Economic Structure of Renegotiation and Dispute Resolution in the World Trade Organization, The Rational Choice and International Law

Alan O. Sykes

Warren F. Schwartz

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

Part of the Law Commons

Recommended Citation


This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
THE ECONOMIC STRUCTURE OF RENEGOTIATION AND DISPUTE RESOLUTION IN THE WORLD TRADE ORGANIZATION

WARREN F. SCHWARTZ and ALAN O. SYKES*

ABSTRACT

The treaty creating the World Trade Organization (WTO) replaced the General Agreement on Tariffs and Trade (GATT) dispute resolution system, which contained no formal sanctions for breach of agreement as a practical matter, with a system that results in centrally authorized sanctions against recalcitrant violators of WTO trade agreements. We examine the important features of the new system and argue that the institutionalization of a sanctioning mechanism was not motivated by a perceived need to increase the penalty for violations, but rather by a need to decrease the penalty. In particular, the GATT system relied on unilateral retaliation and reputation to police the bargain. Toward its end, unilateral retaliation became excessive and interfered with opportunities for efficient breach. The WTO mechanism for arbitrating the magnitude of proposed sanctions is the major innovation under WTO law and ensures that sanctions are not set too high.

This paper is a contribution to the growing theoretical literature on the positive political economy of the World Trade Organization (WTO), which incorporates the General Agreement on Tariffs and Trade (GATT) and supplementary agreements on trade in goods, the General Agreement on Trade in Services (GATS), and the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS).¹ The focus of the paper is on the procedures

for dispute settlement set out in the Understanding on Rules and Procedures Governing the Settlement of Disputes (the Dispute Settlement Understanding, or DSU). Our goal is to develop an economic explanation for the structure of the rules and procedures of the DSU.

The point of departure is the proposition that the WTO agreements are, in effect, contracts among the political actors who negotiated and signed them. As with all contracts, it is in the interest of the signatories to maximize the joint gains from trade, that is, to enable the signatories to attain their Pareto frontier. Drawing on public choice theory, we further posit that the welfare of political officials is best measured by their political support—the factors that affect their ability to retain political power. A treaty on the Pareto frontier for political actors will then have the property that no alternative treaty can increase the political support for one signatory official without decreasing the political support for another. It is our thesis that the rules and procedures for renegotiation and dispute resolution in the WTO, which we set out in detail below, are explicable by this logic of joint political welfare maximization. This claim may seem obvious, but it is the details of the argument that make it interesting. In elaborating the argument, we draw considerably on the public choice literature and on the economic theory of contracts.

The analysis focuses on three central features of the WTO system that we believe have not been assigned sufficient importance or adequately explained by traditional international law scholars. The first can be found in the rules structuring the renegotiation and modification of WTO commitments. A prominent aspect of these provisions is that a member nation that wishes to deviate from its commitments may do so even if it is unable to secure permission from other nations by offers of compensatory trade concessions. If negotiations over compensation reach impasse, the nation wishing to deviate may proceed, and adversely affected nations may then withdraw “substantially equivalent” concessions in response. The second feature on which we focus involves the sanctions for breach of obligations. After a country is adjudged to be in violation of a WTO agreement, sanctions are limited to the withdrawal of substantially equivalent concessions previously granted to the country committing the violation by the country (or countries) harmed by the violation. More severe sanctions, which might at times be necessary if the violator is to be coerced into complying with its obligations, are not permitted. The third, related feature on which we focus concerns the mea-

---

4 Id.
5 We sketch these provisions here and document them in detail below.
measurement of substantially equivalent concessions. After a party has been found to be in violation of its obligations, it has a "reasonable time" to correct the problem. Only if it fails to do so within that time are sanctions allowed at all, and even then the sanctions are limited to measures substantially equivalent to the ongoing harm caused by the violation after the reasonable time for cure has elapsed. No sanctions are allowed for harm caused prior to that point in time.

We believe that these features can be understood using the economic theory of contract remedies. Economic theory teaches that a key objective of an enforcement system is to induce a party to comply with its obligations whenever compliance will yield greater benefits to the promisee than costs to the promisor, while allowing the promisor to depart from its obligations whenever the costs of compliance to the promisor exceed the benefits to the promisee. In the parlance of contract theory, the objective is to deter inefficient breaches but to encourage efficient ones.6 In the sections that follow, we will argue that the WTO provisions respecting renegotiation and the settlement of disputes over breach of obligations are carefully designed to facilitate efficient adjustments to unanticipated circumstances. We also conclude that formal sanctions in the WTO system are relatively unimportant to the other goal of contract remedies—the deterrence of inefficient breach.

Section I provides some pertinent background from the economic theory of contracts. Section II addresses renegotiation within the WTO system. Sections III and IV then consider the provisions for sanctioning nations discovered to be in violation of their commitments.

I. EFFICIENT ADJUSTMENT OF CONTRACTUAL COMMITMENTS: OF DAMAGES AND SPECIFIC PERFORMANCE, LIABILITY RULES AND PROPERTY RULES

Many contracts are negotiated under conditions of considerable complexity and uncertainty, and it is not economical for the parties to specify in advance how they ought to behave under every conceivable contingency. In such contracts, circumstances may arise in which it is in the joint interests of the parties for one of them to deviate from its commitments, or "breach" the contract.

There are essentially two mechanisms that parties to incomplete private contracts employ to encourage efficient performance of commitments while

6 See Robert Cooter & Thomas Ulen, Law and Economics 290 (1988); see also Robert L. Birmingham, Breach of Contract, Damage Measures, and Economic Efficiency, 24 Rutgers L. Rev. 273, 284–86 (1970) ("Repudiation of obligations should be encouraged where the promisor is able to profit from his default after placing his promisee in as good a position as he would have occupied had performance been rendered. . . . To penalize such adjustments through overcompensation of the innocent party is to discourage efficient reallocation of community resources. . . . Rigidity resulting from thus binding a party to his undertaking limits the factor and product mobility essential to proper functioning of the market mechanism.").
facilitating efficient breach of commitments. The first involves the award of expectation damages, which place the promisee in as good a position as it would have been in if the promisor had performed. Expectation damages thus deter inefficient breach because the promisor will not wish to violate and pay expectation damages unless the promisor gains more from the breach than the promisee loses, in which case breach is efficient.\footnote{See Richard A. Posner, Economic Analysis of the Law 117–26 (4th ed. 1992); John H. Barton, The Economic Basis of Damages for Breach of Contract, 1 J. Legal Stud. 277, 283–89 (1972); Birmingham, supra note 6, at 284–86; and Steven Shavell, Damage Measures for Breach of Contract, 11 Bell J. Econ. 466 (1980).} The weakness of this approach is that the measurement of damages by a court is costly, and errors in assessing damages may deter efficient breach if they are too high or permit inefficient breach if they are too low.\footnote{See Daniel Friedmann, The Efficient Breach Fallacy, 18 J. Legal Stud. 1, 6–7 (1989); see generally Alan Schwartz, The Case for Specific Performance, 89 Yale L. J. 271 (1979); Thomas S. Ulen, The Efficiency of Specific Performance: Toward a Unified Theory of Contract Remedies, 83 Mich. L. Rev. 341 (1984).}

The mechanism that employs expectation damages as the means for inducing performance when it is efficient, and breach when it is not, is known as a "liability rule." A party who wishes to deviate from its commitments may do so without the need to secure the permission of any adversely affected party but is liable for damages as a result.\footnote{See Guido Calabresi & A. Douglas Melamed, Property Rules, Liability Rules and Inalienability: One View of the Cathedral, 85 Harv. L. Rev. 1089, 1092–93 (1972); see also Ian Ayres & Eric Talley, Solomonic Bargaining: Dividing a Legal Entitlement to Facilitate Coasian Trade, 104 Yale L. J. 1027, 1036–72 (1995); and Louis Kaplow & Steven Shavell, Property Rules versus Liability Rules: An Economic Analysis, 109 Harv. L. Rev. 713, 715 (1996).}

The second mechanism that can encourage efficient performance while allowing efficient breach involves renegotiation, motivated by an order for specific performance. In this case, the promisor is directed to perform, and a failure to do so will be punished so severely that a party will never prefer violating the order directing performance to complying with its obligation. But the promisor can still avoid its commitments by securing permission from the promisee, usually by paying for it. Since the promisor will pay no more than the value to it of the breach, and the promisee will accept no less than the value of the harm it will suffer from the breach, an agreement that permits the promisor to breach can be reached only when the benefit to the promisor of the breach exceeds the harm to the promisee resulting from the breach, that is, when breach is efficient.

A mechanism whereby the promisor must secure the permission of the promisee before deviating from its obligations is a form of property rule.\footnote{See Calabresi & Melamed, supra note 9, at 1092–93; Kaplow & Shavell, supra note 9, at 715.} The term comes from the analogy to tangible property, which ordinarily cannot legally be taken from one private party by another unless the latter party grants permission, usually by selling it. A property rule here avoids...
the difficulties associated with having a court compute expectation damages. But it introduces other costs—those associated with the transaction costs of bargaining between the promisor and the promisee over the possibility of modifying obligations, including those attributable to strategic behavior during the bargaining process. When these costs are sufficiently low in relation to the judicial and error costs of expectation damages, however, the property rule mechanism will be preferable. It is also likely to be preferable if deviation from obligations is always inefficient.

This background will be quite helpful in our discussion of the WTO system. As we explain below, the system consistently employs liability rules rather than property rules to protect WTO commitments against breach. It does so, we believe, out of concern that an alternative approach would make it too difficult for WTO members to modify their commitments efficiently. Further, the protection of the liability rule is limited in unusual ways, much more so than in the context of private contracts. This feature, we believe, is attributable to the fact that formal sanctions are a relatively unimportant factor in inducing member nations to live up to their commitments. The following sections elaborate these claims.

II. Renegotiation and Modification of Concessions in the World Trade Organization System

Public choice teaches that the objectives that individual countries pursue through international agreements are determined by an interaction among organized interest groups. While this process is not fully understood and assuredly varies across nations, there is wide agreement that producer interests will exercise disproportionately greater influence than will consumer interests, at least in the democracies that dominate the developed world (and thus the trading community). Hence, multilateral agreements that reduce trade barriers are not driven primarily by a desire to benefit consumers (despite the fact that they do) but by a desire to benefit certain producer groups. Trade concessions by one nation are made in exchange for reciprocal trade concessions by other nations that will afford exporters greater access to foreign markets. Exporters will reward their political officials for securing these concessions. Where these political rewards exceed the political costs associated with reduced protection for import-competitive domestic industries,

---


political officials will benefit on balance and conclude agreements for mutual reduction of trade impediments.¹³

But the parties to trade agreements, like the parties to private contracts, enter the bargain under conditions of uncertainty. Economic conditions may change, the strength of interest group organization may change, and so on. Accordingly, officials cannot be certain that the bargain they strike will benefit them in all of its details. Likewise, even where the bargain on a particular issue is initially beneficial, changing circumstances may make it politically unappealing. For these reasons, the drafters of trade agreements may be expected to include devices for adjusting the bargain when it proves mutually disadvantageous.

As noted above, the performance of contractual obligations becomes inefficient when the benefits to the promisee(s) of performance are less than the costs to the promisor of performance. Joint gains then arise if the promisor does not perform, gains that can be distributed ex ante or ex post in any manner that the parties prefer. Although this theory of efficient performance and nonperformance has been developed with reference to private contracts, where the costs and benefits of performance may be measured in money, it applies equally to other bargains such as trade agreements. And the theory of public choice suggests that the metric of welfare for each signatory to a trade agreement will not be money, but instead will be the political welfare (votes, campaign contributions, or graft, as the case may be) of its political officials.¹⁴ Any Pareto optimal trade agreement must maximize a weighted sum of this welfare measure for each signatory government.¹⁵ Implicit, then, in any optimal agreement is a set of weights (called “shadow prices” by economists) that allow the political welfare of one government to be traded off against the welfare of another.

The welfare weights implicit in any Pareto optimal agreement also serve to identify the conditions under which the nonperformance of obligations is efficient. When the political burden of performance to a promisor exceeds the political detriment of nonperformance to the promisee(s), evaluated at the proper weight or shadow price, nonperformance is jointly desirable. Roughly speaking, the political costs of performance may be said to exceed the benefits of performance, just as benefits may exceed costs in the case of a private contract where both are measured in money. It is in the interests


¹⁴ See Mueller, supra note 11, at 2–3.

following our earlier discussion, the parties to any kind of contract can facilitate efficient adjustment of obligations in three ways. First, they can specify in the contract itself the conditions under which performance will not be required or the price for a party to buy out of a particular obligation—force majeure clauses and liquidated damages clauses in private contracts are examples of this approach. Second, when their contract is incomplete as to certain contingencies that may arise, they can agree on (or embrace a legal system that provides them with) a liability rule that encourages efficient nonperformance. As discussed above, the familiar rule of expectation damages in contract law is such a rule. Third, the parties can embrace a property rule and simply renegotiate when performance becomes inefficient. The promisor can buy its way out of the obligation to perform by paying the promisee(s) an amount that makes it whole and still leaves the promisor better off than with performance of the original obligation.

The provisions of the WTO agreements pertaining to renegotiation exhibit aspects of the first two approaches but stop short of creating a property rule. Consider first the Article XIX escape clause, which authorizes temporary measures that would otherwise violate WTO commitments for the protection of industries that are experiencing severe dislocation due to increased import competition. Such industries are likely to have rates of return well below the competitive level and, as a result, to be losing quasi rents on fixed investments. On average, they will lobby more vigorously for protection than other industries because the benefits of protection are less dissipated (if at all) by new entry; to the extent that protection merely raises the rate of return toward the competitive level, no new domestic competitors will be induced to enter the industry. Industries that are profitable and growing are likely to have returns above the competitive level in many cases, which will eventually be dissipated by entry regardless of government policy at home or abroad. Hence, they have less incentive to lobby for domestic protection and less incentive to punish their political leaders for failing to maintain access to foreign markets at historical levels. Accordingly, it will be politically efficient, from the perspective of parties to trade agreements, to afford tran-

16 A formal model that develops these results may be found in Sykes, supra note 12, in the appendix.

17 See GATT 1994, supra note 1, art. XIX(1)(a). The provision reads in full: "If, as a result of unforeseen developments and of the effect of the obligations incurred by a contracting party under this Agreement, including tariff concessions, any product is being imported into the territory of that contracting party in such increased quantities and under such conditions as to cause or threaten serious injury to domestic producers in that territory of like or directly competitive products, the contracting party shall be free, in respect of such product, and to the extent and for such time as may be necessary to prevent or remedy such injury, to suspend the obligation in whole or in part or to withdraw or modify the concession."

18 See Sykes, supra note 12, at 274.
sitory protection, at the expense of growing and prosperous foreign competitors, to import-competing industries that suffer severe dislocation. The escape clause permits such measures and may thus be viewed as an example of the first option above for facilitating efficient nonperformance, a provision written into the contract that excuses performance under specified contingencies.

To be sure, the concern arises that a nation may abuse its right to use the escape clause, imposing protection when it creates more political detriment abroad than can be justified by the benefits it creates at home. A compensation requirement can help to deter such inefficient behavior and was included in the escape clause system until the Uruguay Round. The new, partial exemption from the compensation requirement for the first 3 years of an escape clause measure suggests a judgment by the WTO membership that oversight by the strengthened dispute resolution process can adequately police abuse of such measures and that a compensation requirement is no longer essential to keep the member nations "honest."

A more comprehensive provision for adjustment of the bargain is Article XXVIII of GATT 1994 (and its GATS counterpart, Article XXI). Unlike Article XIX, Article XXVIII does not set out specific contingencies under which deviation from obligations is permissible but instead establishes a procedure under which, subject to certain constraints, any tariff concession can be withdrawn for any reason for an indefinite period of time. It requires as part of this process that nations seeking to withdraw concessions offer compensatory concessions to affected trading partners. But it is noteworthy that Article XXVIII does not require the member who is withdrawing a concession to secure the permission of affected trading partners—it does not create a property rule. Instead, although members are asked to negotiate mutually satisfactory compensation with other members if possible, Article XXVIII provides that a member may proceed to withdraw concessions in cases where negotiations over compensation break down and further provides that adversely affected trading partners may at that point unilaterally withdraw

19 See GATT 1994, supra note 1, art. XIX(3)(a) ("If agreement among the interested contracting parties with respect to the action is not reached, the contracting party which proposes to take or continue the action shall, nevertheless, be free to do so, and if such action is taken or continued, the affected contracting parties shall then be free . . . to suspend . . . such substantially equivalent concessions or other obligations under this Agreement the suspension of which the [GATT membership as a whole does] not disapprove.").


22 See GATT 1994, supra note 1, art. XXVIII(2) ("In such negotiations and agreement, which may include provision for compensatory adjustment with respect to other products, the contracting parties concerned shall endeavour to maintain a general level of reciprocal and mutually advantageous concessions not less favourable to trade than that provided for in this Agreement prior to such negotiations.").
substantially equivalent concessions or other obligations. Ultimately, then, concessions are protected by a liability rule. And the magnitude of “liability” is clearly specified—concessions substantially equivalent to those withdrawn by the member that proceeds under Article XXVIII.

We believe that the explanation for these provisions lies in the desire of signatories to facilitate efficient breach and in the relative superiority of a liability rule approach to that task. At first blush, this claim may seem surprising because the harm done to political officials by a breach of promise in the WTO is no doubt difficult to measure precisely, and when damages are hard to calculate, that fact is usually thought to be a heavy thumb on the scale favoring a property rule over a liability rule. But there is a countervailing consideration here that is compelling. Under the most-favored-nation principle of the WTO, trade concessions must extend equally to all WTO members (WTO membership includes 144 countries at this writing).

Hence, under a property rule, a nation seeking to depart from a prior concession would have to secure the permission of potentially dozens of other nations. It would then face an acute holdout problem as each of the many promisees tried to capture as much as possible of the gain that the promisor could realize from avoiding the concession. Such strategic behavior might prevent agreement from being reached at all, or at least delay it uneconomically while negotiation and posturing dragged along. The liability rule approach of Article XXVIII averts this problem.

Further, by limiting the retaliatory withdrawal of concessions to those substantially equivalent, the system seeks to ensure that the price for non-performance under the liability rule is not too high. Although the phrasing is somewhat vague, a withdrawal of substantially equivalent concessions may be understood as allowing members adversely affected by a withdrawal of concessions under Article XXVIII to raise their level of political welfare by reimposing protection for the benefit of domestic constituencies that will reward them for it, but only up to the point that their level of political welfare is restored to its original level. Indeed, during discussions on Article XXVIII in the Tariff Agreements Committee in 1947, the proposal to include a provision for compensatory withdrawal was explained as follows: “[I]f we wish to take an item out of our Schedule then clearly it is fair and proper that the countries with whom we negotiate should be free to make the corresponding changes in their Schedules in order to restore the balance . . . but we want any such exercise to be limited to what is corresponding and not to be used

23 Id. at art. XXVII(3).

24 Id. at art. I:1 (requiring “that any privilege, advantage, or benefit granted to imports from one [WTO] member be extended to imports of similar products from all other [WTO] members.”); see also GATS, supra note 22, art. II:1 (including nearly identical most favored nation status to members).

in a punitive way."  

In other words, political expectations under the bargain are protected by a rough equivalent of expectation damages, but nations are disabled from insisting on more because any greater level of retaliatory withdrawal would raise the price of nonperformance above the costs to the disadvantaged promisees and thus discourage efficient nonperformance.

III. THE LIABILITY RULE REMEDY FOR VIOLATION OF WORLD TRADE ORGANIZATION OBLIGATIONS

The most intriguing use of a liability rule in the WTO system is pursuant to the DSU, which governs claims by one member nation that another has violated its obligations. Article 21(3) of the DSU provides that a member has a reasonable period of time to bring its policies into conformity with its obligations after it has been found to have violated them.\(^\text{26}\) Article 22(1) then states that compensation or a suspension of concessions may result if compliance has not been achieved within a reasonable period of time.\(^\text{27}\) The first step in the process is a negotiation over compensation, in effect to determine whether the case can be "settled."\(^\text{28}\) Should those negotiations fail, the aggrieved party (or parties) can propose a suspension of concessions, which must be substantially equivalent to the ongoing harm that they suffer from the violation.\(^\text{29}\) An arbitration procedure exists to examine the substantial equivalence question if the member faced with such a suspension of concessions objects that the suspension is excessive.\(^\text{30}\)

Plainly, as with Article XXVIII discussed above, this system is best seen as one embracing a liability rule rather than a property rule. A party found to be in violation of its obligations can, if it so chooses, continue to violate them. The ultimate price to be paid, if the case is not settled, is the withdrawal of substantially equivalent concessions. This structure must, we submit, reflect a collective judgment that a property rule (for example, a threat to expel the recalcitrant violator if it does not cease and desist) would be inferior. The reasons why relate to the considerations discussed above—the large transaction costs and opportunities for strategic behavior that would arise if

---

\(^{26}\) UN Doc. TAC/PV/14 at 20.

\(^{27}\) See DSU, supra note 2, at art. 21(3) ("If it is impractical to comply immediately with the recommendations and rulings, the Member concerned shall have a reasonable period of time in which to do so.").

\(^{28}\) Id. at art. 22(1) ("Compensation and the suspension of concessions or other obligations are temporary measures available in the event that the recommendations and rulings are not implemented within a reasonable period of time.").

\(^{29}\) Id. at art. 22(2) ("Such Member shall, if so requested, and no later than the expiry of the reasonable period of time, enter into negotiations with any party having invoked the dispute settlement procedures, with a view to developing mutually acceptable compensation.").

\(^{30}\) Id. at art. 22(4) ("The level of the suspension of concessions or other obligations authorized by the DSB shall be equivalent to the level of nullification or impairment.").

\(^{31}\) Id. at art. 22(6–7).
DISPUTE RESOLUTION

a member trying to adjust its obligations had to secure the permission of all of the affected members.

Recent WTO decisions make clear that our interpretation of WTO law is correct, even if they do not clearly acknowledge the liability rule nature of the system. In the "bananas" dispute between the United States and the European Union (EU), the EU declined to comply with a panel ruling because it found that its tariff preferences for bananas from certain nations violated WTO law. The United States then invoked its retaliation rights and proposed substantial sanctions that the EU challenged before an arbitration panel as excessive. In defending its proposed sanctions, the United States argued that its "suspension (of trade concessions) is an incentive for prompt compliance. . . . [P]recision in measuring trade damage is not required." The United States thus suggested, in effect, that the purpose of the sanction was to enforce a property rule and that careful calibration of sanctions was unnecessary. The arbitrators rejected this position: "We agree with the United States . . . that it is the purpose of countermeasures to induce compliance. But, this purpose does not mean that the DSB [Dispute Settlement Body] should grant authorization to suspend concessions beyond what is equivalent to the level of nullification or impairment. In our view there is nothing in [the relevant provisions of the DSU] that could be read as a justification for countermeasures of a punitive nature." By refusing to permit the imposition of "punitive" sanctions, the arbitrators impliedly acknowledged that the sanction is more in the nature of compensation than punishment. They set a price for the EU's persistence in its violation of WTO law equal to the harm caused to its trading partners. The system thus allows violations to persist as long as the violator is willing to pay that price, which is the essence of a liability rule approach.

We note that our conclusion is somewhat at odds with the views of other scholars in the field. John Jackson recently addressed the question of whether a WTO member nation that had been found to be in violation of its commitments and that refused to bring its behavior into compliance should be deemed to be in violation of international law. He concludes that a refusal to comply with WTO treaty obligations is indeed a violation of international law, even if compensation is agreed upon or if a retaliatory suspension of concessions is in place. In so concluding, he in effect further concludes that a member of the WTO is obligated to comply with its obligations in all circumstances.

32 Regime for the Importation, Sale, and Distribution of Bananas—Recourse to Arbitration by the European Communities under Article 22.6 of the DSU—Decision by the Arbitrators, WT/DS27/ARB, art. 6.1 (April 9, 1999).
33 Id. at art. 6.3.
35 Id.
Jackson does not base his conclusion on policy or on any articulation of why he believes that strict compliance with all obligations at all points in time should be the preferred outcome for the WTO membership. Rather, he cites 11 textual provisions of the WTO in support of his position. The two that most powerfully support his argument are the following:

Compensation and the suspension of concessions or other obligations are temporary measures available in the event that the recommendations and rulings are not implemented within a reasonable period of time. However, neither compensation nor the suspension of concessions or other obligations is preferred to full implementation of a recommendation to bring a measure into conformity with the covered agreements.\(^{36}\)

The suspension of concession or other obligations shall be temporary. . . . [T]he DSB shall continue to keep under surveillance the implementation of adopted recommendations or rulings . . . [while] the recommendations to bring a measure into conformity with the covered agreements have not been implemented.\(^{37}\)

We acknowledge that the provisions of the WTO relied on by Jackson provide reasonable support for his conclusion that WTO members are obliged to comply with dispute resolution decisions that go against them. Nevertheless, we disagree with that proposition, both as a matter of textual interpretation and for policy reasons implicit in our discussion to this point.

Our arguments from the text are straightforward. The statement in the first passage that compliance is "preferred" is weak—it does not say that compliance is mandatory, and it seems to us that this provision does not exclude the possibility that noncompliance may in some cases be acceptable. The ongoing surveillance discussed in the second passage indeed hints at an obligation to comply, but there is certainly another interpretation. Because circumstances change and the proper calibration of the substantially equivalent concessions may change as well, it is perhaps not surprising that the DSB should exercise some continuing oversight in these cases much as a conventional court might retain jurisdiction over a case where damages are payable over time (such as child support payments under family law or medical monitoring costs in tort). Likewise, ongoing violations may have an impact on parties other than the original disputants. Continued publicity and oversight may thus serve to alert other members who might suffer redressable harm. Finally, and related perhaps to the third-party effects just mentioned, we do not dispute that a "preference" for compliance seems implicit in the system. Ongoing oversight thus serves to check periodically on whether the impasse that led to compensation or retaliation may have lifted. In effect, the violating country is required to persuade the international community that persisting in the violation is desirable. Hence, the existence of continued

\(^{36}\) See DSU, supra note 2, at art. 22(1).

\(^{37}\) Id. at art. 22(8).
oversight by no means excludes the possibility that members have the legal right to opt for paying damages in the form of a loss of trade concessions from other parties.

Our final argument from the overall structure of the text is even more straightforward. We simply note that the provisions of the DSU, taken as a whole, allow a violator to continue a violation in perpetuity, as long as it compensates or is willing to bear the costs of the retaliatory suspension of concessions. If WTO members really wanted to make compliance with dispute resolution findings mandatory, they would have imposed some greater penalty for noncompliance to induce it.

Turning to policy, the starting point is the observation that the textual provisions cited by Jackson both begin by asserting that the withdrawal of concessions is to be viewed as a "temporary" measure. It would seem then that, at least for some temporary period of time, a violation coupled with the withdrawal of concessions is acknowledged to be potentially superior to immediate compliance. Indeed, the fact that violators are given a reasonable period of time to conform their policies before sanctions or compensation become possible further supports the proposition that some period of deviation is seen as potentially valuable.

The reason why is not difficult to divine. World Trade Organization violations are typically the result of domestic laws and regulations enacted by the violating country. Thus, curing the violation requires a new law or regulation that repeals the one that constitutes the violation. For a number of reasons it may be politically difficult, conceivably impossible, to enact such a change. The legislative and regulatory processes are, of course, elaborate and costly. Proposed changes must compete for a place on the agenda. Interest groups who gain from the violation will oppose repeal and be able to exploit differences among supporters of repeal as to what compensating benefits, if any, should be granted to the industries who will lose the benefits of the law.

If these factors make some delay in compliance inevitable, as the system apparently acknowledges and tolerates, there is no reason to think that they may not at times make compliance politically infeasible for an extended period of time. And rather than expel the member who faces such political difficulty or impose some other draconian penalty, the system instead acknowledges that the joint interests of the parties may be better served by compensation or retaliation that restores the benefits of the bargain to aggrieved parties while allowing officials in the violator nation to continue doing what must be done out of political necessity.

Indeed, if one is to claim that the purposes of the WTO members would be better served by compliance in all circumstances, it seems that one must believe that at the time the WTO rules were devised, the drafters were able to anticipate every situation in which the costs of compliance would exceed the benefits of compliance and include provisions to excuse compliance in all of these circumstances. In the parlance of contract theory, the parties
would have had to be able to write a complete contract expressly specifying what would be required in all circumstances that might arise. We think it plainly unrealistic to think that the many parties to the WTO agreement, covering as it does matters of great complexity, could have done so successfully. Knowing that, they framed a dispute resolution system designed to facilitate efficient breach by using a sensible liability rule for that purpose.

But there is one possible response that warrants attention. It might be argued that strict compliance with the rulings of the dispute resolution process is desirable and that adjustments to unanticipated circumstances should always be made via the renegotiation process of Article XXVIII. The drafters did not imagine that they could write a complete contract, the argument runs, but they wanted all changes to occur through tariff renegotiations.

One difficulty with this argument is that it presupposes that changes in most-favored-nation tariff rates can adequately address the political difficulties that arise from unanticipated circumstances. It seems unlikely that this will be true. The recent beef hormones case\(^\text{38}\) is a good illustration. The EU was held to have violated its obligations under the WTO Sanitary and Phytosanitary Measures Code by prohibiting imports of hormone-raised beef, ostensibly because of health concerns. If the continuation of the ban were nevertheless a political necessity for European officials, a uniform change in the tariff rates applicable to all beef from all sources, hormone raised or not, could not replicate its effects. Here, deviation from a nontariff commitment would seem necessary and renegotiation of tariffs a politically unsatisfactory substitute.

In short, it seems clear to us that the WTO system contemplates departures from specified obligations when the costs of compliance exceed the associated benefits, whether those obligations are tariffs or nontariff issues. We can see no other purpose to the provisions that allow departure from obligations when agreement is not reached and confer on the promisee only the right to withdraw substantially equivalent concessions. Such a provision can only represent an institutional means for setting an appropriate price for violating commitments when the price cannot be determined through negotiations.

IV. THE LIMITED SCOPE OF SANCTIONS IN THE WORLD TRADE ORGANIZATION SYSTEM

We have focused thus far on the role of a liability rule in the WTO system in facilitating efficient deviations from commitments following a change in circumstances. We now consider the second role of a liability rule—to deter violations when the benefits of compliance are greater than the costs of compliance.

A. The Absence of Sanctions prior to the World Trade Organization

What is remarkable about the WTO/GATT system is how unimportant formal sanctions have been in encouraging compliance with trade commitments throughout its history. As noted, the WTO succeeded the GATT, which began in 1947. Until 1995, when the WTO agreements superseded the GATT, it was effectively impossible for a nation that was found to have violated the GATT to become subject to formal sanctions. The reason was the consensus rule, which held that any nation could block the authority for the imposition of sanctions, including the nation that had violated the GATT and was threatened with them. Indeed, until 1989, a potential disputant could even block the formation of a dispute resolution panel to hear the merits of a complaint. As a result, GATT dispute resolution was limited to a system that would often (but not always) hear the merits of a complaint and render a decision about the existence of a violation but would never proceed to the point of imposing penalties when a violation was found.

Nevertheless, the GATT system held together rather well. Tariffs in the developed world fell from an average of nearly 50 percent in 1947 to an average of about 5 percent by the end of the GATT. To be sure, some cheating on obligations occurred, but the level of cheating was modest. We are unaware, for example, of any allegation in the history of the system that a nation flagrantly refused to comply with one of its tariff commitments by raising a tariff rate above an agreed tariff limit. Further, where cheating might be said to exist by some, it was often an efficient, tacit amendment of the

39 See sources cited in note 1 supra.
40 See John H. Jackson, The World Trading System: Law and Policy of International Economic Relations 117 (1999) ("[A] disputing nation could block adoption of a report and then argue that no binding requirement exists for it to follow the report."); see also John H. Jackson, William J. Dovey, & Alan O. Sykes, Legal Problems of International Economic Relations: Cases, Materials, and Text 830 (3d ed. 1995) ("Prior to the Uruguay Round, the GATT dispute resolution process could be 'blocked' by one of the disputants under the 'consensus rule'—the losing party to a dispute had to agree to accept the outcome before any formal action could be taken to authorize sanctions.").
42 See Jackson, supra note 40, at 116 ("Although the Contracting Parties were authorized (by majority vote) to suspend concessions (by way of retaliation, retaliation, or 'balancing of benefits'—a term which is not and never has been clear), they actually did so in only one case. That instance was the result of a complaint brought by The Netherlands against the United States for the latter's use, contrary to GATT, of import restraints on imported dairy products from The Netherlands. For seven years in a row, The Netherlands was authorized to utilize restraints against importation of U.S. grain, although it never acted on that authorization. This had no effect on U.S. action, however.").
bargain.\textsuperscript{44} When such tacit modifications are put to one side, the incidence of flagrant cheating under the GATT system was indeed quite low.\textsuperscript{45}

Our explanation for this state of affairs emphasizes that there are strong forces that induce countries to comply with their obligations, although no costs would be formally imposed on them by the GATT if they deviated. Three considerations explain why the system worked as well as it did: the domestic costs of violations, reputational sanctions for noncompliance, and unilateral retaliation against violators.

1. Domestic Costs of Violations

It will often be true that domestic political considerations encourage a country to comply with its commitments under trade agreements. This is true for two sets of reasons.

The first relates to the way the balance of political forces that favor trade protection and trade liberalization will change following the advent of a market-opening trade agreement. As a preliminary, protectionism induces inefficient investments in the domestic production of certain goods and services by importing nations. Those investments commonly entail sunk costs in the form of physical capital that cannot readily transfer to other uses and specific human capital with the same property. The owners of these sunk investments will lose quasi rents on them if protection is removed and will thus devote resources to the political process to protect those rents.\textsuperscript{46} These efforts by import-competing firms and workers may prove insufficient to prevent the lowering of trade barriers, however, because the exporters who benefit from reciprocal trade liberalization may be willing to pay more to their officials to secure access to foreign markets than import-competing interests will pay to keep their market protected. If so, a trade agreement will be struck.

Following the trade agreement, the rate of return to firms and workers in the formerly protected industries will tend to fall below a competitive level because of the introduction of more efficient foreign competitors, and they will begin to exit. Concomitantly, their sunk investments decay over time as physical capital depreciates and specific human capital is replaced by the workers who retrain to work in other industries. The pressure from the owners of sunk investments for the reimposition of protection should fall steadily

\textsuperscript{44} For example, the general failure of parties to demonstrate under the Escape Clause that a surge in import competition was attributable to a particular trade concession and that it was "unforeseen," as required by Article XIX, was really a tacit amendment to the original GATT that all parties accepted and that was later incorporated into the Safeguards Agreement. See Sykes, supra note 12, at 287–88.


\textsuperscript{46} See Alan O. Sykes, The Economics of "Injury" in Antidumping and Countervailing Duty Cases, 16 Int'l Rev. L. & Econ. 5, 24 (1996).
as these specific investments decay. In the limiting case, no specific investments remain, and renewed protection would simply necessitate new investment in an industry that would earn no more than a competitive rate of return. No one would benefit from such protection given the opportunity to earn competitive returns elsewhere, and the pressure for renewed protection would drop to zero. In more realistic cases, some pressure for protection will remain from those whose human or physical capital would earn at least transitory rents from protection, but in most cases the magnitude of that pressure should still be considerably lower than it was before trade was liberalized in the first instance, when sunk investments in the protected industry were much more important.

Reciprocally, political pressure to resist renewed protection may grow with time. In particular, where the imports in question are utilized by producing industries, which tend to be better organized than ordinary consumers, a constituency may develop whose returns on their own fixed investments would be impaired by a significant increase in the price of imports.

In sum, the political balance of interests favoring and opposing the results of a trade agreement may be expected to tilt quite systematically toward those favoring the agreement as time passes. A fortiori, the political pressure to comply with market-opening commitments in trade agreements will tend to strengthen over time, and pressures to deviate from many commitments may simply disappear. We conjecture that this phenomenon also has much to say about the reasons why the WTO system has slowly ratcheted down protection through a series of rounds over the years—each round goes as far as it can given the resistance from import-competing industries, but as that resistance decays after each round is completed, new opportunities for politically profitable deals emerge with time. It may also explain why concessions often have to be implemented only gradually over time.

A second domestic political reason why nations may be inclined to comply with their trade commitments relates to the fact that it may be more costly for interest groups to seek protection than to resist its abolition. In the United States, for example, trade agreements are followed by implementing legislation that conforms federal law (including tariff rates) to the new agreement. A constituency favoring renewed protection, then, must incur the costs of changing a federal statute. Prior to the agreement, by contrast, those favoring trade liberalization must incur the costs of changing the federal statute (as well as of encouraging the international negotiation). It is likely easier to defend an existing statute than to change it for a variety of reasons. Time on the legislative agenda is scarce. Also, individual members of Congress (such as key committee chairs) may have the effective power to veto change, yet they will lack the power by themselves to effect change. Thus, parties resisting change may need fewer political figures to support them than parties seeking change.

As a result, once a trade-liberalizing agreement is reached and imple-
mented, the balance of political power may shift importantly and immediately against those who were previously the beneficiaries of protection. When this fact is coupled with the fact that their sunk investments will begin to decay and their returns to renewed protection will accordingly diminish with time, it seems plausible to us that the constituency for renewed protection will often lack the political muscle to secure it, even if renewed protection would not result in any international sanction of consequence.

We do not suggest, however, that a renewal of protection would always be unilaterally unattractive. After all, the fact that protection had once been afforded indicates that the political forces benefiting from it were powerful enough to secure it. It is certainly unrealistic to suppose that every time a trade agreement is struck, the forces supporting renewed protection lose so much relative influence that they could not effectively prevail on their governments to restore it if there were no international penalty to be paid. We thus turn to other factors that help encourage adherence to trade commitments.

2. Reputation

Nations that renege on their commitments may be expected to face some reputational cost in the form of having to deal with other nations on less favorable terms in the future. This cost will be borne not only in future dealings with the nation aggrieved by a breach of promise, but also in dealings with all other nations that are aware of the breach.

The skeptic might question whether such reputational penalties will be of much importance in the trading system, however, because their costs might seem to be widely diffused. If the United States reneges on a WTO obligation, for example, such behavior might be expected to diminish the opportunity for the United States to strike favorable trade deals in future negotiating sessions. Yet the domestic beneficiaries of those deals might be hard to predict, and the benefits lost by any single organized group might be small in present value. Accordingly, therefore, one might conjecture that no domestic interest group would worry much about the consequences of reneging for the nation's reputation and hence that a fear of reputational damage to the nation will not much constrain domestic political officials who otherwise find it in their interest to renege on promises.

This skeptical view is wrong for three reasons. First, in the WTO/GATT system, negotiations are ongoing more or less in perpetuity. Of late, for example, negotiations over commitments in various service sectors remain very much on the table, as do a number of other topics. In many instances, therefore, the loss of an ability to make credible promises will immediately come back to haunt negotiators representing specific and well-organized groups with a current stake in negotiating progress.

Second, it would be a mistake to suppose that reputation is cabined to the trade area. Nations are engaged in a never-ending series of diplomatic ini-
tiatives on matters ranging from trade to national security to human rights and so on. Typically, the entity that represents a nation in trade negotiations (in the United States, the executive branch) must also negotiate on these other matters. It is quite likely, therefore, that a linkage exists between credibility in the trade-negotiating field and credibility in other matters of current concern.

Third, even if reputational costs in the form of forgone opportunities for future trade liberalization would be borne by a fairly diffuse group of exporters, it does not follow that they will be ineffective at organizing today to protect themselves. Exporting interests can form associations with the mission of overcoming such collective-action difficulties. Entities like the American Chamber of Commerce, for example, can and do serve that function. Likewise, exporters can install political agents who have their interests at heart (either because of the agents' ideology or because of their employment prospects on leaving government). These agents will then act to protect their reputations as effective negotiators for export interests (the U.S. trade representative is illustrative) by actively opposing any proposals to renege on past bargains.

One difficulty with reputational penalties, of course, is that they depend on the quality of information in the trading community about the behavior of violators. Such information may be particularly imperfect as to violations by trading partners that affect third countries; that is, if the United States reneges on an obligation to Brazil, will the EU find out about it and take it into account with respect to future dealings with the United States?

This concern highlights the value of a central dispute authority to hear the merits of complaints, even if that authority has no power to authorize sanctions. By serving as a vehicle for transmitting information about violations throughout the trading system, central dispute resolution enhances the reputational costs of cheating. We think that under the consensus-based system of the old GATT, this was the primary function of the dispute system. Further, the fact that a disputant could block the formation of a panel to hear the merits until 1989 did not destroy the efficacy of the system, for in most cases, the refusal of a disputant to allow the formation of a dispute panel would suffice for an adverse inference by other nations.

If this last claim is correct, however, why did the system evolve in 1989 into one in which disputants could no longer block the formation of dispute panels to hear the merits? Two answers may be given.

Cases may have existed in which blocking did not suffice for an adverse inference. Nations might have claimed, for example, that their decision to block a panel rested on the costliness of the panel process or on some fear that the panel process could not be expected to resolve the case correctly. It may have been difficult for other nations to tell when such assertions were disingenuous, and if this is so, the decision by the members to afford panels
automatically may have indeed been taken to enhance the reputational penalties for violators.

Alternatively, dispute rulings may provide public goods to an extent. Contract theory teaches that parties leave gaps in contracts or use ambiguous language because it is too costly to anticipate what behavior will be in their joint interests or to express precisely what behavior is optimal under every imaginable contingency. When contractual incompleteness leads to disputes, a third-party enforcer can generate joint gains for the parties by resolving the dispute in the way that maximizes joint welfare, that is, in the way that corresponds to what the parties would have chosen for themselves in the absence of transaction costs. Of course, gap filling by a third-party enforcer is not helpful if the resultant gap fillers are chosen improperly. But the agreements that constituted the GATT system were exceedingly elaborate and ultimately grew to hundreds of pages in length, replete with various principles on which panels could draw to guide their analysis (the new WTO agreements are even more elaborate, so the point carries forward). It is thus plausible that panel-generated gap fillers under the WTO/GATT system serve the joint interests of the members in a fairly high percentage of cases.

A party to a dispute might nevertheless have blocked the formation of a panel if it expected that it would probably lose. Even if an adverse inference would then have been drawn by others about the behavior of that party, there may still have been some virtue in ensuring that the panel process could go forward anyway to help clarify the terms of the bargain for everybody. This may have become increasingly true with time, as the GATT system came to encompass more and more side agreements covering many topics. The decision in 1989 to allow complaining parties to obtain a panel decision as a matter of right, therefore, may have reflected increasing benefits in the use of the panel process to generate gap fillers for the membership owing to the increasing complexity of the system.

In summary, we believe that reputation played and continues to play an important role in enforcing the rules of the trading system. But we do not wish to overstate the case. Most of the time, reputation undoubtedly functions as an imperfect check on opportunism here as elsewhere. Thus, the system also has room for sanctions beyond simply reputational costs.

3. Unilateral Sanctions

When a nation breaches a trade commitment and the harm done is material and noticeable to the foreign exporters that benefit from the promise, those interest groups may be expected to complain to their political representatives and to reward those officials for taking action to correct the problem. Re-

gardless of the nature of third-party dispute resolution at the international level, therefore, nations will have an incentive to punish breach of promise by other nations. Furthermore, to the extent that sanctions will take the form of protectionist measures that benefit domestic industries in the sanctioning country, they may actually generate political rewards from those industries for the officials that impose them, further adding to the incentive to employ them.

The prospect of unilateral sanctions is not merely hypothetical. In the United States, Section 301 of the Trade Act of 1974 has long authorized the executive branch to retaliate for breach of trade agreements by other nations.\textsuperscript{48} The EU has a similar statute on the books, and both statutes have been used.\textsuperscript{49} In other nations, statutory enactments that authorize retaliation may be absent, but the inherent powers of political authorities to take action often make formal authority unnecessary.

The growth of unilateral retaliation was a natural response to the consensus rule that prevented centralized sanctions. Indeed, it has been argued elsewhere that Section 301 and its various formal and informal international counterparts (the European Community enacted a mirror statute) were valuable tools for holding the trading system together by affording useful self-help strategies in the face of the limitations in the GATT dispute settlement system.\textsuperscript{50}

The skeptic might again respond, however, that the prospect of unilateral retaliation may be a relatively weak deterrent to violations. The effects of retaliation, the argument might run, will be felt by an unpredictable and diffuse set of export interests who each face relatively small expected costs of retaliation at the time a violation is contemplated. Collective-action problems will impede them from organizing to oppose the violation.

Our answer here is much as before. Exporters will form associations to internalize these diffuse costs and try as best they can to install agents in the political process who will take proper account of them. Furthermore, nations fearful that another nation may violate its commitments will find it in their strategic interest to preannounce targets of retaliation in order to mobilize them. The common practice in Washington, D.C., of drawing up "retaliation lists" in Section 301 cases prior to the imposition of any actual sanctions serves precisely this purpose.\textsuperscript{51}


\textsuperscript{51} See, for example, GATT: U.S., EC Announce Breakthrough in GATT Trade Talks, Oilseeds Dispute, 9 Int'l Trade Rep. (BNA) 1990 (November 25, 1992) (chronicling U.S. trade repre-
4. Summary

For the reasons given here, the level of compliance with trade commitments is quite high even if there is no credible threat of sanctions for misbehavior. The GATT system thus worked quite well without sanctions, and were it not for the recent innovations in the DSU we could end here. But the drafters of the WTO agreements decided to replace the old GATT dispute resolution system with a meaningful prospect of formal sanctions for violations that are not corrected after a reasonable period of time. We now offer an explanation for those changes in the system.

B. The New Prospect of Sanctions for Violations That Are Not Cured within a Reasonable Time

The DSU changes the rules and embraces a “reverse consensus” principle, under which sanctions will be authorized after the dispute process has determined that a violation exists and a reasonable time for cure has elapsed, unless a consensus exists against sanctions (which would have to include the party (or parties) that filed the complaint and prevailed). Consequently, sanctions are a real threat to the recalcitrant violator and have already been employed a number of times. What has not been assigned sufficient importance, however, is that although the DSU made sanctions a real possibility, it did not change another feature of the system that greatly restricts the value of sanctions in inducing nations to comply with their obligations. As we have noted, a sanction cannot be imposed until a dispute panel finds that a violation has occurred, the appellate body affirms the panel’s finding if an appeal is filed, and the violation continues although a reasonable period of time to cure it has elapsed. Thus, the sanction operates only prospectively. As a result, a country can commit a violation and continue it for a considerable time without incurring any formal penalty.

What is the logic of this new system? Does its adoption put the lie to our claim that formal sanctions are not necessary to achieve a high level of compliance? And if formal sanctions are indeed important to deter violations, why limit them to violators who have been caught and continue to cause harm after they have been given a chance to reform their behavior? Does that not invite cheating in hopes of avoiding detection, followed by delay when caught to exploit the reasonable time for cure?

To answer these questions, we begin by reiterating the claim that flagrant cheating has been uncommon in the system through the years. Domestic political constraints, concerns for reputation, and unilateral sanctions indeed
produced a high level of compliance under the GATT, and there is little reason to think that they would not work similarly in the WTO. As a result, many (although not all) of the disputes that arise involve good-faith clashes over ambiguous terms of the bargain. In these circumstances, countries are often genuinely uncertain about what they are obliged to do, and sanctions may have the effect of punishing them for good-faith behavior. Not only is there little deterrence value to such punishment, but it may prove somewhat destabilizing to the trading system and provide further political ammunition to those who would scuttle it on the basis of sovereignty claims and the like.

Indeed, as suggested earlier, there may be instances in which WTO provisions have been intentionally left vague because an expert body, deciding ex post what conduct is value maximizing, may be a better instrument for facilitating mutually advantageous conduct than the ex ante predictions of members as to what will be in their mutual advantage in the many circumstances that may arise. A country found to be in violation of such obligations after the fact may thus have supplied a public good by becoming the test case on a particular issue. The absence of sanctions for behavior prior to an adverse ruling may thus be seen as a way to encourage nations to litigate their disputes to conclusion so as to clarify the rules for everyone.

Once an adverse ruling comes down, however, matters change. If rulings are indeed constructive gap fillers, compliance with them will ordinarily generate joint gains, and renegotiation of most-favored-nation tariff commitments will often (although not always, as discussed earlier) be a better way to protect import-competing industries than ongoing violations of non-tariff provisions.

Does this line of reasoning explain why sanctions are now available, following the lapse of a reasonable time for cure? We must answer no, lest we introduce inconsistency into our argument. Once a ruling adverse to a WTO member is issued, a refusal to comply would otherwise be subject to the same reputational penalties and unilateral sanctions that we discussed earlier. Why are they not enough, and what motivated the drafters of the DSU to introduce the prospect of meaningful sanctions at this stage of the process?

Our answer is to suggest that the innovation of the DSU was intended not so much to deter violations of most substantive rules, for such violations, if clear, were already fairly well deterred, as we have argued. What the new system really adds is the opportunity for the losing disputant to "buy out" of the violation at a price set by an arbitrator who has examined carefully the question of what sanctions are substantially equivalent to the harm done by the violation.

By contrast, the situation immediately prior to the entry into force of the WTO was one in which unilateral retaliation was becoming more and more common. Cases under Section 301 of the U.S. Trade Act of 1974, in particular, were becoming more frequent. And the 1988 amendments to that statute created a timetable for retaliatory action by the United States that could
require it to sanction an alleged violator even before the GATT panel process had run its course. The distinct possibility thus arose that the United States would impose a sanction based on a unilateral determination that another party was in breach of the GATT, even if a dispute panel would find that the U.S. complaint lacked merit. Further, although the statute directed the U.S. trade representative to impose a sanction commensurate with the burden on U.S. commerce caused by the violation, the United States was the sole arbiter of whether its sanction in fact met that test. The international community might thus have reason to be concerned that the United States might impose sanctions for the purpose of foisting an opportunistic construction of the bargain on trading partners or that the unilateral sanctions might be excessive and discourage efficient breach. The political pressures for other trading nations to arm themselves with similar potentially disruptive unilateral strategies were no doubt considerable. Thus, a fear was developing that unilateral sanctions in the name of enforcing the bargain were being co-opted in a way that would allow trading nations to renego on the bargain.

The skeptic will immediately wonder, however, why this type of reneging was not adequately constrained by the factors that we claimed were reasonably effective under the consensus-based system—reputation and unilateral sanctions. In one important respect, we believe it was. In particular, a threat of unilateral sanctions could in principle be employed to induce a country to accept an excessively demanding interpretation of its obligations. But, as one of us has argued, this did not occur. The United States agreed to submit its claims of violations to determination by the GATT, and, unless dispute resolution dragged on too long without result, the United States also committed itself to await the GATT determination and be bound by it. In fact, the United States never acted contrary to a GATT ruling—it never "took the law into its own hands" in finding a violation by another party. Accordingly, one of us previously concluded that U.S. actions under Section 301 did not advance excessively demanding interpretations of the obligations owed to it under the GATT. Such analysis suggests that with respect to the existence of a violation by another nation, the United States was substantially, if not perfectly, constrained by reputational concerns, reinforced by the possibility of authoritative GATT rulings against it and a fear of unilateral retaliation for misbehavior.

There is another dimension to unilateral sanctions and threats, however, that may have been subject to greater strategic manipulation. When a violation occurs, a system of unilateral retaliation leaves it to the aggrieved nation to

---

54 See 19 U.S.C. §§ 2411(c) (requiring that sanctions must "be devised so as to affect goods or services of the foreign country in an amount that is equivalent in value to the burden being imposed by that country in U.S. commerce").
set the magnitude of the sanction. Although the GATT system had always required that any sanction be substantially equivalent to the harm done by the violation, the question of whether an actual or threatened sanction was excessive by this standard might be one about which the members of the trading community have very poor information. Indeed, the harm caused by a violation to another country or to its political officials is almost certainly difficult for other nations to ascertain. If so, a nation injured by a violation might be able to threaten or to impose an excessive sanction without incurring reputational penalties. Consequently, excessive actual or threatened unilateral sanctions may have become an important actual or potential impediment to efficient breaches within the system.  

Under the new DSU, by contrast and as noted previously, a binding arbitration system is established to consider the magnitude of the sanctions. No sanctions can be imposed until the arbitration process has run its course if the violator nation insists on it. The new system thus does a better job of protecting violators from the actual or threatened imposition of excessive sanctions. In turn, it ought to perform better than the old system at ensuring that opportunities for efficient breach are not undermined.

The same conclusion can be reached in another way, which we think bolsters it considerably. In particular, consider the question of what really changes under the new DSU. Penalties for breach existed under the old system, both in the form of reputational costs and unilateral sanctions. The penalties for breach under the new DSU will, as a practical matter, be much the same—reputational costs will attach to roughly the same degree as before (especially after the 1989 understanding that eliminated the ability of disputants to block the formation of a panel), and any sanctions approved by the DSB will continue to take the form of measures by the aggrieved country (or countries) to punish the violator through a withdrawal of trade concessions. The primary difference is that those measures can now be reviewed by a binding arbitral panel for excessiveness before they can be put into place, whereas before they were unilaterally announced and implemented without review by the GATT.

Thus, the innovation of the new DSU is very much consistent with, and we believe motivated by, the perception that unilateral sanctions were in need of greater centralized oversight. If we are right, then the reason for authorizing sanctions against recalcitrant violators in the new DSU is not to punish them so much as to protect them—instead of having to buy their way out in a world of unilateral threats and counterthreats unconstrained by central oversight, the new system ensures that the price for noncompliance will be set in accordance with an honest and unbiased effort to assess the harm to the affected party (or parties).  

V. Conclusion

Treaties are contracts of a sort, and the lessons developed by law and economics scholars regarding the way that private contracting parties structure their bargains accordingly have much to teach us about the structure of treaties. In this paper, we have argued that the WTO system prefers a liability rule to a property rule—roughly, expectation damages to a rule of specific performance—primarily because of the transaction costs and holdup problems that would arise under a property rule in a system with 144 players. We have further advanced a theory as to how formal sanctions are not needed to induce a high level of compliance with most WTO obligations, owing to the domestic pressures for compliance that often exist and to the reputational penalties and unilateral sanctions that further pressure parties to respect their commitments even absent formal sanctions. The value of dispute resolution cases, therefore, may lie more in clarifying the rules and filling in missing terms of the bargain than in detecting and punishing cheaters. It is for this reason that the losing party in a dispute proceeding pays no penalty if it obeys the recommendations of the dispute process. Last, we argue that the recent advent of formal sanctions for parties that lose a dispute proceeding and refuse to conform their policies within a reasonable time is a response not so much to the undercompliance with substantive obligations that arises absent these sanctions, but to the danger of excessive unilateral sanctions that exists in the absence of centralized oversight regarding the magnitude of sanctions. This problem arises because the harm done by a violation is not easily observable absent a careful examination by an arbitrator, and thus a party that imposes an excessive unilateral sanction will be hard to detect and so will not suffer the usual penalties associated with misbehavior. The new arbitral process substitutes an unbiased determination as to the proper magnitude of the sanction for a unilateral judgment about it by the aggrieved party and thus better ensures that the price for deviating from WTO obligations is not set inefficiently high.