannot adjust their prices to reflect liability costs in each state. Out of state product manufacturers cannot do so argues McConnell because arbitrage across jurisdictions prevents price discrimination. Thus, the equilibrium is one in which all states tend to employ suboptimally pro-plaintiff laws. They would benefit from cooperation to avoid the problem, but cooperation is difficult when the number of states is large.

Hay (1992) challenges this analysis, noting that it rests crucially on the assumption that states must employ the law of the jurisdiction where the accident occurs. An alternative choice of law rule has emerged in the United States under the rubric of “interest analysis,” which Hay suggests is sufficiently malleable to allow states to choose between the law of the state in which the accident happened and the law of the state of the defendant. He argues that states will jettison the traditional rule, and use interest analysis to choose whichever law is more favorable to their in-state plaintiff. Their incentive to adopt pro-plaintiff rules of substantive law then diminishes because such rules become a double-edged sword—they benefit in-state plaintiffs in suits against out-of-state firms as before, but they can also be invoked by out-of-state plaintiffs in their home states to disadvantage in-state firms who do business there and become defendants. He further notes that once all states switch to interest analysis, states with relatively pro-plaintiff laws will cause their firms to relocate to jurisdictions with more pro-defendant rules. Hay does not fully model the resulting equilibrium, but suggests that this process will exert powerful discipline on the tendency of states to adopt unduly pro-plaintiff laws.

This work on domestic issues suggests some principles that bear on the international context as well. States can exploit out-of-state sellers of goods and services (or investors) through pro-plaintiff rules of substantive law, or pro-plaintiff conflicts of law rules, only if the out-of-state entities are unable to adjust their prices to recover differences in costs across jurisdictions or if they have made sunk investments in the state
and the legal rule comes as a surprise. As to the first problem, arbitrage may be less of a constraint on international price discrimination given transportation costs and trade barriers, but the problem may still arise. As to the second, unanticipated changes in liability rules are certainly possible, although prices may adjust going forward to restore a competitive return. Investors fearful of unanticipated liability may also charge a risk premium in capital markets, creating an incentive for host states to signal their intention to maintain existing rules if they can do credibly. It is questionable whether BITs have any force in this regard, as it is doubtful that a change in a liability rule would be deemed “expropriation.”

Even if investors anticipate all liability rules, however, inappropriate choice of law rules can distort investment (as well as trade patterns). To return to the Bhopal example, suppose that U.S. law is more generous toward plaintiffs than Indian law. Suppose further that Indian law allows Indian plaintiffs to choose between Indian law and the law of the jurisdiction in which the defendant is incorporated in all suits against foreign companies doing business in India. U.S. firms will then be at a competitive disadvantage relative to Indian firms and perhaps third country firms as well. The result will be to encourage investment, production and/or imports from firms that may be less efficient than U.S. firms. Put differently, if the choice of law rules result in nonuniformity in the rules applied to companies from different countries, the problem of trade diversion arises along with the distortion of investment patterns. The same problem arises if Indian plaintiffs are allowed to come to U.S. courts and invoke U.S. law against U.S. companies when they cannot invoke the same rules against competitors of U.S. companies. The open question is whether the sorts of rules that can produce these distortions will emerge and persist in equilibrium, thus justifying some sort of international cooperation to address the problem.

In short, choice of law rules are potential trade and investment issues. Much more work on the nature of the problem and its empirical importance remains to be done.

ix. The International Commons: Fisheries

I conclude with a short note on an important source of international externalities not yet noted in this chapter: the problem of incomplete property rights. Many valuable resources are unowned, leaving no actor with an incentive to protect them from uneconomic deterioration or more generally to maximize their value. The problem arises with the atmosphere, the oceans, space, Antarctica, the common pasture, oil pools, and
innumerable other resources. I will use fisheries to illustrate the problem because of the rather interesting and discouraging body of international law on the matter.

To crystallize the nature of the externality problem, consider a two-period model of a fishery adapted from classic analyses by Gordon (1954) and Cooper (1975). For a single species fishery, let $H_t$ denote the total number of hours of fishing labor required to catch $Y_t$ fish in fishing season $t$, and let $s_t$ denote the stock of fish at the beginning of season $t$. The relation between $H$, $Y$, and $s$ is given by $H_t = f(s_t)(Y_t)^2$, where $f' < 0$ and $f'' > 0$. Thus, the amount of time per fish invested in fishing increases as the catch increases (reflecting the increasing scarcity of fish), and decreases as the initial stock of fish increases. The beginning of season stock is determined by the size of the stock at the end of the last season plus new “births” given by a transition function that relates the number of new fish added to the fishery between seasons to the end of period stock: $s_t = s_{t-1} - Y_{t-1} + B(s_{t-1} - Y_{t-1})$. We assume that $B' > 0$ at least until some ecological limit on the fishery, and $B'' < 0$. The price of fishing labor is unity, and the price of a fish, assumed constant over time for simplicity, is $p$. The social rate of discount is $r$, and let $\delta = 1/(1 + r)$.

Efficient use of this price-taking fishery requires that the discounted value of profits from the fishery be at a maximum. Using the relation between fishing hours and the catch, the profit function may be written as:

$$\pi = pY_1 - f(s_1)(Y_1)^2 + \delta[pY_2 - f(s_2)(Y_2)^2]$$

The optimization problem is then to select the size of the catch in each fishing season to maximize this function, subject to an initial condition on the stock of fish and to the transition function. The solution to this dynamic programming problem is found by backwards induction. The maximization of season two profits given the initial stock in season two simply requires that price be set equal to marginal cost in season two: $p = 2f(s_2)Y_2$, which allows season two profit ($\pi_2$) to be expressed as a function of $s_2$. Substituting into the profit function above and using the transition function and the initial condition on the stock, the first order condition for the optimal catch in season one (assuming a positive amount of fishing is optimal) may be written as:

$$p = 2f(s_1)Y_1 - \delta(\partial\pi_2/\partial Y_1)$$

This condition also states that price must equal marginal cost, but has two components—the season one marginal cost given the initial stock, plus the effect of marginal fishing in season one on discounted profit in season two (via its effect on the season two stock). Assuming that the fishery is below its ecological limit, the second term on the right hand side is positive. Denoting the first term on the right hand side as “short term” marginal
cost, it is clear that properly computed marginal cost exceeds short term marginal cost, and that the optimal amount of fishing is lower than would be implied by an equality between price and short term marginal cost.

The relationship between this condition for efficient use of the fishery, and the market equilibrium that actually emerges, will depend on the market structure of the industry that uses the fishery. Imagine, for example, that the rights to use the fishery are owned by a single profit-maximizing company. Such a company would confront the profit maximization problem stated above, and would exploit the fishery efficiently (the company’s “monopoly” over the fishery causes no inefficiency because it is a price taker in the market for fish).

As the number of companies using the fishery increases, however, externalities emerge within each season and over time. To isolate the first, assume that each company in each season chooses its fishing effort in Cournot-Nash fashion, taking the fishing effort of other companies as given. Each company will equate price to its private marginal cost, which will incorporate the marginal effect of its own fishing on the hours required to catch fish. But it will neglect the fact that its own fishing also increases the costs to other companies of catching fish. The size of this external effect will tend to be greater as the number of companies using the fishery increases. In the limit, very small companies may behave as if the effect of their own fishing on the hours required to catch fish is zero. The equilibrium for this limiting case may be derived readily (I suppress time subscripts for simplicity). Let $y$ and $h$ denote the catch and hours investment of a small company. Each small company confronts the profit function per season of $\pi = py-h$, where $y$ is given by the average productivity of fishing ($Y/H$) times its hours of fishing, and average productivity is taken to be fixed by each company. Equilibrium requires zero profits, which in turn implies that $p=Y/H$. The last equation states that in equilibrium, price will equal the average cost of fish, in contrast to the condition for efficiency which requires price equal to marginal cost. Because average cost lies below marginal cost, the equilibrium involves overfishing. More generally, it is not difficult to show that with symmetric Cournot-Nash firms, the efficiency of the market equilibrium in each season (taking the initial stock as given) declines steadily as the number of firms increases—the greater the number of firms, the larger the wedge between private short term marginal cost and social short term marginal cost.

The intertemporal externality arises for essentially the same reason. With more than one firm in the fishery, each firm will take account of its own effects on the future stock of fish and its future costs of fishing, but will neglect the fact that greater fishing effort in the current season also raises costs for other firms in future seasons. Thus,
focusing on the second term on the right hand side of the efficiency condition for season one, the private cost of fishing in period one via its effect on future profit is again less than the social cost. This further exacerbates the overfishing problem, and will tend to become more acute as the number of firms rises. In the limiting case, small firms may completely ignore the effect of their fishing effort on the future stock, and the market equilibrium in each season will involve price equal to short term average cost. Any number of trajectories are possible depending on the biology of the fishery, of course, but in the worst case scenario, the stock of fish may shrink steadily until the use of the fishery is no longer economical at all.

Possible corrective measures include all of the usual suspects. Restricting access to the fishery can improve the efficiency with which it is managed. Command and control regulation over the size of the catch can also help, as can taxes on fishing. The choice among these mechanisms turns on familiar considerations that I will not detail here.

To return to the subject of international law, however, the challenges of fisheries management can become all the more acute when no one regulatory authority has jurisdiction, so that a need arises to coordinate across jurisdictions. The history of international cooperation on fisheries management, however, is by and large unimpressive. The Law of the Sea grants each state dominion over a 12-mile area of territorial waters along its coastline, from which it has the right to exclude others. But very little has been done to coordinate fishing in the open oceans. The only substantial cooperation seems to arise when fishing directly or indirectly affects endangered species such as sea turtles or whales.

A number of factors may explain the lack of progress in this area. First, considerations of political economy suggest that support for fisheries management policies may be limited. Even if fishery restrictions would raise the discounted value of profits, existing fishing companies may not favor them. Their time horizons may be limited by the life of their sunk capital, and they may see the benefits of fisheries management inuring to their successors. Other beneficiaries include the future consumers of fish, who seem unlikely to exert much political pull toward cooperation. Companies such as canneries and other fish packing operations may weigh in favor of fisheries management, but the balance of political forces is unclear both within and across jurisdictions.

Second, the fisheries management problem can be highly complex. Interested parties may disagree on the optimal policy. Further, the policy in place in one fishery may have spillover effects on another—a reduction in overfishing in one area may raise prices or divert sunk capital resources to cause greater overfishing in another.
Finally, the enforcement of cooperation is likely difficult, particularly on the open oceans. Defections may be difficult to detect, and enforcement may be plagued by a free rider problem. The suspicion that another party is cheating may lead any agreement to unravel.

In light of these considerations, it is perhaps no surprise that the limited successes in the area have involved measures to protect endangered species. Such measures attract the support of the environmental lobby, which is the most striking example of diffuse interests coming together to organize politically. They may also attract private third-party enforcers such as Greenpeace, which has played a prominent role in publicizing violations of the Whaling Convention.
References


Ethier, Wilfred (2000), Reciprocity, Nondiscrimination and a Multilateral World (mimeo).


Fon, Vincy and Francesco Parisi (2004), Stability and Change in International Customary Law (mimeo).


Readers with comments should address them to:

Professor Alan O. Sykes
University of Chicago Law School
1111 East 60th Street
Chicago, IL 60637
Alan_Sykes@law.uchicago.edu
13. J. Mark Ramseyer, Credibly Committing to Efficiency Wages: Cotton Spinning Cartels in Imperial Japan (March 1993)
16. Lucian Arye Bebchuk and Randal C. Picker, Bankruptcy Rules, Managerial Entrenchment, and Firm-Specific Human Capital (August 1993)
17. J. Mark Ramseyer, Explicit Reasons for Implicit Contracts: The Legal Logic to the Japanese Main Bank System (August 1993)
20. Alan O. Sykes, An Introduction to Regression Analysis (October 1993)
22. Randal C. Picker, An Introduction to Game Theory and the Law (June 1994)
29. Daniel Shaviro, Budget Deficits and the Intergenerational Distribution of Lifetime Consumption (January 1995)
34. J. Mark Ramseyer, Public Choice (November 1995)
41. John R. Lott, Jr. and David B. Mustard, Crime, Deterrence, and Right-to-Carry Concealed Handguns (August 1996)
42. Cass R. Sunstein, Health-Health Tradeoffs (September 1996)
47. John R. Lott, Jr. and Kermit Daniel, Term Limits and Electoral Competitiveness: Evidence from California’s State Legislative Races (May 1997)
48. Randal C. Picker, Simple Games in a Complex World: A Generative Approach to the Adoption of Norms (June 1997)
50. Cass R. Sunstein, Daniel Kahneman, and David Schkade, Assessing Punitive Damages (with Notes on Cognition and Valuation in Law) (December 1997)
52. John R. Lott, Jr., A Simple Explanation for Why Campaign Expenditures are Increasing: The Government is Getting Bigger (February 1998)
60. John R. Lott, Jr., *How Dramatically Did Women's Suffrage Change the Size and Scope of Government?* (September 1998)
64. John R. Lott, Jr., *Public Schooling, Indoctrination, and Totalitarianism* (December 1998)
<table>
<thead>
<tr>
<th>Number</th>
<th>Title and Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>75.</td>
<td>Richard A. Epstein, Deconstructing Privacy: and Putting It Back Together Again (May 1999)</td>
</tr>
<tr>
<td>76.</td>
<td>William M. Landes, Winning the Art Lottery: The Economic Returns to the Ganz Collection (May 1999)</td>
</tr>
<tr>
<td>77.</td>
<td>Cass R. Sunstein, David Schkade, and Daniel Kahneman, Do People Want Optimal Deterrence? (June 1999)</td>
</tr>
<tr>
<td>78.</td>
<td>Tomas J. Philipson and Richard A. Posner, The Long-Run Growth in Obesity as a Function of Technological Change (June 1999)</td>
</tr>
<tr>
<td>79.</td>
<td>David A. Weisbach, Ironing Out the Flat Tax (August 1999)</td>
</tr>
<tr>
<td>81.</td>
<td>David Schkade, Cass R. Sunstein, and Daniel Kahneman, Are Juries Less Erratic than Individuals? Deliberation, Polarization, and Punitive Damages (September 1999)</td>
</tr>
<tr>
<td>82.</td>
<td>Cass R. Sunstein, Nondelegation Canons (September 1999)</td>
</tr>
<tr>
<td>83.</td>
<td>Richard A. Posner, The Theory and Practice of Citations Analysis, with Special Reference to Law and Economics (September 1999)</td>
</tr>
<tr>
<td>84.</td>
<td>Randal C. Picker, Regulating Network Industries: A Look at Intel (October 1999)</td>
</tr>
<tr>
<td>90.</td>
<td>David A. Weisbach, Should the Tax Law Require Current Accrual of Interest on Derivative Financial Instruments? (December 1999)</td>
</tr>
<tr>
<td>95.</td>
<td>David Schkade, Cass R. Sunstein, Daniel Kahneman, Deliberating about Dollars: The Severity Shift (February 2000)</td>
</tr>
</tbody>
</table>
105. Jack Goldsmith and Alan Sykes, The Dormant Commerce Clause and the Internet (November 2000)
110. Saul Levmore, Conjunction and Aggregation (December 2000)
111. Saul Levmore, Puzzling Stock Options and Compensation Norms (December 2000)
112. Richard A. Epstein and Alan O. Sykes, The Assault on Managed Care: Vicarious Liability, Class Actions and the Patient's Bill of Rights (December 2000)
114. Cass R. Sunstein, Switching the Default Rule (January 2001)
116. Jack Goldsmith, Statutory Foreign Affairs Preemption (February 2001)
118. Cass R. Sunstein, Academic Fads and Fashions (with Special Reference to Law) (March 2001)
122. David A. Weisbach, Ten Truths about Tax Shelters (May 2001)
126. Douglas G. Baird and Edward R. Morrison, Bankruptcy Decision Making (June 2001)
127. Cass R. Sunstein, Regulating Risks after ATA (June 2001)
129. Richard A. Epstein, In and Out of Public Solution: The Hidden Perils of Property Transfer (July 2001)
130. Randal C. Picker, Pursuing a Remedy in Microsoft: The Declining Need for Centralized Coordination in a Networked World (July 2001)
131. Cass R. Sunstein, Daniel Kahneman, David Schkade, and Ilana Ritov, Predictably Incoherent Judgments (July 2001)
133. Lisa Bernstein, Private Commercial Law in the Cotton Industry: Creating Cooperation through Rules, Norms, and Institutions (August 2001)
137. Eric A. Posner and George G. Triantis, Covenants Not to Compete from an Incomplete Contracts Perspective (September 2001)
139. Randall S. Kroszner and Philip E. Strahan, Throwing Good Money after Bad? Board Connections and Conflicts in Bank Lending (December 2001)
140. Alan O. Sykes, TRIPs, Pharmaceuticals, Developing Countries, and the Doha “Solution” (February 2002)
141. Edna Ullmann-Margalit and Cass R. Sunstein, Inequality and Indignation (February 2002)
145. David A. Weisbach, Thinking Outside the Little Boxes (March 2002, Texas Law Review)
149. Cass R. Sunstein, Beyond the Precautionary Principle (April 2002)
152. Richard A. Epstein, Steady the Course: Property Rights in Genetic Material (May 2002; revised March 2003)
156. Cass R. Sunstein and Adrian Vermeule, Interpretation and Institutions (July 2002)
159. Randal C. Picker, From Edison to the Broadcast Flag: Mechanisms of Consent and Refusal and the Propertization of Copyright (September 2002)
162. Lior Jacob Strahilevitz, Charismatic Code, Social Norms, and the Emergence of Cooperation on the File-Swapping Networks (September 2002)
163. David A. Weisbach, Does the X-Tax Mark the Spot? (September 2002)
164. Cass R. Sunstein, Conformity and Dissent (September 2002)
166. Douglas Lichtman, Uncertainty and the Standard for Preliminary Relief (October 2002)
171. Richard A. Epstein, Animals as Objects, or Subjects, of Rights (December 2002)
172. David A. Weisbach, Taxation and Risk-Taking with Multiple Tax Rates (December 2002)
181. Amitai Aviram, Regulation by Networks (March 2003)
194. David A. Weisbach and Jacob Nussim, The Integration of Tax and Spending Programs (September 2003)
<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Author(s)</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>199.</td>
<td>The Effects of Taxation on Income Producing Crimes with Variable Leisure Time</td>
<td>Avraham D. Tabbach</td>
<td>October 2003</td>
</tr>
<tr>
<td>200.</td>
<td>Rethinking Prosecution History Estoppel</td>
<td>Douglas Lichtman</td>
<td>October 2003</td>
</tr>
<tr>
<td>203.</td>
<td>The (Non)Taxation of Risk</td>
<td>David A. Weisbach</td>
<td>January 2004</td>
</tr>
<tr>
<td>205.</td>
<td>The Right to Destroy</td>
<td>Lior Jacob Strahilevitz</td>
<td>January 2004</td>
</tr>
<tr>
<td>208.</td>
<td>Disparities and Discrimination in Health Care Coverage; A Critique of the Institute of Medicine Study</td>
<td>Richard A. Epstein</td>
<td>March 2004</td>
</tr>
<tr>
<td>209.</td>
<td>Navigating the Anticommons for Pharmaceutical Patents: Steady the Course on Hatch-Waxman</td>
<td>Richard A. Epstein and Bruce N. Kuhlik</td>
<td>March 2004</td>
</tr>
<tr>
<td>211.</td>
<td>Optimal War and Jus Ad Bellum</td>
<td>Eric A. Posner and Alan O. Sykes</td>
<td>April 2004</td>
</tr>
<tr>
<td>212.</td>
<td>The Persistent Puzzles of Safeguards: Lessons from the Steel Dispute</td>
<td>Alan O’Neil Sykes</td>
<td>May 2004</td>
</tr>
<tr>
<td>213.</td>
<td>Specialization, Firms, and Markets: The Division of Labor within and between Law Firms</td>
<td>Luis Garicano and Thomas N. Hubbard</td>
<td>April 2004</td>
</tr>
<tr>
<td>214.</td>
<td>Hierarchies, Specialization, and the Utilization of Knowledge: Theory and Evidence from the Legal Services Industry</td>
<td>Luis Garicano and Thomas N. Hubbard</td>
<td>April 2004</td>
</tr>
<tr>
<td>216.</td>
<td>The Economics of Public International Law</td>
<td>Alan O. Sykes</td>
<td>July 2004</td>
</tr>
</tbody>
</table>