Public vs. Private Enforcement of International Economic Law: Of Standing and Remedy

Alan O. Sykes
Public vs. Private Enforcement of International Economic Law: Of Standing and Remedy

*Alan O. Sykes*

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

February 2005

This paper can be downloaded without charge at:
and at the Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract_id=671801
Public vs. Private Enforcement of International Economic Law:
Of Standing and Remedy

Alan O. Sykes*

Abstract

This paper develops a positive theory of the rules regarding standing and remedy in international trade and investment agreements. In the investment setting, the paper argues that a central objective of investment treaties is to reduce the risks confronting private investors and thereby to lower the cost of capital for capital importing nations. This objective requires a credible government-to-firm commitment (or signal) that the capital importer will not engage in expropriation or related practices. A private right of action for money damages is the best way to make such a commitment. In the trade setting, by contrast, importing nations have no direct interest in reducing the risks confronting exporters of goods and services, and will desire to make market access promises more secure only if such behavior secures reciprocal benefits for their own exporters. Consequently, commitments in trade agreements are best viewed as government-to-government rather than government-to-firm. The parties to trade agreements can enhance their mutual political welfare by declining to enforce commitments that benefit politically inefficacious exporters, and can most cheaply do so by reserving to themselves the standing to initiate dispute proceedings—a right to act as a “political filter.” The paper also suggests why governments may prefer to utilize trade sanctions rather than money damages as the penalty for breach of a trade agreement.

Like all bodies of law, the public international law of trade and investment requires an enforcement mechanism. The choices to be made in designing such a mechanism are many. Parties to trade and investment agreements must decide whether to create an adjudicative body to hear complaints about alleged breach of obligations, or to rely on informal diplomacy. They must decide whether to create formal sanctions for breach of obligations, or to rely on each party’s concern for its reputation, and perhaps unilateral retaliatory actions, to discourage breach. If they choose to create an adjudicative body, they must decide who has standing to bring complaints before that body. And if they choose to create a formal sanction for breach of obligations, they must select the type of penalty that they will use as well as some mechanism for calibrating its magnitude.

*I am grateful to Roger Saad for able research assistance, and to workshop and conference participants at the University of Chicago and the University of California, Berkeley for many useful suggestions.
A fair amount has been written about various aspects of these issues with specific reference to international economic law, but little has been written from an analytical perspective about what the law and economics literature terms the choice between public and private enforcement of law.\(^1\) This choice becomes relevant once parties to an international agreement elect to allow an adjudicative body to hear complaints. They may then reserve to themselves the exclusive right to petition that body (public enforcement), or allow private actors with a stake in the dispute to petition it (private enforcement). Intertwined with that choice is the parties’ choice of a sanction for breach of obligations—regardless of who has standing before the adjudicative body, the parties to the pertinent agreement may provide no formal sanctions for a finding of breach, they may provide for sanctions that only a governmental party to the agreement has the capacity to administer (such as trade sanctions), or they may provide for money damages that can be paid to the parties injured by violations.

A quick survey of international economic law reveals quite a mixed picture along these dimensions. Private rights of action for money damages have become routine in international investment agreements. In the trade area, by contrast, money damages are much more circumscribed, and provisions for the award of such damages are absent from important multilateral arrangements (although they are always an option for the settlement of disputes as a practical matter). Likewise, some trade agreements afford private actors standing to enforce the rules in a court, while others do not. The goal of this paper is to explain these features of current law from a political economy perspective.

In brief, I argue that in the *investment* arena, the function of international agreements is to reduce the perceived risk of expropriation and related events for private firms operating in global capital markets, and thereby to reduce the cost of capital for

\(^1\) Exceptions include Levy and Srinivasan (1996), Trachtman and Moremen (2003) and Nzelibe (2005), all discussed herein.
capital-importing nations. The most effective mechanism for such risk reduction is a private right of action for compensatory damages should an importing nation engage in proscribed behavior. In the trade arena, by contrast, the function of international agreements is to make credible government to government commitments regarding trade policy, and thereby to raise mutual political welfare relative to an environment without bilateral or multilateral cooperation. For the enforcement of such agreements, it can suffice to provide standing and remedy only to governments, and indeed a private right of action for damages may prove politically counterproductive. By reserving standing to themselves, governments can interpose themselves as “political filters” to exclude certain enforcement actions that private actors might otherwise wish to bring. In some systems, the political filtering process may also occur ex post, as through legislative reversal of politically unpalatable judicial decisions. Regarding the remedy available for breach of trade agreements, I offer several reasons why governments would prefer to employ trade sanctions rather than money damages, and in the process rebut the suggestion by other commentators that trade sanctions are preferred because they are better at coercing compliance.

Section I sets out relevant characteristics of international trade and investment law, and reviews existing commentary on private standing and remedies. Section II then offers a political economy explanation for why investment agreements have private rights of action, emphasizing the importance to investment agreements of inducing private actors to incur new sunk costs. Section III considers the heterogeneity among trade agreements regarding the remedies available to private parties, and explains why governmental parties to trade agreements may prefer to act as political filters in the dispute process by reserving standing to themselves, particularly in the absence of a political body with the power to reverse problematic judicial decisions.
I. Legal and Economic Background

A. Current Law

1. Investment

The public international law of investment is largely a creation of bilateral agreements, and to a lesser extent customary international law. The United States, for example, has relied on a network of bilateral Friendship, Commerce and Navigation (FCN) Treaties to secure limited rights for investors (as well as certain trade and shipping rights) in foreign countries throughout much of its history.\textsuperscript{2} These treaties generally did not create any private rights of action. It was also long thought that customary international law provided foreign investors with protection against expropriation, requiring “prompt, adequate and effective compensation” in the event of any expropriation. Customary law would afford a private right of action to an investor if the host country would allow customary law to be enforced against it in its domestic courts, or would comply with an award by a foreign court.

During the middle of the 20th century, various developing countries began to question whether customary law obliged them to provide “prompt, adequate and effective compensation” for expropriation. This movement culminated with the 1974 U.N. Charter of Economic Rights and Duties of States, adopted by the general Assembly, which provided that compensation for expropriation was to be measured by the law of the expropriating state. These developments created considerable unease among investors in developing countries, and spawned an initiative that began in Europe to negotiate new “Bilateral Investment Treaties” (BITs). The United States began its own program to negotiate BITs in 1977.\textsuperscript{3} BITs typically provide various nondiscrimination commitments, and expressly embrace the old customary law standard of “prompt, adequate and

\textsuperscript{2}For illustrative examples of an FCN treaty, one between the United States and Argentina (1853) and one between the United States and Liberia (1939), see http://www.yale.edu/lawweb/avalon/diplomacy/argentina/argen02.htm and http://170.110.214.18/tcc/data/commerce_html/TCC_2/LiberaFriendship.html.

\textsuperscript{3}For a thorough history, see Vandevelde (1993).
effective compensation” for expropriation. They provide investors with the right to take investment disputes to neutral international arbitration, and commit each party to enforce arbitral awards (including an award of damages).4

Many of the principles found in BITs were incorporated into the investor rights provisions of NAFTA. Chapter Eleven of NAFTA contains non-discrimination obligations respecting investment and an obligation to provide prompt compensation for any “expropriation.”5 It also requires parties to accord investors of another party “treatment in accordance with international law, including fair and equitable treatment and full protection and security.”6 Any dispute under these provisions may be submitted by an investor to arbitration, and the arbitrators have the power to award money damages and restitution.7

NAFTA Chapter Eleven has sparked a number of interesting cases in recent years, which have led some public officials and academic commentators to question the wisdom of the investor rights provisions. I will say more about these controversies below.

Finally, the OECD’s proposed (and now abandoned) Multilateral Agreement on Investment also would have included private rights of action for investors. Its provisions in this regard closely resembled a typical BIT, with investors having the right to proceed to arbitration and to collect monetary compensation from violator states.8

2. International Trade Law

International trade law is a vast area, encompassing numerous bilateral, regional and multilateral agreements. It will suffice for my purposes to note four of these arrangements: the WTO (incorporating GATT); NAFTA; the Treaty Establishing the

---

4The U.N. maintains a searchable database of BITs on the UNCTAD website—see http://www.unctadxi.org/templates/DocSearch____779.aspx. When last visited, the website listed 44 BITS to which the United States is a party.
5NAFTA Art. 1110.
6NAFTA Art. 1105(1). See generally NAFTA Arts. 1102 (national treatment), 1103 (most-favored nation treatment), 1110 (expropriation and compensation).
7See NAFTA Art. 1135.
8The negotiating text of the MAI may be found at 1998 BDIEL AD LEXIS 33.
European Community (the EC Treaty); and the United States Constitution. I recognize, of course, that the EC Treaty and the U.S. Constitution are much more than simply “trade agreements,” but it is their trade-related provisions, as interpreted by their respective high courts, that are of interest here.

Although these trading arrangements differ in many particulars, they are strikingly similar as to many core substantive obligations. All four arrangements expressly limit or eliminate tariffs on trade among their members. 9 All four arrangements place severe limitations on quotas and other quantitative restrictions. 10 And all four systems prohibit discriminatory taxation and regulation that disadvantages commerce from other member states for the purpose of protecting domestic firms against foreign competition. 11 Many other similarities might be noted.

The four systems differ importantly, however, regarding the standing of private parties to invoke the rules. Under the law of the WTO, only member governments may bring disputes into the dispute resolution process. Private parties may lobby their governments to do so, of course, but the ultimate decision to pursue a case is reserved to national governments.

Putting aside the investor rights provisions noted above (as well as heretofore unused private rights of action under the largely meaningless NAFTA side agreements on labor and the environment), NAFTA also denies standing to private parties. Only the three member governments can initiate a case to enforce the core trade commitments

---

10See GATT Arts. XI & XX; NAFTA Art. 309; Treaty of Rome Arts. 28 & 30. Under the U.S. Constitution, protectionist quantitative restrictions are prohibited by judicial interpretation, the so-called “Dormant Commerce Clause.”
11See GATT Art. III; NAFTA Art. 301. Under the EC Treaty, such discriminatory measures will be found to have “equivalent effect” to quantitative restrictions, and thus be prohibited under Article 28 unless they can be justified by certain “mandatory requirements” such as public health. (see Article 30). The leading case remains Rewe-Zentral AG v. Bundesmonopolverwaltung Fur Branntwein (Cassis de Dijon) ECJ Case 120/78, [1979] ECR 649. Similar jurisprudence has evolved in the Dormant Commerce Clause cases in the United States. See David Currie, The Constitution of the United States, Chap. 3.(University of Chicago Press: 2000).
Regarding tariffs, quotas, nondiscrimination commitments, and the like.\footnote{12}{To be sure, private parties can appeal certain disputes arising in national administrative agencies to “binational panels” as an alternative to appellate review in national courts. But these cases simply afford an alternative mechanism for the enforcement of national law, and do not confer standing on private parties to enforce principles of NAFTA law \textit{per se}.}

Within Europe, the situation is different. Private interests can resist the enforcement of laws that violate the EC Treaty in the courts of member states. Controversial issues may be referred to the European Court of Justice for a ruling, which the courts of member states treat as binding.\footnote{13}{See Cassis de Dijon, supra.} In an indirect way, therefore, private parties may be said to have standing to enforce the trade rules of the EC Treaty against member states that would otherwise violate them.

The United States presents a similar picture. Private actors can challenge state laws that they believe to violate the trade-related principles of U.S. Constitutional jurisprudence in state or federal court. A ruling to the effect that a state law in unconstitutional will ordinarily be accompanied by an order directing state enforcement authorities not to enforce it.

With regard to the \textit{remedy} that is available when a party challenging the legality of a member state law obtains a favorable ruling, the four systems also exhibit some important differences. The WTO system requires member states adjudged to be in violation of WTO rules to conform their behavior within a “reasonable period of time.”\footnote{14}{WTO Dispute Settlement Understanding (DSU) Art. 21(3).} If the member state fails to do so (and the “reasonable period will be fixed by arbitration if necessary), the complaining member and the violator must negotiate over the possibility of trade compensation (usually, substitute trade concessions by the violator to “compensate” for the violation). If those negotiations fail, the complainant may withdraw trade concessions that it has made to the violator (i.e., retaliate) in an amount “equivalent” to the harm done by the violation.\footnote{15}{Some commentators argue that the equivalence requirement can be understood, in a rough way, to implement a rule of expectation damages. See Sykes (2000).} The magnitude of retaliatory suspension of concessions is also subject to binding arbitration. In cases where the violator conforms
its behavior within a “reasonable time,” however, the mainstream view is that the complainant has no rights to trade compensation or retaliation, or to compensation of any other sort. The view that some commentators have questioned this proposition, however, finding precedent in international law for retrospective remedies, and noting that a few dispute panels (mostly under the old GATT system) have recommended remedies that are retrospective (such as reimbursement of wrongfully collected antidumping duties). See Mavroidis (2000).

Further, nothing in the structure of the system prevents WTO members from “settling” for monetary compensation. Such a settlement has occurred once to my knowledge, in a case involving a challenge to the U.S. Copyright Act brought by Europe under the WTO TRIPs agreement. U.S. law did not require the collection of royalties for music played at certain smaller eating and drinking establishments, and was adjudged to violate TRIPs. In lieu of amending the Act, the United States ultimately agreed to pay approximately $1 million per year in compensation to European artists. See Bhala and Attard (2003). The Copyright case is discussed at length in Grossman and Mavroidis (2003).

The NAFTA dispute resolution system is quite similar to that of the WTO. Under Chapter 20, disputes that cannot be settled through consultations are referred to an arbitral panel. If the panel rules in favor of the complaining member, the losing party must comply with the ruling, offer compensation to the prevailing party, or else suffer retaliation in the form of the suspension of “benefits of equivalent effect.” There is no monetary remedy.

\[\text{16One might argue that such a system encourages cheating because there is no penalty for it unless the cheater is caught and still refuses to stop within a “reasonable time.” No fully satisfactory explanation for this aspect of the system exists, although Schwartz and Sykes (2002) offer a speculation. They argue that the bulk of disputes involve good faith differences in interpretation of WTO law, and that litigating such disputes to conclusion may create an important positive externality in the form of useful precedent for the hundred-plus other states bound by the same ambiguous “contract.” The rule denying a right of compensation or retaliation for violations unless cured within a reasonable time encourages parties to litigate to a final judgment. It may also assist developing countries, whose legal capacities are limited and who may inadvertently fail to comply with WTO law quite regularly.}\]
The situation is a bit muddled in Europe, but a damages remedy is available in some cases against a member state for a violation by that state of its EC treaty. A leading case on monetary compensation is *Francovich v. Italy*, in which the Court of Justice held that workers injured by the failure of the Italian government to implement a Commission Directive to protect workers in bankrupt firms might have an action for damages against the Italian government. Subsequently, in *Brasserie du Pecheur v. Germany*, the Court of Justice ruled that Community Law afforded a right to damages under three conditions: “the rule of law infringed must be intended to confer rights on individuals; the breach must be sufficiently serious; and there must be a direct causal link between the breach…and the damage.” The second factor is often the key issue, and the question whether a breach is “sufficiently serious” turns in significant part on courts’ assessments of whether the national government in question has committed a flagrant breach of EC law or has instead acted in good faith (albeit illegally) in an area where it has considerable discretion.

The situation in the United States has changed somewhat in recent years. Until fairly recently, cases challenging state laws under the Dormant Commerce Clause sought only declaratory or injunctive relief. Sitting in the background since the Reconstruction Era, however, was 42 U.S.C. §1983, which provides for private rights of action against any “person” who, under “color” of state law, deprives any individual of “any rights, privileges, or immunities secured by the Constitution and laws” of the United States. The remedy for a violation of §1983 includes money damages and attorney fees (pursuant to 42 U.S.C. §1988). Only in modern times did litigants begin to advance the theory that state laws, regulations and administrative actions in violation of the Dormant Commerce

---

17(cases C 6 & 9/90) [1992] IRLR 84.
19Id. ¶51.
Clause amounted to a violation of §1983. The first decisions rejected this theory, holding that the Commerce Clause creates no “individual rights” and merely constrains the activities of states. But in *Dennis v. Higgins*, the Supreme Court held that a violation of the Commerce Clause by a state does give rise to a cause of action under §1983 and to a claim for attorney fees under §1988. Later decisions by lower courts have awarded monetary damages for the injuries suffered by private plaintiffs due to Commerce Clause violations, although reported litigation in the area to date is sparse.

In sum, trade agreements present a mixed landscape on both the issues of standing and remedy. The two entities considered here with the deepest degree of economic integration, Europe and the United States, afford some private rights of action to enjoin member states from enforcing laws and regulations that violate core trade commitments. Both systems also open the door to monetary remedies to a limited extent. The WTO and NAFTA do not provide private rights of action with respect to trade commitments, nor do they provide monetary remedies even for the member governments with standing to bring cases (although nothing precludes monetary settlements).

B. Prior Commentary on Standing and Remedy in International Economic Law

Recent writing on standing and remedy in the trade and investment areas has

---

21See, e.g., Consolidated Freightways Corporation of Delaware v. Kassel, 730 F.2d 1139 (8th Cir. 1984); J & J Anderson, Inc. v. Erie, 767 F.2d 1469 (10th Cir. 1985).
22*498 U.S. 439 (1991).*
24Among the reasons for the paucity of litigation is the confusing body of precedent regarding what constitutes a “person” acting under color of law and what “immunities” such “persons” enjoy. Generally speaking, municipalities are considered “persons” and enjoy no immunity from suit, even when they act in good faith. However, municipal liability is limited to actions that represent the “policy or custom” of the municipality. State governments, by contrast, are completely immune from suits for damages unless they waive their immunity. Individual officials (state or municipal) may be sued in their “individual capacity” (as distinguished from suits against their government employers), but they enjoy various immunities as well, ranging from absolute immunity for some functions (like the actions of legislators) to qualified or good faith immunity for other types of actions. The result of this hodgepodge is that for a plaintiff to recover money damages from a state, either the state must have waived immunity, or the suit must be brought against a state official in their “individual capacity” and the plaintiff must overcome whatever “immunity” is afforded to the defendant. Then, damages may be sought from the individual defendant, and will be obtained from the state itself only if the state has a policy of indemnifying its employees against liability. Given all of these potential hurdles, it is no surprise that successful actions for damages appear to be quite rare. For an introduction to this confusing body of law, see Erwin Chemerinsky, *Federal Jurisdiction*, 4th ed., chapter 8 (Aspen: 2003).
focused on three issues: the wisdom of recent developments in NAFTA investor rights litigation; the proper role of private parties as amicus curiae in the WTO; and the wisdom of trade sanctions for violations of WTO law.

Much of the commentary on NAFTA investor rights litigation has been highly critical of the decisions in cases brought by private parties. The greatest concern is that the concept of “expropriation” is being applied too elastically, and that “regulatory takings” are imprudently found to constitute compensable expropriation. A case of particular concern in this context is Metalclad Corp. v. United Mexican State, which involved the denial of an environmental permit to a waste disposal site, and resulted in a $16 million award to the American complainant. Another notable case was S.D. Myers, Inc. v. Canada, which involved a claim by an American company producing a fuel additive in Canada that a Canadian ban on inter-provincial trade in the additive was enacted for protectionist reasons rather than the stated health reasons. Canada settled the case for $19 million rather than litigate to conclusion.

Commentators make the point that such decisions imply a far broader “takings” doctrine under NAFTA than under U.S. domestic law, and generally argue against compensation for these regulatory takings. Among other things, they contend that the conventional cost internalization argument for compensation is flawed, and that any benefits from compensation as “insurance” are available in the private insurance market. See Been and Beauvais (2003). The same issues feature prominently in earlier literature on domestic takings. See Epstein (1980); Blume and Rubinfeld (1984). Notwithstanding

---

25 Another concern was that arbitral panels would read NAFTA Article 1105 broadly, and that its requirement of “fair and equitable” treatment would become a license to award damages for any government action that the arbitrators viewed as unfair. Much of the concern here flowed from the arbitrators reading of Article 1105 in Metalclad Corp. v. United Mexican States, 40 I.L.M. 36 (2001), see esp. ¶100-01. That issue was perhaps laid to rest by a recent “clarification” adopted by the NAFTA Free Trade Commission, which stipulates that Art. 1105 requires no more than observance of the customary international law standard regarding minimum treatment of aliens. See the discussion of the Metalclad case in John H. Jackson, William J. Davey & Alan O. Sykes, International Economic Relations, 4th ed. (West 2002), 1153-66, [hereafter Jackson, Davey & Sykes].


27 Information on both cases may be found at http://www.naftaclaims.com. Much of the critical commentary is collected in Brower (2003).
the critique of individual decisions, however, there has been little or no criticism of private rights of action in the investment area _per se_. All of the commentators seem to accept their wisdom, as long as they are governed by the proper substantive rules.

In the trade area, the absence of private rights of action in the WTO and NAFTA also seems to be accepted by most of the commentators. See, for example, Trachtman and Moremen (2003). But the United States and many non-governmental organizations favor a more limited opportunity for private actors to participate in the WTO dispute process as amicus curiae. In two controversial decisions, the WTO Appellate Body ruled that dispute panels and the Appellate Body itself could accept such submissions at their discretion. These decisions received a chilly reception from the membership at large, however, and reportedly only the United States spoke in defense of them before the Dispute Settlement Body. Developing countries argued that their result was to shift the balance of power toward the well-funded non-governmental organizations of the developed world, whose agendas were often at odds with the interests of developing countries. One infers from these events that any proposal for creating private standing would be roundly rejected by the WTO membership as a whole.

If there has been little advocacy of private standing, there has been much discussion of changing the remedy for violations of WTO rules. Numerous commentators have observed that trade sanctions cause substantial welfare losses. See Guzman (2002). It is considered a puzzle as to why international trade agreements do not use money damages as a sanction, which are said to constitute “transfers” without the deadweight costs of trade retaliation. See Guzman (2004). Such observations lead various commentators to advocate changes in WTO sanctions. Charnovitz (20001 & 2002); Davey (2001). Prominent practitioners have recently joined the chorus, arguing for

---

28But see Shell (1995).
29See Jackson, Davey & Sykes, supra, 315-17.
30See Statement by Uruguay at the General Council Regarding the Decision by the Appellate Body Concerning Amicus Curiae Briefs, WT/GC/38 (December 12, 2000), excerpted in Jackson, Davey & Sykes, supra, 303-05.
monetary penalties calibrated to the damages suffered by injured exporters (though stopping short of advocating a private right of action).31

Another line of commentary pushes in the same direction. It has long been claimed that developing countries are at a disadvantage in the WTO dispute resolution process. The explanation usually includes the notion that developing countries have small markets and are thus unable to affect the prices received by exporters to their markets (in effect, they lack any “monopsony power” that can be exploited through tariffs). Consistent with this thesis, Bagwell, Mavroidis and Staiger (2004) indicate that in not one instance has a developing country actually exercised its retaliation rights in a WTO dispute.32 This perceived imbalance of power has led a group of African nations to propose that monetary penalties be introduced into the system, and has led Mexico to propose that retaliation rights be subject to auction. See Bagwell, Mavroidis and Staiger (2003).

A few scholars have weighed in on the other side of this debate, suggesting that trade retaliation is preferable to monetary remedies because it will be more effective at inducing compliance with trade commitments. The most thorough exposition of this argument is that of Nzelibe (2005). See also Goldstein and Martin (2000). I will return to these arguments below.

II. Investment

This section offers a rationale for the existence of private rights of action for money damages in investment agreements, a rationale that does not apply to trade agreements as Section III will demonstrate. As noted earlier, the impetus for modern BITs (which provide the model for NAFTA Chapter Eleven) was a growing concern

31See Marco Bronckers and Naboth van den Broek, Trade Retaliation is a Poor Way to Get Even, Financial Times, June 24, 2004, p. 15.
32Busch and Reinhardt (2003) look further at the experience of developing countries in WTO dispute settlement, and contend that they are disadvantaged by their relative incapacity to pursue effective strategies in the early phases of disputes.
about the expropriation of foreign investments in developing countries during the mid-20th century. The concern reached its zenith when developing countries as a group took the position in the United Nations that customary law did not require compensation for expropriation. Investors with sunk investments in those countries faced greater risk, and had an incentive to lobby their governments for international agreements to reduce it. This political pressure was the impetus for developed nations to seek to enter BITs with developing countries.

From the developing countries’ perspective, BITs were a double-edged sword. They plainly limited the capacity of governments to expropriate existing foreign investments without compensation, a limitation that worked to their disadvantage, other things being equal. But they also yielded an important benefit. Investors in developing countries (as elsewhere) will require a risk premium on their investments to ensure themselves an expected competitive rate of return. A risk of uncompensated expropriation thus increases the price of imported capital to developing countries. A reduction in this risk likewise lowers the cost of foreign capital, which creates rents for domestic factors of production that work with foreign capital. Those factors will in turn offer political support for any policy that reduces expropriation risk. This explanation of why many developing countries agreed to BITs is a conventional one in the literature. See Guzman (1998), Elkins, Guzman and Simmons (2005).34

33For a simple diagrammatic exposition of the benefits of capital inflow to the country that hosts new foreign investment, see Peter Lindert, International Economics 9th ed. 547-49 (Irwin: 1991).
34Guzman (1998) argues that developing countries accepted BITs because they lowered the cost of foreign capital in this fashion, and made direct foreign investment in the territory of the signatory more attractive than in the territories of other nations that had not executed BITs. Guzman further argues that the net effect on BITs on developing countries as a group may have been adverse – if they collectively possessed monopsony power in the capital market, a policy that permitted uncompensated expropriation might have allowed them to exploit it. He argues that proposed BITs induced the abandonment of this collectively valuable policy by forcing each developing nation into a sort of Prisoner’s Dilemma. Each was tempted to defect from the collectively preferred regime by the prospect of obtaining a competitive advantage over others, and once they all defected they were collectively worse off than before. Whether or
Because a central objective of the investment agreements was to induce foreign investors to make new investments in developing countries at a lower interest rate, the utility of a private right of action for money damages is obvious. To see why, consider a world of BITs without the private action. In the event of an uncompensated expropriation or similar action, an investor would have to lobby its own government to take some sort of action against the violator state. The investor might be politically inefficient in this process for any number of reasons. It may be unable to offer enough political benefits in return for the governments’ assistance. Its government may have diplomatic reasons for declining to take any action, or for declining to retaliate against the violator in any effective way. And even if some retaliation were forthcoming, it might do nothing to compensate the investor for its losses. Considerable risk for investors would remain, and the risk premium on new investments would reflect it. A credible promise of monetary compensation to investors, by contrast, in an amount set by neutral arbitrators, goes much farther to reduce investment risk and to achieve the developing countries’ goal of lowering the cost of foreign capital.

Moreover, for any developing country that does not plan to engage in significant expropriation or other prohibited activity, a credible promise of monetary compensation to investors imposes few if any offsetting costs. If expropriation never occurs, compensation and the attendant litigation costs need never be paid. The possibility of socially excessive litigation because of a divergence between the private and social costs of suit—a concern explored in Section III below—is then minimal. A promise of monetary compensation to investors is thus a cheap commitment device for states with benign intentions toward investors, and a cheap way for states with more benign

---

not Guzman is correct in this account, it seems that on any account the developing countries accepted BITs
intentions than others to signal their “type.” As long as the capital importing nation is confident that it does no wish to engage in prohibited behavior, therefore, the private right of action on behalf of foreign investors is not a burden on it but a clear benefit.

**Implications: The NAFTA Chapter Eleven Controversy.** The explanation developed here for private rights of action under investment agreements offers some additional support to the argument of Been and Beauvais (2003) that a “regulatory takings” doctrine under NAFTA Chapter Eleven is undesirable. The commitment by a developing country to a private right of action for investors is a low-cost commitment or signal that it will respect investors’ property rights only to the extent that “expropriation” is defined to include acts that the national government is unlikely to want to undertake. An expansive regulatory takings doctrine, by contrast, can sweep in many acts of “expropriation” that arguably result from the resolution of regulatory uncertainty (like the denial of the environmental operating permit in *Metalclad*) or the emergence of new information about the subject of regulation (such as the health issues associated with the gasoline additive in *S.D. Myers*).35

Been and Beauvais note that such regulatory policy changes typically do not confer private rights of action for damages under U.S. law, and it seems unlikely that developing countries would wish to provide broader “insurance” against regulatory policy changes. The extensive literature on legal transitions suggests that compensation for policy changes (or insurance against them) can encourage over-reliance on policies that may change, and may chill desirable change as well. See, e.g., Kaplow (1986b); Shaviro (2000). Further, the capacity of international dispute panels accurately to distinguish

---

35I stipulate that some commentators have a less benign view of the actions of the respective governments in those cases, and take no position on the ultimate merits of either case.
well-intentioned or desirable regulatory policies from those driven by protectionism or other forms of capture is limited. The tendency in international law is to restrict the scrutiny of domestic regulation by international dispute panels to fairly narrow issues such as the presence of clear discrimination or the violation of various procedural requirements, and to foreclose open-ended inquiry into the wisdom or legitimacy of regulation through cost-benefit balancing and the like. See Sykes (1999, 2003), Trebilcock and Soloway (2002). An open-ended “expropriation” doctrine that permits challenges to regulatory outcomes under investment agreements would thus stand in contrast to the treatment of national regulation in other areas of international economic law, a fact that casts additional doubt on the wisdom of compensation for regulatory takings under NAFTA.

III. International Trade

The (capital) importing nation that is party to an investment agreement wishes to commit or signal to private (capital) exporters that their investments are secure against government interference. The private right of action for compensatory damages facilitates this government-to-firm commitment, in a way that a mere government-to-government commitment cannot. Trade agreements are different in an important way—(goods and services) importing nations have no direct interest in making (goods and services) exporters more secure or confident that market access commitments will be respected. An importing nation, call it A, will make commitments that benefit exporters in another nation, call it B, only to the degree that the government of B will make reciprocal commitments that benefit A’s exporters. For this reason, trade agreements are better structured as government-to-government commitments, and a private right of
action for damages may actually be counterproductive. Indeed, private standing irrespective of the remedy may be counterproductive. Section A develops this argument.

Even if standing under trade agreements is best limited to governments, it remains to consider what remedy the agreement should select. Section B offer several new reasons why trade sanctions might be preferred to money damages, despite what some commentators note as the apparent inefficiency of trade sanctions.

A. Standing

The political rationale for trade agreements differs importantly from the rationale for investment agreements. In the absence of trade agreements, trade policy reflects the political equilibrium between the interest groups opposed to imports (domestic import-competing industries) and interest groups that favor imports (domestic import-consuming industries and consumers). Export industries have only a modest stake in policy, reflected in the degree to which a more restrictive import policy may result in some retaliation against them by foreign governments acting unilaterally. When trade agreements become possible, by contrast, nations can negotiate reductions in their trade barriers in exchange for reciprocal reductions by others. Exporters thus have a more direct stake in policy, and will lobby their home governments to secure market access concessions abroad. The new political equilibrium will result in a more open trading system.

Unlike the situation with investment agreements, importing nations do not enter trade agreements out of a desire to lower the price of imports. They view any reduction in their own trade barriers, and the attendant price of their imports, as a “concession” that is

---

36 The proposition that trade agreements arise to facilitate government-to-government market access commitments is a standard one in the modern economics literature. See, e.g., Bagwell and Staiger (2002). A smaller strand of literature, however, explores whether trade agreements may also facilitate valuable commitments by governments to their domestic firms. Because such theories do not suggest any role for private rights of action by foreign firms, however, I do not address them here. Useful references include Maggi and Rodriguez-Clare (1998) and Staiger and Tabellini (1999).
only tolerable in return for concessions by trading partners. Holding constant the concessions by trading partners, therefore, importing nations will prefer that foreign exporters face greater risk regarding the security of market access commitments. Indeed, importing nations would gladly reimpose their barriers to imports if they could do so without suffering any retaliation or punishment by foreign governments. From the importing nations’ perspective, therefore, a private right of action to compensate foreign exporters for violations of trade agreements is undesirable, other things being equal.

The reader may well observe, however, that even if there is little desire by the governments of importing nations to make their own market access concessions more secure, every importing nation is also an exporting nation. Exporters value secure market access commitments (or a promise of compensation should they be withdrawn), at least to the degree that the risks associated with their withdrawal cannot be priced up front to ensure an expected competitive return on investment. As in the investment area, therefore, exporters with sunk investments (or inelastically supplied factors of production that earn inframarginal rents), the returns to which would be threatened by the violation of trade agreements, might be expected to lobby their governments for a private right of action on their behalf.37

The distinction identified thus far between investment agreements and trade agreements may be summarized as in Table 1:

---

37There is perhaps a shade of difference between the investment and trade areas in this regard, because the impetus for modern investment agreements arose during a time when developing countries were retreating from their commitment to compensate for expropriation and related actions. This retreat may have been unanticipated by many investors and may thus have threatened the quasi rents on substantial amounts of sunk investment, leading to especially strong political pressure for measures to restore the previous commitment by developing countries to compensation. Liberalization in the trade area has proceeded steadily, by contrast, and exporters never enjoyed a private right of action. The degree to which
Table 1

<table>
<thead>
<tr>
<th>Investment</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trade</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The private damages action, which makes exporters more secure, is critical to achieving the goals of the importing nation in the investment area, but objectionable to the importing nation in the trade area. This key difference suggests that the private damages action will be *more attractive* to states contemplating an investment agreement. The distinction drawn here, however, is insufficient to establish that the private damages action will not also be attractive (if less so) to states contemplating a trade agreement. I now turn to some further considerations that weigh against the private damages action in the trade area.

*The Value of Political Filters.* Recall the fact that goods and services importing nations are interested in making commitments that benefit foreign exporters only if they result in reciprocal commitments to benefit their own exporters. Likewise, an importing nation may have little interest in honoring a previously-made commitment to a foreign exporter unless the violation of that commitment would result in some detriment to its own exporters, perhaps in the form of direct retaliation or at least some damage to the importing nation’s reputation that would harm its ability to secure future trade commitments. This fact suggests the virtue for trade agreements of what I will term a “political filter.”

To elaborate, a central tenet of public choice theory is that some interest groups sunk investments have been threatened by potential violations of trade agreements may be more modest, therefore, although this point is surely somewhat speculative.
are better organized than others, and thus better able to influence the political process. Indeed, this observation affords the standard explanation for protectionist trade policies, which (perhaps with rare exception) lower the national economic welfare of nations that employ them—the harm from protectionist policy is often borne by a diffuse group of poorly-organized consumers, while the benefits inure to well-organized import-competing industries.

Not all industries are equally well organized either. Industries comprised of a large number of small producers, for example, may have difficulty overcoming the transaction costs and free rider problems associated with efforts to influence the political process. Industries with a few larger firms may be better able to overcome free rider problems and transaction costs, by contrast, and individual firms in such industries may have high enough stakes in the outcome of policy decisions that they are willing to incur the costs of acting alone to influence political representatives.

Thus, consider some generally applicable legal rule under a trade agreement, such as a requirement that forbids regulatory measures that discriminate against foreign commerce (the “national treatment” obligation in WTO parlance), and imagine that a party to a trade agreement has violated that obligation. In the absence of any private standing to enforce the rule, political officials in the nation whose exporters are aggrieved by the violation must decide whether to bring a compliance action themselves. If the violation affects a well-organized export industry, officials may expect significant political rewards from such an action. If the violation affects a poorly-organized industry, by contrast, they may expect few political rewards from bringing the enforcement action.

The violation of the national treatment rule will also have political consequences for officials in the violating state. An inadvertent violation that confers rents on a poorly-
organized import-competing industry, for example, will generate little political support for those officials. But if the violation confers substantial rents on a well-organized import-competing industry, it may be a source of considerable political benefits to the officials who can claim credit for it. These observations suggest that the political officials who become parties to trade agreements can enhance their mutual welfare by retaining the right to determine which enforcement actions are brought.

The basic idea can be formalized very simply as follows: Consider two states, A and B (although the analysis readily generalizes to the multilateral case). Assume for simplicity that states are unitary actors, each with a “political utility” function—let \( U() \) denote the utility function for A and \( V() \) denote the utility function for B. Each function is additive and its arguments will become clear in a moment.

Industries in each state may be divided into two categories, exporting industries and import-competiting industries. State A has comparative advantage in the set of industries that export to B, call them industries in the set \( AB \), with its members denoted \( \alpha_i, i = 1 \text{ to } n \). State B has comparative advantage in the industries that export to A, call them industries in the set \( BA \), with its members denoted \( \beta_j, j = 1 \text{ to } m \).

States A and B have entered a trade agreement, which contains some generally applicable legal rule (say, national treatment). Import-competing industries in each state do not like the national treatment rule (which forecloses a range of protectionist policies that could benefit them), while exporting industries in each state benefit from the national treatment rule. Political officials in state A can thus gain political utility from a national treatment rule that is enforced for industries in \( AB \), and but will suffer a loss of political utility from a national treatment rule that is enforced for industries in \( BA \). Let \( u(\alpha_i) \) denote the political benefit derived by officials in state A from the enforcement of the
national treatment rule in (its exporting) industry $\alpha_i$, and let $u(\beta_j)$ be the political detriment suffered by officials in state A from the enforcement of the national treatment rule in (its import-competing) industry $\beta_j$. Choose the utility scale so that $u(\alpha_i) \geq 0$, all $\alpha$, and $u(\beta_j) \leq 0$, all $\beta$.

The situation in state B is exactly the opposite. Let $v(\alpha_i)$ denote the political detriment suffered by officials in state B from the enforcement of a national treatment rule in (its import-competing) industry $\alpha_i$, and let $u(\beta_j)$ be the political benefit derived by officials in state A from the enforcement of the national treatment rule in (its export) industry $\beta_j$. Choose the utility scale so that $v(\alpha_i) \leq 0$, all $i$, and $v(\beta_j) \geq 0$, all $j$.

If the national treatment rule is enforced in all industries, officials in each state enjoy total utility equal to the sum of their respective utilities from enforcement in each industry:

$$U = \sum_{i=1}^{n} u(\alpha_i) + \sum_{j=1}^{m} u(\beta_j) \quad \text{and} \quad V = \sum_{i=1}^{n} v(\alpha_i) + \sum_{j=1}^{m} v(\beta_j).$$

But the parties can almost certainly do better. Imagine that some exporters in state A are very poorly organized, say, those in industry $\alpha_2$. Suppose for simplicity that $u(\alpha_2)=0$. But imagine that the import-competing firms in state B are better-organized, so that $v(\alpha_2) < 0$. The parties to the trade agreement would then experience a (political) Pareto improvement if they agreed not to enforce the national treatment rule in industry $\alpha_2$—officials in state A would be no worse off, while officials in state B would enjoy utility gains.

The point is much more general, and does not depend on the presence of export industries for which the political utility of enforcement is zero (indeed the utility scale is completely arbitrary). State A can agree to forego enforcement of the national treatment
rule on behalf of its industries where the political gains to its officials are “small” and the political costs to officials abroad are “large,” while state B can make a reciprocal promise. The political costs in each state from foregone enforcement on behalf of export industries under such an arrangement can be far outweighed by the political gains from foregone enforcement in politically powerful import-competing industries.

One can take the analysis a step further and imagine that the parties write an elaborate contingent contract, specifying industry-by-industry where the national treatment rule applies and where it does not. Sykes (1991) solves a mathematically analogous continuous problem (here one might imagine a continuum of industries). For present purposes, it is enough to note that the solution has the following properties: For any point on the parties’ Pareto frontier (and thus associated with an optimal treaty), a shadow price exists that allows units of each party’s utility to be converted into units of the other’s utility. An optimal treaty will provide that the parties forego enforcement of the national treatment rule (or any other rule, for that matter) in any industry for which the political utility gain to officials in the violator state “outweighs” the political utility loss to officials in the state harmed by the violation, using the treaty’s shadow price to convert utilities into the same units.

A detailed contingent contract of that sort would be extremely costly to write, however, especially given the wide array of rules found in modern trade agreements. The parties may thus prefer cruder and cheaper rules. A simple rule with considerable potential appeal is that parties will not bring enforcement actions on behalf of politically weak industries. As long as the poorly-organized export industries in state A are not as poorly organized on average in state B, and vice-versa, an exchange of reciprocal promises to forego enforcement on behalf of poorly-organized industries can leave both
sides at a considerably higher level of utility.

How might the parties implement such a rule? The obvious way is to omit any private rights of action from their agreement. Because it is costly for the parties to bring enforcement actions themselves, they will only bring actions on behalf of exporters who offer sufficient political rewards in exchange. It is precisely the group of politically weakest exporters whose cases will be ignored under this arrangement.

It is perhaps instructive here to note how the investment situation is different. Because the goal of the capital importing nation is to lower its cost of capital, it has an interest in assuring all capital exporters that it will respect their investor rights, not just the ones who are politically efficacious in their home countries. The reason is the distinction emphasized earlier—capital importing countries benefit directly from making foreign exporters more secure, but goods and services importing countries only benefit to the extent that foreign exporters reward their governments for greater security and induce them to grant reciprocal commitments.

Two further considerations support this general line of analysis. First, in newer trade agreements such as NAFTA and much of the WTO, each member state confronts a number of transition issues. The task of ferreting out all national and sub-national regulation that might run afoul of the WTO Technical Barriers Agreement, the Sanitary and Phytosanitary Measures Agreement, and so on, is not a trivial one. Thousands of regulatory measures are potentially in play (especially after the WTO agreement made clear that its obligations apply to state and local regulation as well as national regulation), and many nations may be constrained in their capacity to make conforming changes quickly (especially developing countries). In the face of many potential “transition” violations, therefore, the parties may well desire to limit enforcement actions to those of
the greatest political importance to aggrieved exporters. Private rights of action, by contrast, might tax the resources of nations seeking to comply with their obligations, and distort the timing of the compliance agenda (from the standpoint of maximizing joint political gains). This problem is likely to be far less acute in the investment area, where the basic rules have been established for a long time (although NAFTA decisions have shaken them up, perhaps unwisely, as noted earlier).

Second, legal interpretations developed in one case inevitably have consequences for others. Officials concerned with their political welfare must worry that precedents established in an enforcement action brought on behalf of their exporters will come back to haunt them in an action brought against them by foreign exporters. Only by reserving standing to themselves, and thereby retaining control over the arguments put forward in litigation, can officials ensure that their export interests do not advance legal theories that are lacking in “net” political value. The same issue can arise in the investment arena, to be sure, and one might interpret the uneasy reaction of NAFTA member governments to recent decisions as a manifestation of this problem. Trachtman and Moremen (2003) make a similar point. Levy and Srinivasan (1996) make the related point that private actors may bring actions when, for diplomatic reasons, their governments might prefer forbearance.

Objections. One possible objection to this line of reasoning is that it ignores the resources that political officials must expend to bring compliance actions. Perhaps the direct costs of briefing and arguing cases, and the related expenses of deciding which claims to bring, more than offset the benefits from the “political filter” mechanism outlined above.

Although this objection surely raises a logical possibility, it should not be
exaggerated. As a practical matter, governments can (if they wish) push much of the cost of compliance actions back onto the private sector. Indeed, it is routine in WTO practice for the private interests who seek compliance actions on their behalf to supply legal assistance for the purpose of developing the legal analysis, writing briefs, and the like. Nations could also agree to allow privately funded counsel to argue the cases (although neither NAFTA or WTO practice allows private parties to appear at present). It is simply not the case, therefore, that governments must bear high litigation costs when they choose to act as a political filter.

A second possible objection is that the analysis proves too much. If governments can gain by interposing themselves as political filters between private interests and the dispute resolution process under trade agreements, why do private rights of action emerge in some contexts nevertheless? As noted earlier, both the European Union and the U.S. constitutional system afford limited private rights of action to parties aggrieved by violations of trade rules.

A partial explanation for the disparity across systems is that private rights of action may not have been contemplated by the framers of either the U.S. Constitution or the EC Treaty. In the United States, the important “free trade” principles associated with the Dormant Commerce Clause were developed by the Supreme Court years after the Constitution was written. Likewise, the modern powers of the European Court of Justice in the trade area may not have been anticipated during the founding period, and decisions such as Cassis de Dijon, analogizing regulatory trade barriers to measures having an “equivalent effect” to quantitative restrictions, may have come as a surprise. Private rights of action in each system may thus be an unanticipated creation of the courts rather than a mechanism desired by the political founders.
Even if this suggestion is correct, however, it is at best an incomplete answer to the puzzle. Had private rights of action been viewed as seriously problematic in either system, political actors could have changed the law to extinguish them. This observation suggests a second important consideration bearing on the wisdom of private rights of action.

*Ex Post Political Oversight.* Both Europe and the United States have political bodies (the Congress and the EC Commission) with the capacity to address sensitive trade matters directly. NAFTA and the WTO do not have such bodies—only a costly process of amending the pertinent treaty text, which requires unanimous acceptance, can override judicial decisions that political actors find unpalatable.

The presence of a body like the Congress or the Commission reduces the value of interposing a political filter at the front end of the dispute resolution process. As long as the body has not lost too much flexibility to constitutional rules that it cannot change, it can handle the most politically charged matters directly, and can correct judicial decisions that become politically unacceptable over time (decisions under the Dormant Commerce Clause in the United States, for example, can always be overridden by an affirmative exercise of the Commerce power). In such a system, the danger of politically objectionable judicial decisions surviving for any length of time is much less, and the opportunity to shift the costs of bringing cases onto the private sector is more attractive. Put differently, bodies like the Congress and the Commission can serve as *ex post* political filters, undoing the undesirable decisions while avoiding the costs of sorting the larger number of cases that arise *ex ante*.

A possible objection to this line of reasoning, however, is that constitutional restrictions may preclude Congress and the Commission from discriminating among
industries in a politically desirable fashion. The Equal Protection clause of the U.S. Constitution, for example, likely precludes policies that treat industries differently solely in accordance with their political efficacy (some other “rational basis” is almost certainly required to justify disparate treatment). An *ex ante* political filter may have the advantage of circumventing such “bothersome” legal constraints. Still, the *ex post* mechanism has a potentially sizeable cost advantage that may on balance favor private standing.

**B. Remedy: Trade Sanctions and Money Damages**

Regardless of the rules governing the standing of private parties, some remedy must be made available to parties who have standing. We thus return to the puzzle of why trade agreements such as NAFTA and the WTO select the seemingly inefficient remedy of trade sanctions with their attendant deadweight costs, in preference to a remedy like money damages that is a mere “transfer.”

As noted at the outset, nothing in the current law of NAFTA or the WTO precludes the use of monetary payments to resolve disputes. “Compensation” is allowable (and indeed preferred) to trade retaliation, and nothing restricts the form that compensation might take. Yet, with very rare exception, WTO and NAFTA cases are not settled in this fashion (putting aside NAFTA investor rights cases). Thus, by “revealed preference” of sorts, we must infer that trade retaliation is preferred in general by violator states to the alternative option of money compensation, at least given the implicit reservation prices of complainants. This section suggests a few considerations that may help explain this state of affairs. Before turning to those issues, however, I wish to rebut one argument that has appeared elsewhere.

*Are Trade Sanctions Preferred Because They Better Induce Compliance with Trade Agreements?* As noted earlier, some scholars, most notably Nzelibe (2005), have
argued that trade agreements utilize trade sanctions rather than money damages because the sanctions mechanism is more effective at inducing parties to comply with their commitments. Trade retaliation can be targeted at powerful export groups in the violator country, the argument runs, which will motivate them to encourage their government to cease the violation. Likewise, trade retaliation provides at least temporary benefits to import-competing firms that compete with the imports that are targeted. Their political support for retaliation makes the use of the retaliatory sanction a credible threat. Money damages, it is argued, are inferior in both respects. Their costs are borne by a diffuse group of taxpayers in the violator state, who will not organize to lobby their government to avoid monetary liability—money is too “cheap” from the perspective of the violator state for monetary penalties to be an effective deterrent to misconduct. Further, the beneficiary of a monetary sanction will be the national treasury of the complaining state, and no interest group in that state will have much interest in pushing for its government to pursue the sanction.

I find this line of reasoning unconvincing for two reasons. First, to the degree that damages are a “cheap” penalty from the perspective of violators, they can always be increased. At some point, monetary penalties would become a sufficient burden on the treasury of even the wealthiest trading nations that a violator nation would prefer to comply with its commitments rather than to pay damages, and the prospect of collecting the money would be an appealing prospect for potential complainants. The claim that damages will not induce compliance, therefore, must be modified to something like the following: *compensatory* damages (perhaps measured by the loss of rents to foreign exporters injured by the violation) will not suffice to induce nations to comply with their commitments.
Even this modified claim is problematic, for it implies that parties to trade agreements should actually prefer to utilize a compensatory damages remedy. To see why, assume, arguendo, that the claim is correct. Then, parties prefer a state of the world in which they violate the rules and pay compensatory damages, to state of the world in which they comply with their commitments and avoid damages. If this is true, however, then there is a clear opportunity for efficient breach (in a political sense), and the parties could enjoy Pareto gains by incorporating a compensatory damages remedy into the agreement. Violators by hypothesis enjoy higher utility even when bearing the cost of damages, and injured exporters should be indifferent between compliance with the agreement and full compensation for violations. Under these circumstances, a trade retaliation remedy that sufficed to induced compliance would eliminate this source of joint gains, and it would not be in the mutual interest of the parties to employ it.

The suggestion that trade retaliation is preferred over damages in trade agreements because it is a more effective mechanism for inducing compliance thus suffers from two logical problems. It is not a convincing rationale for the widespread use of trade retaliation in my view. I now turn to some other considerations that weigh against monetary remedies.

*Are Monetary Remedies a Mere “Transfer?”* The claim that money damages are a simple transfer is misleading for one well-known reason and one perhaps not so well-known. First, governments face budget constraints, and must finance money damages through taxation. Putting aside “lump sum” taxes, oft invoked by economists but hardly ever used in practice, taxes create their own distortions. The choice between trade sanctions and money damages then is not a simple choice between a mechanism that creates welfare costs and one that does not. And although it may well be true that
methods to raise the money for a damages award can be found that cause less distortion than the trade sanctions for which the award substitutes, the magnitude of the difference need not be dramatic. This may be especially true in developing countries, some of which still rely heavily on trade taxes to raise funds for their national treasuries—trade sanctions against them raise inefficient barriers to their exports, but monetary awards against them might well force an increase in inefficient tariffs on their imports (which might even hurt the exporters from the complaining nation).

Second, the claim that money damages are a simple transfer neglects the lessons of the economic literature on the private versus social value of litigation. Landes and Posner (1975) note that the Becker/Stigler prescription in the criminal law area for higher penalties and lower probabilities of sanction (to save on administrative costs) becomes problematic if private enforcers could collect the higher penalties—increases in the penalty may induce more expenditure on enforcement rather than less. See also Polinsky (1980). In the same spirit, Shavell (1982) explores the difference between the private incentive to file civil claims for damages, and their social value, in a tort setting. The social return to a prospective lawsuit is the value of the ex ante change in risky behavior, which can be measured by the reduction in the expected costs of accidents induced by prospective liability. The social cost of a lawsuit is its litigation cost—the damages payment from a defendant to a plaintiff is simply a transfer. Because there is no necessary relationship between the damages payment (and the attendant incentive to file suit) on the one hand, and the social gains from prospective liability on the other, litigation may occur too often or too infrequently from an economic standpoint. The analysis is refined somewhat in Menell (1983) and Kaplow (1986a), but the essential point remains the same. See also Shavell (2004).
These insights have potentially important implications for the wisdom of private damages actions in international economic law. The social gains from such actions and the private gains may be quite different. Consider a simple illustration. Figure I depicts a market for imported goods in some importing state. The curve $I_s$ is the import supply curve, reflecting the quantity of imports available to the state in question at each price. The curve $I_d$ is an import demand curve, constructed from the excess of domestic demand over domestic supply for the good in question at every price. Assume that a pertinent trade agreement requires “free” trade, which would produce an equilibrium price $P_f$ and quantity of imports $Q_f$ where $I_s$ and $I_d$ intersect. But the importing nation violates the agreement, and imposes a tax of $t$ on imported goods, yielding an equilibrium price to
domestic consumers of \( P_v \) and net price to foreign sellers of \( P_v-t \); import quantity drops to \( Q_v \). The deadweight cost of the violation in this simple framework is equal to the area \( abd \). The rectangle \( P_v db(P_v-t) \) is tax revenue.

Consider the incentives to file a case under these assumptions. Foreign sellers lose surplus in the amount \( P_v ab(P_v-t) \). If the case will succeed with certainty, they will bring it as long as their litigation costs do not exceed this value. Because their gains from suit exceed the deadweight cost [triangle \( acd \) appears to be considerably smaller than rectangle \( P_v cb(P_v-t) \)], the distinct possibility arises that foreign sellers will file a case even if their litigation costs would exceed the social gains from correcting the violation. And if foreign sellers do not bear the litigation costs of the violator state and the tribunal (as in NAFTA or the WTO), there is an even greater danger that a case will be filed even though the social costs of litigation exceed the gains from correcting the violation.\(^3\)\(^8\)

The analysis is not quite so simple, however, because the mere threat of litigation may deter the violation in the first place. If violations never occur, and thus litigation never ensues, the danger of socially excessive litigation evaporates. This point is akin to the observation in Shavell (2004) that the danger of socially excessive tort litigation is diminished when injurers are subject to a negligence rule rather than a strict liability rule, and can avoid suit by exercising due care.

It follows that the efficiency of private damages actions for violations of trade agreements turns, inter alia, on two empirical questions. First, when violations occur, how will the likely costs of litigation compare to the social gains of correcting the violation (assuming litigation will lead to its correction)? Second, to what extent will the prospect of private damages actions deter violations altogether?
Regarding the first question, Figure I hints at a potentially important general concern. Many violations of trade agreements result in large transfers—tax revenue for the government at the expense of foreign producers and domestic consumers, quota rents to domestic importers who hold importation rights at the expense of foreign producers and domestic consumers, producer surplus for favored trading partners at the expense of other trading partners who are discriminated against, and so on. Deadweight losses arise in all of these scenarios to be sure, but they may be much smaller than the transfers away from injured foreign exporters. Such situations present a clear danger of socially excessive litigation costs. The problem is all the more acute, of course, to the degree that nations would prefer to pay damages in perpetuity than to conform their behavior, a possibility that cannot be completely discounted.

Regarding the second question, it is questionable whether trading nations can avoid litigation costs by simply “complying” with the rules. Flagrant cheating occurs in the system to be sure, but many disputes involve good faith disputes over the interpretation of legal obligations or the application of legal principles to the facts. Many violations may simply be “accidental,” especially by developing countries with limited compliance capacity. In other cases, the law is unclear or its application under the circumstances is fairly debatable—trading nations can then “comply” with the rules only by abandoning potentially legitimate policies, and they may be unlikely to do so in many cases.

Money Damages and Developing Countries. Developing countries might be quite threatened by a system of monetary damages, at least if they were calibrated to compensate foreign exporters for the damages they suffer due to illegal acts. Many

---

38 A dynamic perspective does not change the basic conclusion. Think of the amounts in Figure I as “per
developing countries regularly suffer severe shortages of hard currency, and will be hard-pressed to pay significant damages in any currency that they cannot print. And because their legal infrastructure is often weak, their ability to avoid damages judgments by simply “complying” with the law is limited as noted above—indeed, within the WTO, developing countries often complain that they do not fully understand how to implement their obligations, and seek assistance for “capacity-building” in this regard. Monetary damages might thus produce a significant chill on their domestic regulatory initiatives, and expose them to considerable liability for inadvertent violations.

Although some developing countries have nevertheless suggested that a monetary remedy be introduced into WTO law as noted earlier, Davey (2001) explains the nature of the remedy that they have in mind. They do not advocate a system of “compensatory damages,” but a sliding scale system of fines set on the basis of factors such as GDP and per capita income. Mexico’s proposal for auctioned retaliation rights provides similar protection to the treasuries of developing countries, in that no nation would be compelled to participate in any auction. Existing proposals for monetary remedies, therefore, are aimed at imposing substantial monetary costs on violations by wealthier states while imposing small costs on violations by developing countries. It is thus possible that no system of monetary damages lies in the “core” of the bargaining game between North and South, at least not as the game has been structured so far. The problem might be overcome in future bargaining by some sort of issue linkage, but the opportunity for such an arrangement perhaps has not yet arisen.

Is it Possible to Substitute Money Damages for Trade Sanctions? Finally, one must reflect on the practical challenges of designing a credible system of monetary period” amounts, sum them over time and discount to present value—the same issues arise.
penalties to replace trade sanctions. Imagine that monetary penalties were assessed against a violator state, which then declined to pay the judgment. What consequence would follow? Unless nations are confident that reputation or repeat play concerns will alone ensure respect for monetary awards—and they plainly lack confidence in these forces to police the rest of their obligations—it seems that any system of monetary liability would have to be backed up with the threat of trade sanctions. At the end of a dispute, violator states would then have the opportunity to choose between paying the assessed damages or suffering the trade retaliation that follows from non-payment. In many ways, such a system is the mirror image of the *status quo*—violator states today choose between retaliation under agreed standards for calibrating it, and negotiated compensation, which can be monetary or anything else. It is not clear what advantage an alternative system of monetary penalties would really offer.

To be sure, part of the problem might be avoided if trading nations posted bonds *ex ante*. But this system too could quickly break down without a threat of sanctions to induce bonds to be posted in timely fashion and to be replenished when needed.

The Mexican auction proposal, of course, does embody a blend of monetary payments and trade retaliation. The proposal is apparently making little headway in the WTO at large, however, perhaps because of the concern noted above—it would amount to a sizable transfer from North to South, and may not yet lie in the core of any near-term global bargain.

I conclude with one last point about the contrast between trade and investment. If a threat of trade sanctions is necessary for monetary penalties to be enforceable in a trade setting, why not also in the investment context? The answer is that an alternative threat sits in the background of an investment agreement to induce governments to pay
damages, wielded by the global capital market. Any capital importing nation that defaults on its obligations under a BIT or similar instrument will face an increased cost of capital that lowers its welfare. Because the importing nation entered the BIT in the first place to avoid paying such risk premia, it will typically respect a judgment against it absent some substantial change in circumstances.

_Damages in Europe and the United States._ The remarks above perhaps prove too much in light of the difference between NAFTA and the WTO on the one hand, and Europe and the United States on the other. Damages awards in the U.S. and European systems are sometimes available to successful plaintiffs, and are collected, without the need for a trade sanctions regime in the background.\(^39\) The reason is likely multifaceted, relating to the existence of a political body that sits above the member states with various coercive options, the power of judges to order individual public officials to act and to punish those who do not, and perhaps a stronger pull of reputation in these more deeply-integrated systems. Monetary penalties assuredly may have more appeal, other things being equal, where the member states can be trusted to pay them without a fight.

Conclusion

The role of private actors in the enforcement of international economic law varies greatly between investment and trade agreements, and considerably within the trade area as well. A private right of action for money damages is especially valuable in the investment arena, where a key objective of the parties is to lower the cost of capital for new, irreversible investments. Nations can cheaply achieve that objective by credibly promising to compensate for “expropriation” and related practices, so long as those

\(^{39}\) Recall _Dennis v. Higgins_ in the United States and the _Francovich_ line of cases in Europe, discussed supra.
practices are clearly and appropriately defined. This objective is absent in the trade area, where trading nations will often prefer to act as “political filters” by denying standing to private parties and thereby retaining the ability to block private enforcement actions that can reduce joint political welfare. The existence of a political body with the capacity to reverse politically unfortunate judicial decisions (an *ex post* political filter) is at least an imperfect substitute for the denial of private standing, and will make private rights of action more attractive, other things being equal. Finally, it is by no means clear that private rights of action for damages would be superior to trade sanctions, or that they could be substituted for trade sanctions as practical matter in a system such as the WTO or NAFTA. These observations can begin to explain, in broad brush, the patchwork of rules regarding private standing and remedies under existing investment and trade agreements.
REFERENCES


Guzman, Andrew (2002). The Cost of Credibility: Explaining Resistance to Interstate Dispute

Guzman, Andrew (2004). The Design of International Agreements (mimeo).


Readers with comments should address them to:

Alan O. Sykes
University of Chicago Law School
1111 East 60th Street
Chicago, IL  60637
Alan_Sykes@law.uchicago.edu
Chicago Working Papers in Law and Economics
(Second Series)

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)
67. Yannis Bakos, Erik Brynjolfsson, Douglas Lichtman, Shared Information Goods (February 1999)
68. Kenneth W. Dam, Intellectual Property and the Academic Enterprise (February 1999)
70. Cass R. Sunstein, Must Formalism Be Defended Empirically? (March 1999)
71. Jonathan M. Karpoff, John R. Lott, Jr., and Graeme Rankine, Environmental Violations, Legal Penalties, and Reputation Costs (March 1999)
75. Richard A. Epstein, Deconstructing Privacy: and Putting It Back Together Again (May 1999)
76. William M. Landes, Winning the Art Lottery: The Economic Returns to the Ganz Collection (May 1999)
77. Cass R. Sunstein, David Schkade, and Daniel Kahneman, Do People Want Optimal Deterrence? (June 1999)
78. Tomas J. Philipson and Richard A. Posner, The Long-Run Growth in Obesity as a Function of Technological Change (June 1999)
79. David A. Weisbach, Ironing out the Flat Tax (August 1999)
81. David Schkade, Cass R. Sunstein, and Daniel Kahneman, Are Juries Less Erratic than Individuals? Deliberation, Polarization, and Punitive Damages (September 1999)
82. Cass R. Sunstein, Nondelegation Canons (September 1999)
83. Richard A. Posner, The Theory and Practice of Citations Analysis, with Special Reference to Law and Economics (September 1999)
84. Randal C. Picker, Regulating Network Industries: A Look at Intel (October 1999)
90. David A. Weisbach, Should the Tax Law Require Current Accrual of Interest on Derivative Financial Instruments? (December 1999)
95. David Schkade, Cass R. Sunstein, Daniel Kahneman, Deliberating about Dollars: The Severity Shift (February 2000)
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)
113. William M. Landes, Copyright, Borrowed Images and Appropriation Art: An Economic Approach
    (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)
    Perspective (April 2001)
121. Douglas G. Baird and Robert K. Rasmussen, Control Rights, Priority Rights and the Conceptual
    Foundations of Corporate Reorganization (April 2001)
122. David A. Weisbach, Ten Truths about Tax Shelters (May 2001)
    Judicial Law Clerks (June 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)
134. Richard A. Epstein, The Allocation of the Commons: Parking and Stopping on the Commons
    (August 2001)
    Laws (September 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)
142. Daniel N. Shaviro and David A. Weisbach, The Fifth Circuit Gets It Wrong in Compaq v.
    Commissioner (February 2002) (Published in Tax Notes, January 28, 2002)
143. Warren F. Schwartz and Alan O. Sykes, The Economic Structure of Renegotiation and Dispute
144. Richard A. Epstein, HIPAA on Privacy: Its Unintended and Intended Consequences (March 2002,
    forthcoming Cato Journal, summer 2002)
145. David A. Weisbach, Thinking Outside the Little Boxes (March 2002, Texas Law Review)
146. Eric A. Posner, Economic Analysis of Contract Law after Three Decades: Success or Failure (March
    2002)
| 194. | David A. Weisbach and Jacob Nussim, The Integration of Tax and Spending Programs (September 2003) |
| 200. | Douglas Lichtman, Rethinking Prosecution History Estoppel (October 2003) |
| 201. | Douglas G. Baird and Robert K. Rasmussen, Chapter 11 at Twilight (October 2003) |
| 205. | Lior Jacob Strahilevitz, The Right to Destroy (January 2004) |
| 208. | Richard A. Epstein, Disparities and Discrimination in Health Care Coverage; A Critique of the Institute of Medicine Study (March 2004) |
| 209. | Richard A. Epstein and Bruce N. Kuhlik, Navigating the Anticommons for Pharmaceutical Patents: Steady the Course on Hatch-Waxman (March 2004) |
| 213. | Luis Garicano and Thomas N. Hubbard, Specialization, Firms, and Markets: The Division of Labor within and between Law Firms (April 2004) |
| 216. | Alan O. Sykes, The Economics of Public International Law (July 2004) |
225. Christine Jolls and Cass R. Sunstein, Debiasing through Law (September 2004)
228. Kenneth W. Dam, Cordell Hull, the Reciprocal Trade Agreement Act, and the WTO (October 2004)
230. Lior Jacob Strahilevitz, A Social Networks Theory of Privacy (December 2004)
231. Cass R. Sunstein, Minimalism at War (December 2004)