Optimal Sanctions in the WTO: The Case for Decoupling (and the Uneasy Case for the Status Quo)

Alan O. Sykes
Stanford University
Optimal Sanctions in the WTO: The Case for Decoupling (and the Uneasy Case for the Status Quo)

Alan O. Sykes*

Abstract

Various commentators have suggested that the current system of trade sanctions for violation of WTO obligations be replaced with financial compensation. The details of these proposals vary, but one option is to allow firms injured by violations to recover damages. This paper questions the wisdom of such proposals, and argues that the current system in which those injured by violations do not reap the benefit of sanctions – a “decoupled” sanctions regime in economic parlance – may well be superior for a number of reasons. The paper also reviews and refines the view of current WTO practice as an analogue to expectation damages in private contracts. The original version of this paper was prepared for the interdisciplinary workshop on The Calculation and Design of Trade Sanctions in WTO Dispute Resolution, at the Graduate Institute in Geneva, 2008. The revised version will appear in The Law, Economics and Politics of Retaliation in WTO Dispute Settlement, forthcoming from Cambridge University Press.

My assigned topic for this conference on WTO dispute resolution is “Improvements and New Approaches from a Legal Perspective.” As shall quickly become clear, however, I am not an advocate of any particular “improvements” or “new approaches.” In general, I believe that the WTO dispute resolution system has evolved in a reasonably defensible and satisfactory fashion in response to a number of competing pressures. The overall level of compliance with WTO obligations is high (though by no means perfect), disputes are reasonably modest in number and formal

* James & Patricia Kowal Professor of Law, Stanford University. I have received thoughtful comments on earlier versions of this paper from Chad Bown, William Davey, John Jackson, Petros Mavroidis, Joost Pauwelyn, Robert Staiger, and from the participants in the conference on the Calculation and Design of Trade Sanctions in WTO Dispute Settlement at the Graduate Institute of International Studies, Geneva, July, 2008.
sanctions are comparatively rare. The system is not broken and does not require fixing.

I do not mean to suggest that everything about the WTO dispute system functions perfectly. Certainly, one can quibble with the details of dispute resolution in individual cases, as a number of other contributors to this conference do quite persuasively. No doubt more can be done in cases that proceed to a formal suspension of concessions, for example, to devise better compliance counterfactuals and to measure their trade implications more accurately.¹ But nothing in these thoughtful critiques makes out a case for any fundamental change in the design and operation of the dispute resolution process.

Rather than advocate for particular reforms, therefore, I take this opportunity to raise doubts about some that have been proposed. In particular, the burgeoning commentary on the dispute settlement system includes much criticism of trade sanctions as the “punishment” for breach of WTO obligations. Among other things, critics point out that trade retaliation is economically costly and may harm the nation that retaliates. Some critics further observe that trade sanctions do not compensate the victims of breach, taken to be the injured exporters, but instead insulate other import-competing firms from competition at the expense of consumers while harming the “innocent” exporters targeted by sanctions. Such thinking often leads to the suggestion that, in one fashion or another, monetary compensation should replace or supplement trade sanctions. See, e.g., Pauwelyn (2000); Charnovitz (2001); Lawrence (2003); Davey (2005); Bronckers & van den Broek (2005); Bagwell, Mavroidis & Staiger (2007).

A system of monetary compensation must address a range of issues. Would exporters have a private right of action? How would monetary damages be calculated? Who would receive the monetary compensation, the treasury of the complaining nation(s), the injured exporters, or someone else? Would compensation extend to all harm caused by the violation (“retroactive compensation”), or just the prospective harm that occurs after a violator nation refuses to comply with an adverse ruling within a reasonable

¹ See, e.g., the contributions to this volume by Sebastian, Keck, Evenett, and Bown & Ruta.
period of time? What would happen if a WTO member refused to pay a monetary judgment against it? This list of issues is no doubt incomplete.

I will not take on all of these matters in this brief contribution, but instead will focus on the narrower question whether it would be desirable for exporters injured by a violation to receive financial compensation for their losses. The current system of sanctions in the WTO may be characterized as “decoupled” in the sense that recalcitrant violator nations may pay a price in the form of trade sanctions, but the exporters injured by violations do not receive compensation for the harm caused by violations (except indirectly if sanctions eventually induce compliance). The purpose of this paper is to examine the wisdom of moving away from a decoupled system in the direction of a system that provides direct compensation to exporters. Although the analysis is not conclusive, decoupling has a number of virtues that may counsel in favor of retaining it. The remainder of this essay will explicate those virtues and, along the way, lay out the basic argument for the system as is, or what one might term the somewhat “uneasy case” for the status quo.

I. The Expectation Damages Analogy

As other participants in this conference have noted, it is difficult to assess the effectiveness of WTO sanctions without a conception of their purpose. A long and inconclusive debate exists in the literature in this regard, with some commentators taking the view that the purpose of formal sanctions is to induce compliance with WTO obligations, and others taking the view that centralized oversight of sanctions is intended to ensure that the price for deviation from WTO obligations is not “too high.” Efforts to

---

2 Such compensation might be provided through a variety of mechanisms. One can imagine a system that requires violator nations to compensate exporters harmed by a violation, a system in which violators simply choose to negotiate compensation for exporters in lieu of some other penalty (in principle possible now, although not employed to date outside of the U.S. – Copyright case), or a system in which complainant governments use retaliatory tariffs to raise revenue to provide compensation to injured exporters. Each of these possibilities raises both distinct and common issues. The analysis to follow will elide these details to a considerable degree, but is aimed primarily at proposals to make financial compensation mandatory in lieu of the existing system of trade sanctions.

3 See, e.g., the contribution to this volume by Pauwelyn.
resolve this debate based on the treaty text seem inconclusive because the

text is confused and contradictory. Snippets of text can support a variety of

Schwartz & Sykes (2002). Arbitrators seeking to ascertain “purpose” from
the text have noted these inconsistencies as well. See United States –
Continuing Dumping and Subsidy Offset Act of 2000, Recourse to
Arbitration by the United States under Article 22.6 of the DSU (Byrd
Amendment), ¶¶6.2-6.7.

In my view, much more can be learned about the purpose of sanctions
from a thoughtful examination of the history and structure of WTO dispute
settlement, assisted by some basic economic observations about the
challenges of designing an optimal treaty mechanism under conditions of
uncertainty. Following this approach, one quickly comes to an appreciation
of how, in a limited and imperfect sense, the system of sanctions under the
WTO Dispute Settlement Understanding (DSU) plays a role somewhat akin
to that of expectation damages in private contracts, in that it facilitates
arguably desirable deviations from the letter of the bargain under politically
exigent circumstances (what economists often term “preference shocks”).
This observation is not new, but it is useful to review the basis for it here as
it serves as a building block for consideration of decoupling.

A. Sanctions During the GATT Years

The WTO dispute settlement system is now 13 years old, but it was
preceded by 48 years of disputes under GATT. During that time,
collectively authorized sanctions played no material role in the system.
Disputes were resolved (if at all) through diplomacy, often aided by GATT
panel reports. Despite the absence of a dispute resolution system with
“teeth,” however, the GATT held together well and brought about dramatic
reductions in average tariff rates around the world, particularly among the
developed economies.

GATT members respected most of their commitments most of the
time for two reasons. First, it is in the interest of nations that seek
international cooperation on trade (and other matters) to demonstrate their
reliability to other nations by adhering to their commitments. Nations that
fail to do so, at least with sufficient frequency, will be deemed unreliable
and other nations will not cooperate with them in the future. In economic
parlance, nations adhere to commitments because it is in their interest to develop a *reputation* for cooperative behavior. [See. e.g., Guzman (2008).]

Second, and related, GATT members harmed or threatened by violations could employ “self help” by engaging or threatening to engage in unilateral retaliation against violators. It is well known in both the theoretical and experimental literature on game theory that threats of retaliatory defection can (though they need not) sustain cooperation over time when strategic interaction is open ended (“infinitely repeated”) and when parties to strategic interaction are sufficiently patient so that the short term gains from cheating do not outweigh the long term costs of undermining cooperation. Previous commentators have made the argument that unilateral self help was a useful part of the “glue” holding the GATT system together, see Sykes (1992); Bown (2004), and indeed the now standard economic models of trade agreements presuppose that mutual threats of retaliatory defection are key to sustaining cooperation. E.g., Bagwell & Staiger (2002).

If the incentives created by reputational concerns and unilateral retaliation enabled GATT to flourish without formal sanctions, what was the role of dispute resolution under GATT? The answer lies in the fact that the informal methods for sustaining cooperation can only function well if cooperative behavior can be identified accurately. Disputes can arise over the question whether actions breach the agreement due to differences of opinion about the facts or the law. GATT panels (and the consultations that preceded them) provided an opportunity for the disputants to clarify the facts regarding the conduct at issue and to seek legal interpretations of ambiguous obligations. The fact that panel formation could be “blocked” over much of the history of GATT did little to undermine this role, in that concerned parties could generally draw a (correct) adverse inference from the decision by a disputant to block the panel process.

In short, the role of dispute resolution during the GATT years was to provide the information necessary for cooperation to be sustained through informal mechanisms grounded in reputation and self help. Centrally administered sanctions played no role, and their absence was of not great consequence.

B. The WTO Innovation
Why did the parties to the Uruguay Round negotiations see fit to replace the GATT dispute system with the DSU, which introduced a real prospect of centrally authorized sanctions which, indeed, has resulted in the imposition of sanctions in a number of cases? A naïve and incorrect answer is that the GATT suffered from too much “cheating,” so that a new system was needed to introduce tougher penalties for breach. A moment’s reflection will reveal that the new system did no such thing.

As under GATT, WTO disputes are initiated by signatories who claim to have been harmed by a violation (to have had their benefits from the bargain “nullified or impaired.”) There is no “public prosecutor.” And just as under GATT, if a respondent loses its case and declines to comply with the ruling against it, it suffers two adverse consequences – damage to its reputation, and the costs of any trade (or other) sanctions that the complaining nation(s) may impose. Nothing in the WTO system augments the ability of complaining nations to impose sanctions, and there is no authority for collective retaliation by other nations. If a complaining nation is weak and lacks the ability to impose meaningful retaliation, therefore, it must content itself with the favorable ruling alone, just as under the old GATT system. It cannot fairly be said that the DSU enhances the penalties for breach.

So what has really changed? Aside from an appellate process to review disputed legal interpretations, the primary change lies in the fact that the magnitude of trade (or other) sanctions is now subject to central oversight through an arbitration process under DSU Article 22. The systematic effect of this arbitral process – both in its design and in practice – has been to reduce the magnitude of the trade (or other) sanctions that the complaining nation(s) may impose.

This innovation was, as argued elsewhere, a response to the “aggressive unilateralism” that developed during the waning years of GATT. See generally Bhagwati & Patrick (1990); Schwartz & Sykes (2002). Under GATT, nations aggrieved by an actual or alleged breach of obligations would employ self-help remedies to a degree of their own choosing because nothing in the GATT dispute system offered any practical constraint on the magnitude of those measures. The United States, through Section 301 of the Trade Act of 1974, was particularly aggressive in this regard. Reputational concerns were at best an imperfect constraint on unilateral sanctions because it was difficult for other nations to observe when sanctions were “excessive”
– indeed, GATT itself did not establish any clear principle for determining what level of self-help was acceptable (consider GATT Article XXIII, which provided for the suspension of “such concessions or other obligations under this Agreement as the [Contracting Parties collectively] determine to be appropriate in the circumstances.”). Likewise, it was questionable whether excessive retaliation could be adequately disciplined by counter-retaliation, and the worry instead was that such scenarios could escalate into trade wars that might unravel cooperation.

Consequently, toward the end of GATT, many signatories came to believe that some mechanism was needed to reign in unilateral retaliation. This proposition is, I believe, quite widely accepted. But implicit in the proposition is the notion that retaliation can be “excessive.” This observation raises a core issue that has proven a source of controversy through the years – why might retaliation be “excessive?”

To frame the issue, suppose that a WTO member breaches its obligations. The facts and the law clearly establish the breach, and the violator refuses to change its behavior. Why should the sanction for such recalcitrance be limited at all? If the complaining nation wants to engage in “massive” retaliation for the purpose of inducing the violator to comply with its commitments, why should the system object? Indeed, why not permit all WTO members to retaliate collectively for the purpose of coercing compliance, even the members whose interests are not impaired by the breach? Two ostensibly different answers have been offered through the years, although as I will suggest below, they actually collapse into a single logical point.

Economically oriented commentators offer an answer grounded in familiar ideas from the literature on contracts. Contracts are entered under conditions of uncertainty and positive transactions costs. It is prohibitively costly for the parties to contracts to anticipate every possible contingency that may materialize and to write down the optimal response to every such contingency. Accordingly, even contracts that include built-in flexibility mechanisms will generally fail to anticipate every situation in which deviation from the bargain is desirable. The parties can thus benefit from

---

4 See Maggi (1999) for a model in which multilateral trade agreements are valuable because they orchestrate collective retaliation.
alternative mechanisms that facilitate deviation. Renegotiation is always an option, but it may be subject to bargaining failure and hold up problems that in turn have unfortunate *ex ante* consequences. As an alternative, contracting parties may limit the remedies for breach, as by subjecting themselves to a legal system that awards only expectation damages. Such a system requires the breaching party to compensate the party harmed by breach, and thus encourages parties to breach precisely when net gains arise.

An essentially identical story may be told about many treaties, and the WTO in particular. The set of issues addressed by the WTO is vast, and WTO agreements are negotiated in the face of enormous uncertainty about the future. Even with built-in flexibility mechanisms such as GATT Articles XIX, XX, XXI, and XXVIII, it is unlikely that the negotiators can anticipate all of the circumstances in which deviation from the agreement is jointly optimal (for purposes of this argument, the welfare metric is arbitrary). Renegotiation is not an attractive way to deal with this problem given the vast number of WTO members. Hold-up problems become all the more acute as the number of parties to renegotiation increases, a problem compounded by the most-favored nation obligation. Thus, some other device is needed to allow deviation from commitments that are not optimal *ex post*. A mechanism akin to the expectation damages regime for private contracts can strike a sensible balance. Of course, the task of designing such a mechanism is immensely more complex here, but the observation that WTO agreements are analogous to “incomplete contracts” offers a clear rationale for a system that does not penalize breach “excessively.”

Some commentators, however, balk at the claim that compliance with international law generally, or WTO law in particular, may at times be “inefficient.” When asked why the WTO should nevertheless constrain the magnitude of retaliation for breach, these commentators typically answer that a need arises to cabin retaliation to maintain the stability of the system – to prevent trade wars that might cause an unraveling of cooperation. This argument for limited retaliation, of course, is simply another account of efficient breach. It says, in effect, that the parties to the arrangement are jointly better off by limiting retaliation so as to sustain greater cooperation.

---

5 See Horn, Maggi & Staiger (2006) for more on the incomplete contract perspective on trade agreements.
over time, even if the result is long term deviation from the law by some parties in some situations – *a fortiori*, violations are sometimes efficient.

In sum, a strong theoretical case can be made for limiting the penalty for breach of obligations under GATT and other WTO agreements. If signatories can deviate from the bargain without paying an excessive price, the bargain becomes more valuable and more trade concessions can be expected *ex ante*. It thus comes as no surprise that when the DSU introduced a serious prospect of formal sanctions into the dispute settlement system, it concurrently limited their magnitude – DSU Article 22.4 states that such measures should be “equivalent to the level of nullification or impairment.” As under GATT, aggrieved parties can still penalize violations with trade sanctions, but those sanctions are subject to the ceiling implied by “equivalence” (putting aside the prohibited subsidies cases\(^6\)), to be enforced by an arbitral process.

As argued elsewhere, this particular innovation in the DSU may be understood as a crude analog of expectation damages under private contracts. Just like expectation damages, the ceiling on retaliation associated with Article 22.4 can work to prevent the price of breach from becoming too high. Schwartz & Sykes (2002).

To be sure, the analogy to expectation damages fails in other respects. Expectation damages in the private contract setting not only ensure that the price of breach is not too high, but also that it does not become too low. Neither Article 22.4 nor any other part of the DSU address the latter issue -- if a complaining nation lacks the capacity to retaliate in a meaningful way, nothing in the DSU strengthens its hand.\(^7\)

The analogy to expectation damages is imperfect in other ways as well. Most importantly, the question of how to measure and operationalize “equivalence” is much less clear than in the private contract setting. The

\(^6\) Three arbitral panels have ruled that the countermeasures available in prohibited subsidies cases need not satisfy the “equivalence” standard that applies to other cases, relying on language in Article 4.10 of the WTO Agreement on Subsidies and Countervailing Measures. See Jackson, Davey & Sykes (2008), Section 7.5.

\(^7\) Of course, as in the days of GATT, reputational considerations may do much to protect the interests of weaker nations. And the addition of TRIPs to the WTO system may empower weaker nations considerably with a prospect of cross-retaliation.
next section discusses this issue further, and explains why it is unnecessary for injured exporters to be compensated for the system to achieve a reasonable calibration of sanctions.

II. “Equivalent” Trade Sanctions

In the private contract setting, the measurement of expectation damages is conceptually straightforward. The court determines the monetary value of what the promisee would have received if the contract had been performed, and deducts the value of the promisee’s position following breach. What is the analogous calculation in the WTO setting?

To answer that question, one must first specify who is to be conceived as the injured promisee. On one view, the promisees are the exporters whose market access (or some other right) has been impaired by a violation. This line of thinking leads quickly to the idea that payment of compensation to exporters for economic losses might be a sensible mechanism for imposing sanctions for breach, an idea to which we will return in the next section.

But it is not obvious that the “promisees” in the WTO bargain are properly viewed as private actors. Treaties are entered by political officials, and the field of public choice teaches that such officials commonly behave in such a way as to maximize their political interests, as measured by votes, campaign contributions, and the like. An “optimal” treaty, therefore, might be expected to maximize the joint welfare of signatory officials in a rough sense. Those same officials may also be viewed as the beneficiaries of the treaty bargain, and from this perspective, the cost of violations arises in the form of a loss of political welfare for officials in nations harmed by the violation. The connection between the loss suffered by political officials, and the loss suffered by private actors, may be close or attenuated, depending on such factors as whether the private actors in question are well organized politically. See Sykes (2005).

This line of thinking implies that sanctions might ideally take the form of measures that allow aggrieved political officials to restore their political welfare in the face of a breach. Of course, compensation to exporters may serve that function, but it is by no means the only option.
In particular, if officials in a nation harmed by a violation may impose trade sanctions, they can use them to pursue one of two politically valuable strategies. One option is to target the sanctions at foreign exporters who seem likely to lobby their own governments to cure the violation in order to end the sanctions. Political officials who pursue this strategy successfully will reap political rewards from exporters harmed by the original violation. A second option is to employ trade sanctions to protect import-competing firms that value such protection, and that will directly reward their officials for providing it. Savvy political actors may be expected to choose between these two strategies according to which one benefits them the most.

Plainly, trade sanctions also impose political costs on officials in the violator nation. Exporters harmed by trade sanctions may be expected to withdraw political support from the domestic officials whose policies result in the violation that begets the sanctions.

Thus, if one shifts the focus to political officials and treats them as the beneficiaries of the bargain, trade sanctions seem a perfectly plausible enforcement mechanism.\(^8\) Trade sanctions impose political costs on violators, and allow officials in nations harmed by a violation to reap political benefits. If the sanctions are calibrated properly, they would seem to have as much potential to create proper incentives for compliance or breach (in relation to the political welfare “maximand” in the background) as any monetary compensation mechanism.\(^9\)

But the last statement contains a big “if.” How can the WTO ensure that trade sanctions are calibrated properly – large enough to encourage an

---

\(^8\) In this respect, one must distinguish investment agreements from trade agreements. As to the former, I have argued elsewhere that a private right of actions for damages suffered by individual investors (as is common in bilateral investment treaties, for example) makes good sense. If a capital-importer wishes to lower its cost of capital most effectively, a promise of compensation to private investors for acts such as “expropriation” is a more effective mechanism than a right of retaliation vested in the investor’s home government. For elaboration of this argument and an analysis of why it does not apply to trade agreements, see Sykes (2005).

\(^9\) Of course, the political welfare maximand will not in general correspond to economic welfare conventionally defined. Some of us may lament the disparity, but there is nothing that we can do about it. For this reason, I view efforts to ground the dispute settlement mechanism in conventional welfare economics as futile or counterproductive.
appropriate degree of compliance, and not so large as to discourage efficient (in the political sense) deviation from the bargain?

The modern theory of trade agreements suggests some useful principles. Bagwell and Staiger (2002) develop the concept of “reciprocity” in trade agreements. Roughly, reciprocity refers to a process whereby trade concessions are exchanged in a manner that expands the volume of trade but holds the terms of trade constant. Although they do not explicitly consider how reciprocity might work in the event of a breach of a trade agreement, Howse and Staiger (2006) extend the basic idea to consider the optimal response to changed circumstances (a “preference shock”) that lead a party to a trade agreement to contemplate breach. In their model, each government maximizes a welfare function that is quite general in its specification, and allows political considerations to enter in the form of arbitrary preferences regarding local prices (which reflect the distribution of rents among interest groups). A regime of “expectation damages” — i.e., a regime that restores the welfare of the government harmed by a violation — can be approximated in this framework with a rule whereby breach is followed by trade sanctions that reduce the exports of the breaching party in an amount equal to the reduction in exports caused by the violation (valued at the pre-violation prices). This analysis offers a rough theoretical defense of the approach that has been used in WTO arbitrations to date (again putting aside the prohibited subsidies cases) — calculate the export trade lost as a result of the violation by the complaining nation(s), and then permit sanctions in an amount that would reduce the violator’s exports by the same amount.

To be sure, the real world is more complicated than the simple Howse and Staiger model, which involves trade in two goods with four countries, and in which the only policy variables are tariffs. The WTO involves trade in thousands of goods and services, among over 140 countries, with many different border and domestic policy instruments in play. Nevertheless, their analysis suggests how a principle of “equivalence” that seeks to calibrate the trade effects of sanctions to the trade effects of the original violation may

---

10 In the basic Bagwell/Staiger model, cooperation is sustained by a mutual, reciprocal threat to withdraw trade concessions and return to the noncooperative (Nash) equilibrium that precedes the agreement. The withdrawal of reciprocal concessions is thus an “off equilibrium” threat that does not occur when cooperation is sustained.
represent a sensible, simple rule of thumb that explains the basic approach seen in a number of WTO arbitrations.

III. Compensating Exporters as an Alternative

Even if trade sanctions linked to the trade effects of the violation are a plausible mechanism, might a system of monetary compensation for aggrieved exporters be better yet? Such a system, one might argue, could also restore the political welfare of officials harmed by a violation by alleviating the injury suffered by their exporters. It would also arguably facilitate efficient breach (if a nation wishing to deviate from its commitments can compensate injured exporters and still remain better off itself, breach seems efficient). Further, as others have noted, a system of compensation for exporters would have the advantage of avoiding the deadweight costs of protection associated with sanctions.

Such thinking is open to question for a variety of reasons. First, it is important to recall a well-known inefficiency associated with expectation damages. Although expectation damages may in principle induce breach when and only when it is desirable, they also insure promisees against the loss of their expectancy. By compensating promises for breach even when breach is efficient, “over-reliance” will result – promisees will make excessive irreversible investments that enhance the value of performance to them. This is why, as a general matter, it cannot be shown that expectation damages are superior to a rule that excuses liability for breach altogether. See Shavell (1980). The inefficiency that results from expectation damages disappears, however, if damages are decoupled. If the breaching promisor must bear the cost of breach to the promisee, but the promisee does not receive that cost, both the breach decision and the reliance decision can be optimized.

How might excessive “reliance” manifest itself in the international trade setting? Firms assured of market access to a particular market or full compensation for its withdrawal might well sink excessive costs in anticipation of serving that market. These could include investments in customizing goods or services for the market, or investments in production, sales and distribution arrangements for that market. The notion that international trade can entail significant relationship-specific investments of this sort is well established in the literature. See, e.g., Yarbrough & Yarbrough (1992); Antras & Staiger (2007).
Second, the costs of measuring the economic losses caused by violations could be enormous. It is one thing to do the analysis to construct a counterfactual to measure the aggregate trade volume effects of a violation, as is done now. To take that analysis to the next level, and calculate the economic losses for every exporting firm affected by the violation, would add another layer of complexity. Among other things, the costs for each firm would be needed to determine lost profit, along with an examination of opportunities in each firm to mitigate losses. Furthermore, the lost profits of exporting firms do not capture the full costs of a violation. Rather, the loss is borne by all factors of production that earn rents or quasi-rents from the export opportunity, including the owners of all specific capital involved in the export industry (such as workers with specific human capital). The task of valuing such losses, and aggregating across what might be many firms, workers and others, would be error prone and quite expensive.

Of course, this problem could be averted by a system of “liquidated” damages [as some commentators have proposed, see Bronckers & van den Broek (2005)], whereby a violator pays a fixed amount per period regardless of the nature and extent of the violation. But such a system forfeits almost all of the benefits of a sensibly designed sanctions mechanism. The price of breach would be unrelated to the harm it causes, leading to situations in which the price of breach is excessive and to others in which nations could buy out of their obligations too cheaply.

A third caution regarding the compensation of exporters concerns the claim that monetary compensation is more efficient because it avoids the deadweight costs of protection. This is at best a misleading claim for it ignores the deadweight costs of the tax system that would raise the funds to pay compensation. Once we put the economists’ fiction of lump sum taxation to the side, it becomes apparent that all tax systems in practice

---

11 Limao & Saggi (2008) analyze a formal model of a trade agreement in which cooperation can be sustained either by retaliatory defection on tariff commitments, or by a monetary fine (paid to the other government). They conclude that both mechanisms are capable of sustaining the same degree of tariff cooperation, but that the monetary enforcement mechanism Pareto dominates retaliatory tariffs because fines create no deadweight losses. That result, however, depends on the assumption that the money to pay fines is raised by lump sum taxation. See id., p. 53.
generate deadweight costs. It is possible that the taxes imposed to raise funds for compensation will have smaller deadweight costs than trade protection, but one cannot be sure.

A fourth concern regarding the compensation of exporters relates to the cost and volume of litigation. How would a prospect of compensation affect the number of cases filed in the WTO and the resources invested in them? Simple economic theory suggests that as the returns to litigation increase, the number of cases filed and the resources invested in each case will increase. If exporters were given the equivalent of a private right of action, one can imagine legal entrepreneurs searching far and wide for compensable violations and filing numerous cases on something akin to a contingency fee basis. And even if the right of action remains with governments alone, the prospect of compensation for harm suffered might well ratchet up the political pressure on governments to bring actions (especially if compensation were extended to past harm and not simply to the future harm that follows from a nation’s refusal to comply with an adverse ruling).

Of course, the fact that litigation and litigation costs might increase is not undesirable a priori. Perhaps current levels of litigation and enforcement are too low from a social perspective. But violations of trade agreements often result in private losses that exceed social losses. Protection for domestic industries diverts profits from foreign firms to domestic firms, and the loss suffered by foreign firms may greatly exceed the deadweight costs of protection. Violations can also shift business among foreign firms in ways that cause a loss of profits to some of them far exceeding the social costs of the trade diversion. Indeed, some violations of WTO law (such as export subsidies) may actually enhance global economic welfare. These considerations lead to the distinct possibility of excessive litigation in the WTO system, whereby the amount spent on the pursuit of compensation exceeds the social value of discouraging non-compliant practices. See Shavell (1982).

I do not wish to overstate the case here. It is also possible that the mere threat of litigation will induce a higher level of compliance in many settings such that actual litigation is unnecessary, and the gains from greater compliance might exceed the costs of the occasional cases. The point is simply that the net costs and benefits of an increased incentive to litigate and
to invest resources in litigation is very much uncertain, and will depend on
the balance of various empirical considerations. See generally Sykes (2005).

Fifth, and finally, any consideration of compensation for exporters
must take into account the situation facing developing countries. With
regard to developing countries as potential defendants, compensation
requirements might represent a worrisome threat to their public sector
budgets, doubly so when “violations” may be more accidental than
deliberate due to limited compliance capacity. To be sure, these issues
might be addressed by yet another form of “special and differential
treatment,” but exemptions for developing countries would weaken the
utility of the system and complicate the politics of agreeing on any sort of
reform.

As for developing country plaintiffs, by contrast, compensation
requirements are sometimes thought to be potentially useful because of the
limited capacity of developing countries to retaliate with trade sanctions.
See, e.g., Bagwell, Mavroidis & Staiger (2007). Such arguments are not
fully convincing, however, given the ability of small nations to obtain
authority for cross-retaliation on matters such as TRIPs rights (as in EC –
Bananas and United States – Gambling).

Furthermore, it is a mistake to imagine that substantive rights can be
divorced from the remedies that enforce them. If developing countries lack
retaliatory capacity, it is because they have not made concessions of value to
trading partners that they can take back. And if that is the case, it is not
necessarily in the interest of developing countries to make the obligations
running in their direction more “enforceable.” In particular, if developing
countries receive concessions for which they have not “paid” with
meaningful concessions of their own, a prospect of stiffer sanctions for
breach of those concessions may lead the countries that would otherwise
make the concessions to withhold them in the first place, leaving developing

IV. Conclusion

For the reasons developed above, I believe that the case for replacing
the current system of trade sanctions with some form of direct compensation
to injured exporters is a weak one. Despite its imperfections, the current
“decoupled” mechanism of trade sanctions has a number of considerable
virtues. It is also noteworthy that WTO members could always agree to compensate aggrieved exporters presently in settlement of cases, but do not elect to do so (putting aside the United States – Copyright Act case, which involved the payment of a modest sum to an entity that already existed for the purpose of distributing such monies among the relevant private actors). This “revealed preference” on the part of the WTO membership as a whole suggests further reason to question whether an exporter compensation system would be an improvement over the current state of affairs.

Of course, it is possible to design monetary compensation mechanisms that remain decoupled and do not compensate exporters. Some of the objections above would not apply to such a system, and indeed there are many moving parts in any monetary compensation mechanism that would have to be considered as noted in the introduction. But once any system to provide monetary “compensation” for harm caused by WTO violations was introduced, the political pressure for governments to turn those funds over to injured exporters might become considerable. This observation raises a further caution about any sort of remedial system grounded in monetary compensation.
References


WTO, United States – Continuing Dumping and Subsidy Offset Act of 2000, Recourse to Arbitration by the United States under Article 22.6 of the DSU.