
Role of Remedies in the WTO System

An inherent tension exists in the drafting of trade rules. On the one hand, *precise* formulations, which are explicit about purpose and clear about meaning, facilitate implementation and adjudication. On the other hand, *vague* formulations, open to a variety of interpretations, facilitate agreement and adoption. By convention, major WTO decisions require a consensus, and since members often have differing views and interests, it is not at all surprising that the rules are characterized by what is sometimes called constructive ambiguity. The result is that agreements are subject to very different interpretations that reveal more about the perspectives of the interpreters than about the meaning of the text. This is particularly the case when those approaching the WTO have strong normative preconceptions that color their views of what the system should be. Three sets of such preconceptions are considered in this chapter. They relate to the legal nature of the system, the purpose of WTO dispute settlement remedies, and the relationship between the systems for obtaining and enforcing agreements.¹

1. Had this study focused on the specifics of panel findings, I would have emphasized another preconception that relates to the very strong presumption that the General Agreement on Tariffs and Trade (GATT) was intended to promote a liberal trade ethos, regardless of what might have been written in the text. Robert Howse (2000, 39) argues that this perspective heavily colored the panel findings in the *Tuna-Dolphins* case. He also quotes Robert Hudec, who states that “GATT rulings often expressed an intuitive sort of law based on shared experiences and unspoken assumptions.”

Another powerful notion, discussed below, relates to the idea that the GATT has successfully channeled mercantilist folly to the purposes of achieving freer trade.

This chapter starts by outlining these preconceptions. It then describes the central role the paradigm of reciprocal concessions plays in the WTO both when negotiating agreements and when responding to violations. Finally, it considers rationales for the system's design and the degree to which the system actually accords with the preconceptions about how it is supposed to work.

Preconditions

Perspectives differ radically on the nature of the WTO rules and, in particular, the consequences of their breach. In a public international law sense, the WTO agreement is a treaty, and WTO obligations are obligations under public international law. But should the implications of breaching WTO rules be treated like breaches of a commercial contract—albeit among countries—or should they be treated as infractions of obligations of more traditional legal codes? Leading legal scholars can be found on both sides of this debate.² Informed by their perspectives, they hold very different views on the nature of WTO commitments and the manner in which remedies should be applied.

One side of this debate emphasizes that the WTO agreement resembles a commercial contract.³ Proponents of this view sometimes point out that members of the General Agreement on Tariffs and Trade (GATT) were referred to as “contracting parties.” Similarly the process is called “dispute settlement” rather than adjudication. Those with this view would see the WTO commitments as reciprocal obligations like those assumed in commercial contracts.⁴ The parties agree to constrain their behavior because they believe it will be mutually beneficial. If, subsequently, one of the parties decides compliance is no longer in its interest, it can provide adequate compensation and be relieved of its obligations.

2. Kenneth Dam (1970, 251), in his pioneering study of GATT, emphasized that it was an unusual international agreement. “The fundamental decision was taken at the very beginning to refrain from the adjudicatory approach to dispute settlement. The General Agreement contains no provision for references of disputes or questions of interpretation to the International Court of Justice. Nor is there any provision in the General Agreement for an internal tribunal to . . . promulgate authoritative interpretations on questions of interpretation.” “Illegality” in GATT is “an uncertain and ambiguous concept. Although the substantive provisions . . . are drafted in conventional terms, including a rather liberal use of prohibitory language, the remedy provisions are not drawn in terms of sanctions.” “A failure to respect a tariff concession is usually not regarded as a transgression to be punished but rather as an event giving contracting parties the privilege, subject to the approval of the Contracting Parties, of suspending reciprocal concessions” (Dam 1970, 352).

3. See, for example, Bello (1996) and Schwartz and Sykes (2002). For a different view, see Jackson (1997).

4. For an argument that WTO obligations are reciprocal, see Pauwelyn (2002).

In this construct, the party that runs afoul of an agreement might be likened to a renter who breaks a lease. Such a breach is not an act of negligence nor is it a crime or infraction that should be punished. Contracts are, after all, not social obligations. Breach is a matter between the parties and does not violate commitments to the community (members) as a whole. Once it provides adequate compensation, the promisor in a contract is usually fully relieved of its obligations. To be sure, its reputation and future credibility may suffer as a result of the breach. However, if it decides that it would be better off by breaching the contract, and if the promisee can be left in a position that is no worse than had the contract been fulfilled, the society benefits from allowing such a breach.

The other side has a different perspective on WTO rules—that they are binding in the same way as other international agreements. Implicitly, therefore, parties should comply with WTO rules even when continued performance of the agreement is no longer in their interest and even if their behavior has no measurable impact on the trade of other members. In particular, even if a losing member provides compensation, is subject to retaliation, or strikes some other political bargain that leads the plaintiffs to drop their case, it is not relieved from the obligation to comply with the original agreement. Viewed from this orientation, violations are akin to statute infractions. They are ultimately committed against WTO members as a whole—technically obligations are multilateral, or *erga omnes partes*—and not simply against members with a direct interest in the matter. The proponents of this view would like the WTO system to evolve even further in the direction of legalization and constitutionalization.⁵ One argument is that this would create greater certainty for private economic rights, broadening the institution's purview from an agreement that is only among nation-states. Some would like the WTO Dispute Settlement Body (DSB) to follow precedent and establish a common law. Some others advocate responses to violations that exceed the impact on the trade of a complainant—provide for reparations, induce compliance, and even punish the violator.

The tension between the views on the purpose of the responses to violations has a long history. John Jackson's (1969) authoritative description of the preparatory work for the original GATT Articles XXII and XXIII notes that drafters had different positions. He observes, "Domestic law analogies to these two positions might be the law of contract damages on the one hand (redressing the damage done to expectations but generally banning punitive damages) and the criminal or tort law on the other hand (whereby sanctions or punitive responses are authorized and not limited by the actual amount of damage caused)" (Jackson 1969, 170).

5. See, for example, Pauwelyn (2001). For a cautionary note, see Goldstein and Martin (2002) and Howse and Nicolaidis (2001).

These differences in perspective may have important policy implications. In particular, those with the contract view will be more likely to see the WTO as more compatible with national sovereignty, less inclined toward punitive retaliatory responses, and more willing to see a role for diplomatic solutions to conflict. By contrast, those endorsing the alternative perspective see a greater role for international legal rules and are more likely to endorse stronger and even punitive responses to infractions.

Why Remedies?

Differences on how to respond to violations come out clearly in a related dispute about the purpose of retaliation in the WTO system. Since the purpose is not spelled out explicitly in the agreement, this creates lots of opportunities for disagreement. In one case, the WTO talks of allowing “appropriate countermeasures.”⁶ But what is “appropriate”?

Compliance

Some naturally presume that retaliation should be a trade sanction imposed to induce compliance. This view follows naturally if the system is presumed to be similar to other enforceable rules-based systems, such as codes of criminal law, although in principle it is also compatible with the idea of the WTO rules as a contract in which there are consequences for breach. This perspective certainly meets the political needs of the officials from a country that wins a trade dispute but is unable to obtain a satisfactory response from the loser. It also fits naturally into the classic prisoner’s dilemma paradigm used by economic theorists, which assumes that self-interested parties that could benefit from breach must be induced to comply with the prospect of punitive consequences in response to infractions.⁷

Those who express concerns about the erosion of national sovereignty generally take the punitive nature of the system for granted. The complaints of Ralph Nader (cited in chapter 1) and others exemplify this view, which is shared by those who experience what Steve Charnovitz (2001) aptly terms “sanctions envy” and would like the WTO to enforce their particular rules such as environmental or labor standards.

But others argue that punishing sovereign nations is both ineffective and unnecessary. Abram Chayes and Antonia Chayes (1995), for example,

6. See WTO Agreement on Subsidies and Countervailing Measures (SCM) Article 4.10.

7. The theoretical literature usually models compliance with trade agreements as a repeated game in which each country weighs the benefits of deviating from its commitments against the costs implied by the responses of other countries. In this system, the amount of liberalization is determined by how severe and credible the likelihood of punishment is. See Staiger (1995).

suggest that compliance with international agreements is generally preferred for three principal reasons: (a) it enhances efficiency by saving on transaction costs, (b) it occurs because treaties reflect national interest, and (c) there is a widely held social norm that the law should be obeyed. An additional reason is the concern that countries have about their reputations (Kovenock and Thursby 1992). Breaching an agreement even when the specific costs of compliance are greater than the benefits can nonetheless do long-term damage to a nation's ability to obtain future agreements. It could also damage its nontrade relationships with other countries.

The relationship between the severity of retaliation and compliance may not be linear. If enforcement were made too strong, and a country were to retaliate in a manner viewed as wholly disproportional, this could spark (unauthorized) counterretaliation by its trading partner and result in an escalating trade war.⁸ Thus, paradoxically, keeping responses within bounds is another role of the system.

Compensation

A second preconception is that retaliation is meant to provide compensation for the wronged party. Indeed, as will be discussed in greater detail, some adopt the contracting perspective, arguing that with retaliation, the equivalent of "expectation damages" are provided to the promisee in the event of a contract breach. Viewed from a mercantilist perspective, this notion appears justified. If the defendant has improved its balance of trade as a result of the violation, it is only proper that the plaintiff be allowed to recoup its losses by raising barriers of its own. To be sure, this view is resisted by those who see mercantilism as folly and the impact of protectionism as harmful to the retaliating country. In particular, it is hard to square the idea that allowing the plaintiff to "shoot itself in the foot" is a mechanism for providing compensation. But as noted later, countries may indeed incur both political and economic costs in making agreements and may reduce these costs through retaliation.

Legal Breach

A third and less common notion is that WTO remedies should provide a legal mechanism for countries to escape from their commitments. Members may commit to agreements in good faith, but if circumstances change, they may find compliance too onerous. Proponents of this approach argue that the system should allow countries to escape from such burdens under certain conditions.⁹ To be sure, these conditions should not be too

8. This possibility, characterized as "off-equilibrium path retaliation," is discussed in Bagwell and Staiger (2000).

9. See, for example, Rosendorf and Milner (2001).

lenient since this might undermine compliance. But they should also not be too onerous since this might discourage members from negotiating agreements in the first place.

Balancing Concessions

A fourth notion is that retaliation is simply a mechanism to maintain the balance of concessions negotiated through the WTO. WTO members are presumed to accept obligations only because they have received rights in return. In particular, they have agreed to reduce their own trade barriers and in return have received increased access to other members' markets. When a member violates an agreement in a manner that nullifies concessions to which it has agreed, it is only fitting that others be permitted to withdraw concessions they have made.¹⁰

An Integrated System?

Another preconception in many discussions of the WTO system is that the system for dispute settlement can be considered separately from the system for negotiations. Some of the concerns about the system of retaliation described in chapter 1 have this character. For example, some who portray retaliation as a penalty neglect to consider that parties losing a WTO case have in principle reneged on an agreement in which others made concessions. As elaborated in greater detail below, when plaintiffs retaliate, they are not applying a punishment but returning to the status quo before the original agreement.

Similarly, those who express concern that smaller countries have greater difficulties in retaliating sometimes overlook that retaliation is the mirror image of (a) the advantages these countries might enjoy from free riding in a negotiating system based on the most favored nation (MFN) principle and (b) the advantages developing countries might enjoy from the special treatment that allows them to receive the benefits of WTO concessions without being required to make reciprocal concessions.

Those who advocate making the enforcement either stronger or weaker sometimes fail to understand or consider that the mechanisms used in the dispute settlement system will affect the negotiations system by influencing the extent to which members are willing to liberalize in the first place. On the one hand, if the dispute settlement system is too weak, it could discourage liberalization by raising concerns about the likelihood of compliance and the credibility of commitments.

10. As elaborated later, balancing concessions is not the same as providing compensation. Compensation could be expected to leave the injured party in a condition no worse than had the agreement been fulfilled. In contrast, the balance of concessions could be achieved either through providing new concessions or restoring the status quo ante.

On the other hand, if it is too strong, it could create disincentives to make commitments because of concerns of the costs of breach when circumstances change. Stronger penalties, for example, are presumed to improve compliance, but if they put a damper on new negotiations, their impact could actually be perverse because ongoing negotiations encourage compliance. Members comply, in part, because maintaining a reputation for compliance is important not only to preserve existing agreements but also to extract concessions in the future.

In sum, the WTO system is complex and subtle. The merit of the criticisms of the dispute settlement system and suggestions for the reform of dispute remedies will depend on the kind of legal system that should be implemented, the purposes that the remedies should serve, and the links that are desirable between the dispute settlement and negotiations systems. Thus far, however, the discussion has dealt with preconceptions about what the system should do. Let us now analyze what it actually does and then evaluate the degree to which the system accords with the preconceptions.

The Reciprocity Paradigm

The preamble to the Agreement Establishing the World Trade Organization lists the goals of parties to the agreement. These include “raising standards of living, ensuring full employment and a large and steadily growing volume of real income and effective demand, and expanding the production of and trade in goods and services.” It notes the parties’ desire to achieve these goals “by entering into *reciprocal and mutually advantageous* arrangements directed to the substantial reduction of tariffs and other barriers to trade and to the elimination of discriminatory treatment in international trade relations” (emphasis added). Thus WTO agreements are presumed to result from countries making concessions on a reciprocal basis. This presumption lies at the heart of what I shall term the WTO paradigm, the crucial notion that guides both agreements and dispute settlements.

Concessions

A central conception in the paradigm is that when a country agrees to reduce a trade barrier, it is making a concession. Article II of GATT, for example, is not entitled “schedules of tariff reductions” but “Schedules of Concessions.” This idea is fundamental to how the WTO operates. Since reducing import barriers will generally increase national welfare, economists sometimes scoff at this notion of a concession. Why should a

measure that is in the nation's interest and should in any case be adopted unilaterally be considered a concession?

Mercantilism. Some see this talk of concessions as a brilliant subterfuge that harnesses mercantilist sentiment in the name of free trade.¹¹ In this view, as vividly spelled out by Paul Krugman (1997), the WTO tacitly accepts the naïve (and incorrect) notion that “exports are good” and “imports are bad” and therefore treats measures that increase imports as concessions. Economists sometimes rationalize playing along with mercantilist views when they see these serving a greater good. But their deeply held view that this perspective is fundamentally misguided sometimes leads them to not take certain elements in the WTO system very seriously. In particular, the ideas that trade liberalization entails a concession and that retaliation could actually operate as a form of compensation for the winner of a trade dispute appear highly questionable.¹²

Terms of Trade. There are, however, alternative explanations that rest on more solid conceptual grounds. Tariff reductions will improve domestic efficiency, reducing distortions to consumption and production decisions. Under competitive conditions, therefore, if a small country removes tariffs, it will raise national welfare. However, for countries large enough to affect prices in world markets, the calculation is more complicated because tariff reductions may also worsen a nation's terms of trade. When a large country lowers a tariff barrier, it increases demand for the product on world markets. This increased demand makes its imports more expensive to the country.¹³ Moreover, to pay for these imports and restore balanced trade, the country will have to export more, and this could reduce export prices. These responses will reduce the nation's terms of trade (the ratio of export to import prices) and offset the welfare benefits that occur through the greater efficiency associated with reducing the trade barrier.¹⁴ Thus countries reducing trade barriers could be making “concessions” in the sense of taking actions that, everything else being equal, could reduce their welfare.

11. Krugman (1997) observes in GATT-think: “(1) exports are good; (2) imports are bad; and (3) other things equal, an equal increase in imports and exports is good. In other words, GATT-think is enlightened mercantilism.”

12. According to Krugman (1997, 115), “If economists are sometimes indulgent toward the mercantilist language of trade negotiations, it is not because they have accepted its intellectual legitimacy but either because they have grown weary of saying the obvious or because they have found that in practice this particular set of bad ideas has led to pretty good results.”

13. The domestic price will equal the world price plus the tariff. When tariffs are reduced, the world price will rise but the domestic price paid by consumers will generally fall.

14. For an analysis of how tariffs affect the terms of trade, see Krugman and Obstfeld (2003, 195–96).

Domestic Politics. The third explanation is based on domestic policies. Reducing trade barriers could create losers as well as winners. With lower import barriers, consumers will gain but import-competing producers will lose. Under competitive conditions, the gains to consumers will outweigh the losses to producers. However, producers are generally well organized politically, while consumers face serious problems in acting collectively. Accordingly, policymakers may find liberalization politically painful. From their standpoint, the creation of losers among producers who compete with imports could therefore represent a concession.

Realpolitik. A fourth explanation reflects international concerns of the realist school of international relations. Conventionally, economists assume the goal is only to maximize absolute national incomes, but in “realist” political views, countries are also concerned with their relative incomes.¹⁵ Realists see states as living in an anarchic world in which competition and conflict are inevitable. The win-win character of trade, which economists find so pleasing, actually presents a problem from this perspective. Reducing import barriers may raise domestic incomes, but it could also raise incomes abroad by more than it raises incomes at home. In this framework, any measure that risks enhancing foreign welfare by more than it raises welfare at home could therefore be seen as a concession.¹⁶

Reciprocity

A second key notion in the WTO system is the paradigm that all concessions in the WTO are made on a reciprocal basis.¹⁷ What does “reciprocity” mean? According to economists Kyle Bagwell and Robert Staiger (2000, 37), “The principle of reciprocity in GATT refers to the “ideal” of mutual changes in trade policy, which bring about changes in the volume of each country’s imports that are of equal value to changes in the volume of each country’s exports.”

Thus concessions are balanced or reciprocated when they result in equal trade flows. Although it is nowhere defined explicitly, implicit reciprocity

15. For a more detailed description, see Grieco (1990) and Stein (1990).

16. A fifth reason is that by signing an agreement (that binds tariffs at a particular level or accepts certain rules), countries are agreeing to limit their future policies, even if circumstances change. Such limitations are concessions from countries that could otherwise change their policies freely.

17. Dam (1970, 58) recalls that the Havana Charter emphasized “no Member shall be required to grant unilateral concessions.” He later notes, “from the formal legal principle that a country need make concessions only when other contracting parties offer reciprocal concessions considered to be mutually advantageous has been derived the informal principle that exchanges of concessions must entail reciprocity” (p. 59). Thus while the GATT does not formally require that negotiations produce balanced concessions, it is implicitly assumed that they have done so.

in the WTO is thus used in a specific sense. WTO members are neither required to completely remove their trade barriers nor generally required to have the same tariff levels, either on average or in specific commodities. Instead, as a result of each negotiation, members are expected to give, in value, the same *new* trading opportunities as they receive. Such a system is based on what Jagdish Bhagwati (1991) has termed “first difference” reciprocity. Countries take their initial starting positions as given and then establish equivalence between the levels of concessions, not the overall level of trade barriers.¹⁸

For example, the United States agrees to lower its tariffs on imports of a product—say, a particular type of industrial chemical sold only by Poland. In return Poland agrees to lower its tariffs on a product the US exports—say, a particular type of computer that is made only in the United States.¹⁹ If the concessions are balanced and the lower Polish tariffs on computers result in an extra billion dollars’ worth of US exports, then the lower US tariffs on industrial chemicals should result in an extra billion dollars’ worth of Polish exports.

The value of trade induced by a tariff reduction will reflect (a) the volume of trade covered by the reduction, (b) the size of the tariff reduction, and (c) the behavioral responses these reductions will induce. In practice, however, some negotiators may never actually estimate these variables. Instead, some may simply use readily available proxies such as trade coverage—the volume of trade covered by the tariff reductions—taken alone or each multiplied by the percentage reduction in the relevant tariff. Since they are incomplete, such measures may be flawed. Other negotiators may prefer not to be precise about their estimates so that they can apply different political weights to particular concessions (Dam 1970, 59). (Returning to the example, if the US industrial chemicals lobby is strong, an additional billion dollars of computer exports is worth only half a billion dollars of industrial chemicals imports.) Nonetheless, the *idea* of reciprocity through balanced concessions plays a crucial role in the WTO paradigm. Countries that consent to a trade agreement are presumed to have received “concessions” that offset exactly those that they have made.²⁰

Reciprocity plays an important role not only in the negotiations for new trade agreements but also in negotiations for new memberships. As Joost Pauwelyn (2000) emphasizes, the WTO differs from many other international agreements in that membership does not simply entail a

18. Ironically, the idea that agreements are supposed to be balanced does not stop politicians from implying that they will boost aggregate employment and create jobs!

19. Typically, these concessions are made to several principal suppliers, and the concessions received may be from other parties.

20. For more on this, see Bagwell and Staiger (2000). For skeptical views on reciprocity, see the papers in Bhagwati (2002).

willingness to accept certain obligations in return for certain rights. Acceding members are required to make concessions when joining in order to balance the concessions that existing members are presumed to have made.

The principle that countries are not expected to make concessions unless other nations provide them with reciprocal benefits comports with the interpretations of concessions already discussed. From a mercantilist view, if imports are bad, they should be increased only in return for additional exports. In this view, agreements that increase the trade balance are particularly good, while those reducing it are especially bad. If a country obtains an agreement that boosts its trade balance, others must be experiencing declines. However, a system that preserves balanced trade is more plausibly fair for all.

From an economic standpoint, reciprocity takes care of the terms-of-trade problem that large countries face when liberalizing unilaterally. As Bagwell and Staiger (2000) have emphasized in a model with two large countries, world relative prices will remain unchanged if the value of increased imports due to an agreement is offset by an equal value of exports.²¹ Thus negotiations based on reciprocity allow countries to set their tariffs in a manner that is desirable from a domestic viewpoint without having to worry about their terms of trade.²²

From a domestic political standpoint, reciprocity helps deal with the collective action problem that hinders unilateral liberalization because consumers are poorly organized. The reduction in foreign trade barriers generates support from producers in export industries, which can offset the opposition of producers competing with imports.²³ From an international political viewpoint, reciprocity appears to be a rough-and-ready way of ensuring that relative gains are similar.²⁴

21. Bagwell and Staiger (2000) show that if two countries act independently and take each other's tariffs as given, they will reach a Nash equilibrium that is not Pareto optimal. Each is reluctant to liberalize further because it will be hurt by the associated decline in its terms of trade. However, in a cooperative negotiation based on reciprocity, each can be sure that its trading partner's actions will keep the terms of trade constant. This allows both countries to liberalize further, and they will achieve an agreement that is Pareto optimal.

22. If the countries are simply maximizing domestic welfare they will move to free trade. If, for political reasons, they apply different weights to output, they may prefer to have tariffs.

23. In the formal economic explanation that relies only on the terms-of-trade effects, there is no need for a separate domestic politics rationale for reciprocity. However, despite the very impressive theoretical accomplishment of Bagwell and Staiger (2000) there remains considerable skepticism about explanations giving terms-of-trade effects a dominant role particularly since these are almost never mentioned (and rarely recognized) by trade policymakers. See the discussion on the practical relevance of terms-of-trade considerations in Bagwell and Staiger (2000, chapter 11).

24. Actually, even if concessions are balanced this way, the gains could be very different.

Politically, relative gains may matter. In a bilateral trade negotiation, if nation A receives some benefits, but the benefits to nation B are perceived to be even greater, then nation A may sense that the deal is unfair. In addition, if the other side is left with large benefits, the negotiators for nation A may face complaints that they have not done as well as they might have. Reciprocity helps quell these concerns and aids in the normally difficult relationship between negotiators (the agents) and their domestic constituents (the principals).

Making reciprocity a foundation for negotiations also has a practical side. As John McMillan (1988) has pointed out, multilateral trade negotiations are extremely complex, and members may have to consider a wide variety of outcomes. But if the principle of reciprocity is taken as the basis of the negotiations, only those options that yield balanced trade need be evaluated.²⁵ Drawing on the work of Thomas Schelling (1963), McMillan argues that reciprocity serves as a focal point that limits potential options for agreement. He comments that even if the notion is inherently arbitrary, it serves to reduce transaction costs and thus facilitate agreement.

To be sure, there are some departures from reciprocity in the WTO rules. Developing countries, in particular, are supposed to enjoy “special and differential treatment.” This principle is enforced in many areas relating to rules. Also, developing countries are not expected to make reciprocal concessions and are allowed to obtain special preferences under the Generalized System of Preferences (GSP), which is separately administered by developed countries and does not require reciprocal concessions.²⁶ In addition, in practice, since the WTO is based on unconditional MFN treatment, small countries may be able to free ride on reductions in tariffs negotiated by others.

However, basing negotiations on reciprocity also has a downside for these countries. With reciprocity, what nations get is determined by what they give. It is therefore no surprise that the most significant process in reducing trade barriers has been in the products that are of concern to the developed countries—that is, tariffs on industrial products. In contrast, barriers remain significantly high in agriculture, textiles, and other labor-intensive manufactures produced by developing countries.²⁷ Moreover, as

25. According to McMillan (1989), “In the case of GATT talks the focal point is generally nothing more than a measure that takes into account the relative size of different countries and is simple to calculate using data that are readily available.”

26. The WTO ministerial declaration on the Doha Round, for example, states, “the negotiations shall take fully into account the special needs and interests of developing and least-developed country participants, including through less than full reciprocity in reduction commitments.” www.wto.org/english/thewto_e/minist_e/min01_e/mindecl_e.htm.

27. To be sure, the failure to liberalize these sectors reflects the domestic strength of protectionist forces in developed countries in addition to the reluctance of developing countries to negotiate.

elaborated later, the inability to retaliate is inversely related to the ability to free ride.

Violations

What role does reciprocity defined in terms-of-trade effects play in the response to violations? It is important to distinguish between a breach of a legal obligation or rule and the “nullification and impairment of benefits” that result from such a breach. The former refers simply to the act of violating the agreement; the latter is generally interpreted to refer to its impact on trade.

As John Jackson (1998, 332) has documented, “Originally the key to invoking the GATT dispute settlement mechanism was almost always ‘nullification and impairment.’” However, this changed over time and Article 3.8 of the Uruguay Round Understanding on Rules and Procedures Governing the Settlement of Disputes (DSU) states, “In cases where there is an infringement of the obligations assumed under a covered agreement, the action is considered prima facie to constitute a case of nullification and impairment. This means there is normally a presumption that a breach of the rules has had an adverse impact on other Members’ parties to that covered agreement, and in such cases, it shall be up to the Member against whom the complaint has been brought to rebut the charge.”

Jackson observes, “This makes the presumption of nullification and impairment derive ipso facto from a violation, thus almost discarding the nullification and impairment concept in favor of a focus on whether or not a ‘violation’ or ‘breach’ of obligation exists.”²⁸

Nonetheless, even if violation does create a presumption of nullification and impairment, *the trade effects of the nullification* play a critical role in determining the responses that are allowed because of such violations. Indeed, the manner in which the WTO responds to violations in the agreement follows quite naturally from the manner in which the agreement was negotiated in the first place. Reciprocity is a central principle not only in negotiations but also in allowed responses under the DSU. Such allowed responses, particularly retaliation, relate to nullification and impairment of benefits between the parties rather than violations of the

28. In a case brought by Uruguay in 1962, as noted by Jackson (1998), a panel argued that “any violation of the GATT would be considered a prima facie nullification or impairment.” This represented a shift in the direction of rule orientation rather than simply facilitating settlements. Jackson also notes that in 1988, in the Superfund case, the United States conceded that it had violated the national treatment provisions of Article III by imposing a higher tax on imported products. However, it tried, unsuccessfully, to have the case rejected on the grounds that the tax had had no measurable impact on trade flows. However, in that case, the panel said that such a violation created a presumption of nullification that was not rebuttable. More recently, the US 1916 Anti-Dumping Law was held to be a violation even though it had never been applied.

rules in general. As Joost Pauwelyn (2000) observes, “What is *actionable* under the WTO is not so much the breach of obligations, but the upsetting of the negotiated balance of benefits consisting of rights, obligations and additional trade concessions” (emphasis added).

Although reciprocity is a desired outcome, and negotiations are presumed to result in reciprocity, it is not formally required in negotiating rounds. There are no officially published estimates of the trade volumes granted and received. Indeed, since WTO agreements increasingly relate to the acceptance of obligations to adhere to rules rather than to reduce tariffs, the reciprocity paradigm has less relevance. The agreements relating, for example, to Application of Sanitary and Phytosanitary Measures (SPS) and Technical Barriers to Trade (TBT) do not readily fit into the reciprocity framework. However, reciprocity *is* formalized in the responses the WTO allows other members to make when agreements are violated.²⁹ (See box 2.1.)

If a defendant is found to have nullified its commitments, the agreement would no longer provide reciprocal benefits. How could the agreement be rebalanced? There are three ways: (1) the defendant could eliminate the regulation and comply with the agreement; (2) the defendant could grant the plaintiff another concession (compensation); or (3) the plaintiff could withdraw concessions to the defendant (suspension of concession). Any of these three options could achieve a rebalancing.

The WTO seeks the first solution—having the defendant comply with the agreement. As noted in Article 3.7 of the DSU, “A solution mutually acceptable to the parties to a dispute and consistent with the covered agreements is clearly to be preferred.” If, however, immediate withdrawal of the offending measure is “impracticable,” compensation (in the form of other concessions) can be provided. “The last resort . . . is the possibility of suspending the application of concessions or other obligations under the covered agreements on a discriminatory basis vis-à-vis the other Member.” However, DSU Article 22 emphasizes:

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 is preferred to full implementation of a recommendation to bring a measure into conformity with the covered agreements.

29. The rules for dispute settlement in the WTO build on those enshrined in Article XXIII of the GATT. WTO members typically bring complaints in one of two forms—either they allege that a specific WTO discipline or negotiated commitment has been violated or they claim that although no specific WTO rule has been violated, a government measure has nullified a previously granted concession (a so-called nonviolation complaint). There is also a catch-all third possibility, a so-called situation complaint, in which a member argues that “any other situation” not captured by the other two categories has led to nullification or impairment of a negotiated benefit.

Box 2.1 Settling disputes

According to the Uruguay Round Understanding on Rules and Procedures Governing the Settlement of Disputes (DSU), the parties to the dispute are first required to engage in consultation and, if they wish, to ask the director-general to mediate in their dispute. If these consultations are unsatisfactory, within 60 days a complainant can request the establishment of a panel to hear the case. The WTO Dispute Settlement Body (DSB) would establish such a panel (drawing on its roster of potential panelists nominated by WTO members). The panel would then examine the case, and after an interim report, issue a final report with conclusions and, if they so decide, with suggestions as to how to comply. If the panel finds that a member has failed to comply, absent an appeal by that member,¹ it could make a recommendation as to how it could comply. If it is impractical to comply immediately, the member would be given "a reasonable period of time to do so" (DSU Article 21:3). The finding could then be appealed. If the member loses the appeal and fails to act within this period, the rules call for the parties to negotiate compensation, "pending full implementation" (DSU Article 22:2). Compensation is, however, "voluntary" (DSU Article 22:1). Moreover, it is generally understood that any compensation provided should be on a most favored nation (MFN) basis.² If after 20 days compensation cannot be agreed, the complainant may request authorization from the DSB to suspend equivalent concessions.³ In particular "the level of the suspension of concessions . . . shall be equivalent to the level of nullification and impairment" (DSU Article 22:4). The DSB determines the magnitude of the retaliation generally on the recommendation of the original panel. Arbitration, to be completed within 60 days, may be sought on the level of suspension and the procedures and principles of retaliation (DSU Article 22:6). Conspicuous by its absence in this approach is a requirement that the violating country provide "compensation" for the trade that has been lost over the period before the resolution of the dispute settlement case. Adjustments have almost always been prospective.⁴

1. Panel reports must be adopted within 60 days, unless a consensus exists not to adopt, or a party appeals the findings. Appeals are limited to issues of law and legal interpretation and are heard by an Appellate Body composed of seven members. Appeals proceedings must be completed within 90 days.

2. DSU Article 22:1 states, "Compensation is voluntary and, if granted, shall be consistent with the covered agreements." This is generally understood to require that it be based on MFN principles. See Anderson (2002).

3. According to DSU Article 22:3, the complaining party should first seek to suspend concessions with respect to the same sector as that in which the panel body has found a violation. If that party considers it is not practicable or effective it may seek to suspend concessions in other sectors under the same agreement (if this is not practicable or effective, then obligations under another covered agreement may be suspended).

4. According to Pauwelyn (2000, 337, footnote 21), "In six of the GATT panel reports that addressed either antidumping or countervailing duties against alleged subsidies (only three of which were adopted), the panel found in one way or the other that duties levied in breach of the rules had to be reimbursed."

Two observations on the suspension of concessions are in order. First, rebalancing restores the “balance of concessions,” but it is a rough measure and does not precisely restore the status quo ante. WTO negotiations are multilateral, but suspending concessions in response to violations is bilateral or at most plurilateral and is therefore discriminatory. (In contrast, compensation is supposed to be provided on an MFN basis.) Whereas reciprocity may be diffuse, retaliation occurs only between principal suppliers to the violating member. A country that retaliates may well rebalance its bilateral trade with the defendant country, but it could well experience new inflows from third parties. Moreover, members are not obligated to respond by raising tariffs on the same products as they originally provide concessions. Instead they are required to consider suspension first in the sector in which the violation occurred,³⁰ second in the agreement (e.g., GATT), and only then in other covered agreements (e.g., the Trade-Related Aspects of Intellectual Property Rights [TRIPs] or the General Agreement on Trade in Services [GATS]).

Second, the WTO also allows countries to suspend concessions under circumstances in which there are no disputes or violations. On these occasions, reciprocity again plays a key role as responses are allowed in order to maintain the balance of concessions. One circumstance occurs under GATT Article XXVIII, which allows countries legally to modify their tariff schedules. If a country withdraws concessions it has previously granted, principal suppliers of the product on which tariffs have been raised are allowed to obtain permission to respond (on an MFN basis) with higher tariffs of their own. Clearly, in this case, no attempt is being made to compel the country to restore the tariffs it has modified. The purpose here is simply to permit the injured suppliers to rebalance their concessions. Yet the magnitude of the response is commensurate with the response to Article XXIII (nullification) violations.³¹

A second circumstance occurs in response to safeguard actions under GATT Article XIX (Emergency Action on Imports of Particular Products). If a member finds that “as a result of unforeseen developments and of the effect of obligations it has incurred” imports are (or threaten to be) “a cause of serious injury,” the member may “to the extent and for such time as may be necessary to prevent or remedy such injury suspend the obligation or modify the concession.” Again, with certain qualifications,

30. DSU Article 22:3 (f) defines “sector” as “all goods” with respect to goods and refers to a list of 11 sectors with respect to services and several categories with respect to trade-related intellectual property rights.

31. Another important difference between rebalancing under GATT Article XXVIII and rebalancing under Article 22 of the DSU is that renegotiations under GATT Article XXVIII, in addition to being done on an MFN basis, are permanent, whereas suspensions under DSU Article 22 (and GATT Article XIX) are expected to be temporary and are undertaken on a discriminatory basis.

suppliers affected by a safeguard have the right either to obtain compensation through MFN tariff reductions in other sectors or to rebalance concessions through raising tariffs discriminatorily or suspending obligations elsewhere.³²

The similarities between the escape clause and GATT Article XXIII are not coincidental. John Jackson (1969) notes that in its 1946 proposal for a draft of the International Trade Organization (ITO) charter, the United States included an article on nullification and impairment. He adds, "This article was placed in close proximity to the escape clause article and discussions at the London session indicated the opinion that there was a strong resemblance [between] the purposes of the escape clause emergency relief from obligations (in cases of serious injury from imports) and the 'nullification and impairment clause' (where expected benefits under the agreement were not materializing)" (Jackson 1969, 167).

In the case of rescheduling both negotiations and safeguards, there is obviously no punitive purpose.³³ It is significant, therefore, that if an agreement is nullified or impaired and not brought into compliance, the response is of a magnitude similar to that when countries legally modify their tariff schedules under Article XXVIII of the GATT and when they implement a safeguard action under Article XIX.³⁴ Responses under GATT Articles XXVIII and XIX are both required to be "substantially equivalent" to the trade opportunities lost in response to rescheduling or safeguard actions, respectively. Responses under DSU Article 22 are required to be "equivalent" to the level of nullification and impairment.³⁵

32. As a result of the Uruguay Round Agreement on Safeguards, compensation is not required, nor a suspension of concessions allowed if the safeguard is limited to three years and if the injury results from an absolute increase in the quantity of imports.

33. Standing to retaliate also differs in these cases. For actions taken under DSU Article 22 eligibility is confined to parties that bring the case. Any substantial supplier (typically a supplier with a 10 percent share) may retaliate in response to an Article XIX safeguard. Principal suppliers, substantial suppliers, and the member that originally negotiated the concession may participate in Article XXVIII renegotiations.

34. DSU Article 22:4 states that "the level of the suspension of the concessions or other obligations authorized by the DSB shall be equivalent to the level of the nullification or impairment." Article XXVIII:4 (d) of GATT states that if a contracting party withdraws a concession, other parties can withdraw "substantially equivalent" concessions initially negotiated with that contracting party. Article XIX:3 of the GATT talks of allowing affected parties to suspend "substantially equivalent concessions or other obligations."

35. In practice, responses to renegotiations and safeguards are undertaken on a duties-collected basis—i.e., the increase in tariff times the average trade in that item for the last three years. In contrast, retaliation in response to nullification and impairment could entail more sophisticated economic reasoning. I am indebted to Amy Porges for clarifying this for me.

WTO Dispute Settlement Remedies: A Deconstruction

When parties breach a contract, they are generally required to provide compensation. This requirement actually accomplishes three tasks *simultaneously*. It provides an incentive for compliance by the promisor; it provides compensation to the promisee in the event of breach; and it also permits breach by the promisor if circumstances warrant. How well, then, do the remedies provided for in the WTO system meet the goals that participants generally ask of it? In particular, how successful is it as a means for inducing compliance, providing compensation, affording an escape clause, and maintaining reciprocity?

All four of these goals may be desirable in practice, and any system that could simultaneously achieve them all would be particularly successful. However, to hit four targets, one generally needs an equal number of instruments. If one has just one instrument, unless the targets just happen to be aligned in a particular way, some compromises may have to be made. Similarly, some trade-offs may have to be made if one mechanism—rebalancing concessions—is used to accomplish all these tasks. In addition, the instruments that are practicable and available in an international setting are different from those that might be feasible domestically. In particular, parties that are sovereign governments face some constraints—WTO rulings have no direct legal effect on domestic laws, and force cannot be used to ensure remedies are implemented.

It should not be surprising, therefore, that several of these goals are not completely achieved. Nonetheless, rebalancing concessions in response to violations helps improve compliance, mitigate some of the costs that violators impose on other countries, provide a legal mechanism to permit at least temporary breach of WTO rules, and ensure that reciprocity is maintained. But it does not compel (or ensure) compliance. Nor does it fully compensate other countries for their losses. Its role as an escape clause remains controversial, and it maintains reciprocity fairly imprecisely and sometimes only with considerable delay.

Compliance

Economic theory suggests that rebalancing through the suspension of concessions eliminates the encouragement the WTO system might otherwise provide to large countries to engage in violations that improved their terms of trade. But rebalancing does not necessarily *punish* violations, and theory also predicts that if the internal benefits outweigh the costs, countries will violate the agreement, a practice WTO remedies do not necessarily prevent.³⁶

Suppose, for example, a tenant rents an apartment on a month-to-month basis (without making a security deposit) and then fails to keep up on the rent payments. The landlord responds by evicting the tenant. Has the tenant been subject to a *punitive* sanction? No, the landlord and tenant have both taken actions with equivalent value. It's simply that the arrangement has collapsed, and the landlord has been able to restore the status quo ante. To be sure, the tenant is more likely to pay the rent knowing that eviction is an option; the threat of eviction induces compliance. But eviction is clearly not a surefire method of ensuring compliance, and the tenant may prefer eviction to paying the rent.

The WTO's "suspension of concessions" operates in a similar fashion. It has the effect of restoring the balance of concessions that existed before the adoption of the rule (or agreement) that has been nullified. It does not entail punishment, particularly when trade flows of similar value are assumed to be equivalent. Indeed, the WTO never speaks of sanctions or penalties. It describes retaliation simply as the "suspension of concessions." Moreover, it restricts the value of trade eliminated by suspension to the value of trade nullified by the violation, an amount implicitly equal to the trade value of concessions initially provided by the plaintiff.

In calculating whether or not to defect in a system without any retaliation at all, a country might only weigh the benefits and costs of defection *to itself*. One of the costs is the impact of the higher trade barriers on the efficiency of resource allocation. Another is the damage that defection will inflict on the country's reputation—which could make other WTO members less likely to provide it with concessions in the future and which could do more damage to its international relations generally. Included in the benefits would be the perceived (political) advantages of raising trade barriers and shifting resource allocation toward favored domestic constituents.³⁷ In addition, however, on the benefit side, a large country could also see an improvement in its terms of trade. This would provide *an added incentive* to commit the violation. Therefore, if suspending concessions

36. In his study of the preparatory work on GATT Articles XXII and XXIII, John Jackson (1969, 170) notes, "It was clear that the draftsmen had in mind that Article XXIII would play an important role in obtaining compliance with the GATT obligations. The customary international law analogy of *retorsion* was used but throughout the various drafting sessions, there seems to have been some conflict as to the nature of the role that Article XXIII should play, particularly with respect to whether the 'suspension' provisions should be limited to the equivalence of the damage done by the action of an offending state or whether they authorized more extensive suspensions in the nature of a sanction. . . . In some statements, draftsmen suggested the need for 'more rigorous retorsion' if the offending action is 'abusive.' Yet other statements were made to the effect that measures under the article were not 'sanctions' and were not 'punitive.'"

37. Countries could have noneconomic reasons for preferring more production (or less consumption) of a particular product. See Johnson (1965).

were not allowed, the system could reward violators at the expense of members that adhered to their commitments and this could actually encourage violations. To offset such a loss in its terms of trade, the plaintiff member would have to respond with tariff hikes of its own in a manner that reduced its imports by the same volume as its export loss valued at original world prices. This is precisely the amount of retaliation that WTO rebalancing allows. As Bagwell and Staiger (1997, 3) show, "GATT's insistence on reciprocity in this circumstance can guide governments to efficient politically optimal outcomes, since by neutralizing the world-price effects of a government's decision to raise tariffs, reciprocity eliminates the externality that causes governments to make inefficient trade policy choices."

While they may not actually be encouraging defection, the absence of punitive responses renders WTO remedies rather weak instruments for compelling compliance. Indeed, in this regard several trade theorists have come to quite negative conclusions about the role of the dispute settlement system. In models typically used in this literature, low tariff levels can result when tariffs are determined noncooperatively, provided that participants fear that tariff increases will result in retaliatory responses that lead to lower payoffs.³⁸ In these models, introducing a dispute settlement system that limits and delays punitive responses makes potential retaliation less likely and thus leads to *less* liberalization (Ludema 1990). As Wilfred Ethier (2001a, 12) observes, "Trade theory focuses virtually exclusively on punishments as trigger strategies to sustain agreements for significant liberalization. [Since it does not allow for punishment] actual practice is inexplicable from this view."

One definition of an economist is "someone who sees something working in practice and wants to know if it works in theory." In general, nations seem on the whole to comply with WTO rules (Hudec 2002), but the absence of more punitive retaliatory responses in the WTO puzzles theorists. Dan Kovenock and Marie Thursby (1992) have posited an additional "sense of international obligation" to explain why widespread defections are not a feature of the system. They demonstrate that if the dispute settlement system increases the value of such an "obligation cost," it could serve to enhance compliance and encourage the negotiation of lower tariffs. But these benefits could be obtained from a dispute settlement system that simply rendered public judgments and without retaliation at all.

All in all, therefore, it appears that WTO remedies are likely to make some modest contribution to compliance.³⁹ When combined with the fact that in practice most members are unlikely to retaliate and that in addition an extensive period could pass before retaliation, the system ap-

38. For an excellent survey of this literature, see Staiger (1995).

39. See Bown (forthcoming) and Bown (2003) for evidence that concerns about retaliation under the GATT/WTO did affect behavior.

appears particularly poorly designed for this purpose. From some viewpoints, however, this is not necessarily a bad feature of the system. In particular, it implies that countries generally will comply only when they believe that compliance is in their interest. This level of compliance is actually quite healthy for a system in which sovereign nations are meant to cooperate. (See box 2.2.)

Steve Charnovitz (2001) has claimed that the WTO has become more punitive than GATT. Charnovitz recognizes that the GATT dispute settlement system was based on a “rebalancing paradigm” designed to maintain a balance of concessions. However, he claims that the legal provisions of the Uruguay Round moved away from this paradigm and established a system based on “sanctions intended to achieve compliance.” Charnovitz provides an impressive amount of evidence that policymakers widely perceive the system to endorse (punitive) sanctions. But his efforts to find actual textual support in the DSU and other appeals body statements for his claim that the purpose has been changed are less convincing.

While there are some subtle differences in language between GATT Article XXIII and the DSU, which Charnovitz interprets to indicate that the intent of suspension has been altered, he offers no evidence that larger suspensions would now be authorized under DSU Article 22. Yet, regardless of the language used, suspension has always provided *both* an incentive for compliance and a mechanism for maintaining reciprocity through rebalancing. The dual function of the system as facilitating rebalancing and inducing compliance has long been recognized. As Kenneth Dam noted in his study published in 1970, “The essence of the GATT system lies not in the abstract legal relationships created by a tariff concession but rather in the enforcement mechanism.” Charnovitz reads great significance into the statement by the arbitrators in the bananas case: “We agree with the United States that the temporary nature [of countermeasures] indicates that it is the purpose of countermeasures to induce compliance.” However, it then goes on to state, “But this purpose does not mean that the DSB should grant authorization to suspend concessions beyond what is equivalent to the level of nullification and impairment. In our view, there is nothing in the [DSU] that could be read as justification for countermeasures of a punitive nature.”

In fact, authorized retaliation under the DSU could well be *less* punitive in the new system than it might have been under GATT. With the important exception of the way in which prohibited export subsidies are handled,⁴⁰ the magnitude of the authorizations to retaliate in response to violations is actually now more tightly constrained in the Uruguay Round rules than it was in GATT. The DSU eliminated any possibilities that may have existed under GATT to allow the response to the suspension of concessions to exceed “substantially equivalent measures.” The original

40. For a more complete discussion of this exception, see chapter 3.

Box 2.2 Prisoner's dilemma in trade negotiations

Trade negotiations are sometimes modeled in game theory as a prisoner's dilemma.¹ The essence of this game is the ranking of payoffs in a game between two countries. Let C denote the home country cooperating, C^* the foreign country cooperating, D the home country defecting, and D^* the foreign country defecting. The home country's ranking would be $DC^* > CC^* > DD^* > CD^*$. The foreign country's ranking is the opposite and symmetrical. Each side would prefer to defect, provided the other side complies (DC^*). One way of justifying this is that such actions lead to improvements in the terms of trade. But mutual cooperation (CC^*) is preferred to mutual defection (DD^*). Worst of all is compliance when the other party defects (CD^*).

If the game is only played once, both parties will defect. This way each avoids its least desirable outcome. But if the game is played repeatedly, cooperation may become more attractive. Players will factor in the effect of defection today in the likelihood that their partners will defect tomorrow. The more the future is taken into account, *and the more costly future defection* is likely to be, the greater the likelihood of cooperation today.

Robert Axelrod (1984) has conducted several contests in which players engaged in such multiperiod prisoner's dilemma games. It is interesting that he found those that played tit for tat were generally the most successful in achieving the most cooperative solutions over time. If each party can be assured that if it complies, the other party will do likewise, then by complying it will raise the likelihood of mutual compliance—an outcome it prefers to mutual defection. Conversely, parties will be discouraged from defecting if they realize that defecting raises the possibility that their trading partners will defect in the future. It is striking, therefore, that a system with rebalancing is in fact institutionalizing responses that are tit for tat. WTO members do indeed interact continuously. There is reason to believe therefore that to the degree the prisoner's dilemma model accurately captures national preferences, rebalancing will similarly lead to cooperation over time.

But the prisoner's dilemma assumes parties rank outcomes in a particular manner. In particular, they value mutual cooperation (both do not confess) above mutual defection (both confess). However, this may not always be the case. In some circumstances, parties may prefer mutual defection to cooperation. For example, it could well be the case that European governments prefer to ban hormone-treated beef even if the ban subjects them to US retaliation. Under these circumstances, the United States cannot compel compliance by simply withdrawing equivalent concessions. More punitive responses are more likely to overcome such a deadlock. But they need not. In any case, they could only do so at the risk of undermining reciprocity.

1. The name stems from a game in which two prisoners are each given the choice to confess or not to confess. If neither confesses, they will both be convicted, say for two years each. However, if only one confesses, he will obtain a lighter sentence of one year, and the other receives four years. If both confess, their sentences will both be long, say three years. If the game is only played once and they are not allowed to collude, the optimal strategy is to confess. If the other prisoner confesses, the first does better to confess (three years) than not to confess (four years). If the other does not confess, the first still does better to confess (one year) than not to confess (two years). Since both think alike they both confess and do worse than if they both did not.

GATT Article XXIII that covered nullification and impairment used more expansive language. In response to nullification violations, GATT Article XXIII spoke of allowing such suspensions of concessions as the GATT Contracting Parties “determine to be appropriate in the circumstances.” In contrast, Article XXVIII (Rescheduling of Concessions) spoke of providing “substantially equivalent concessions” in response to tariff re-scheduling. These differences in wording led the legal adviser to the director-general of GATT to argue that in responses under Article XXIII “other factors could be taken into account.” Implicitly, therefore, the response could be *greater* than equivalent to the amount of trade subject to nullification and impairment by the violation.⁴¹ However, the DSU eliminated this discrepancy and made it clear that retaliation can be no larger than the nullification and impairment resulting from the violation.⁴² Specifically, Article 22:4 of the DSU states that “the level of suspension shall be equivalent to the level of nullification and impairment.” In addition, while retaliation was constrained in the GATT, no mechanism existed for subjecting the size of such retaliation to review.⁴³

The DSU also established a provision allowing for binding arbitration when retaliation through suspending concessions was authorized. In particular, if a member objects to the level of the suspension of concessions, or the principles and procedures by which they have been determined, the agreement calls for binding arbitration within 60 days, either by the original panel if available or an arbitrator appointed by the WTO director-general.⁴⁴ By affording defendants the ability to argue for limitations on retaliation, the DSU raised the possibility that these could be smaller than would have occurred under GATT.⁴⁵

In sum, it is hard to square this tightening of the system in the Uruguay Round with the idea that the intention was to have a system more oriented toward inducing compliance. Finally, it is instructive to recall that under the DSU, the same procedures (of compensation and suspension of concessions) are applied to cases where members violate an agreement as

41. For a more complete discussion, see *GATT Analytical Index: Guide to GATT Law and Practice*, 696–700.

42. According to Friedl Weiss (2000, 389), “[The ITO Charter originally made] clear that retaliation was not to exceed the amount needed for the harm done.” He observes that “the DSU text thus returns to the final intentions of the GATT/ITO negotiators.”

The arbitrators in the *EC-Bananas* case emphasize that the equivalence standard required by DSU Article 22:7 is stricter than under GATT (paras 6.3–5).

43. Chad P. Bown (2002) analyzes the implications of giving countries more scope for responses in retaliation to violations than to safeguards.

44. DSU Articles 22:6 and 7.

45. Indeed, in the *Canada-Aircraft II* case (Article 22:6), Brazil requested authorization for countermeasures against Canadian aircraft export subsidies valued at \$3.36 billion. The arbitrator eventually authorized \$247.9 million.

when they take measures that nullify and impair benefits even when such measures do not violate or conflict with any provisions of a covered agreement.⁴⁶ When nullification and impairment are found to result under “nonviolation complaints,” the DSU makes it clear that a member is not obligated to remove the offending measure.⁴⁷ However, members *experiencing* nullification and impairment under violation complaints are entitled to respond in the same way under violation complaints as they do under nonviolation complaints.⁴⁸ This suggests that in violation complaints, responses are not designed to induce compliance.

Compensation?

The ability to retaliate in response to a violation by suspending concessions could compensate a country for losses brought about by the impact of the violation in worsening its terms of trade. However, offsetting this benefit would be the losses in efficiency inflicted by the higher tariffs. While a large plaintiff can therefore offset its external (terms-of-trade) losses, *it cannot recoup the internal efficiency benefits it would have enjoyed had there been no violation.* In other words, it is not fully compensated for the effects of the violation, and in effect the best it can do is to return to the status quo ante the original agreement. Since small countries generally face terms of trade that are fixed, retaliation may provide them with no compensation at all. Responding with increased tariffs of their own will lead to less efficient levels of production and consumption and no offsetting terms-of-trade improvements. Such countries will be unable to receive any compensation at all.

Warren Schwartz and Alan Sykes (2002) have undertaken an interesting analysis of the WTO system using the economic theory of contract remedies. They note that in an efficient enforcement system

a party will be induced 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.

One mechanism for doing this is the award of what are termed “expectation damages” that place the promisee in as good a position as it would have been if the promisor had performed (Schwartz and Sykes

46. For an elaboration of this distinction, see Hoekman and Kostecki (2001, 75).

47. DSU Article 26:1 (b) states, “Where a measure has been found to nullify or impair benefits under, or impede the attainment of objectives, of the relevant covered agreement without violation thereof, there is no obligation to withdraw the measure.”

48. Article 26:1 notes that in cases of nonviolation “the procedures in this Understanding shall apply.”

2002).⁴⁹ If the promisor is obliged to make the promisee no worse off than she would have been had the contract been fulfilled and can fulfill this obligation even while breaking the contract, then breach is efficient.

Schwartz and Sykes (2002) claim that the WTO provisions respecting the settlement of disputes over breach of obligations are carefully designed to facilitate efficient social adjustments to unanticipated circumstances. But the WTO system is not actually efficient in this sense. By allowing victorious plaintiff countries to rebalance their concessions through suspension of concessions, the WTO does not actually provide them with “expectation damages.”⁵⁰ As already noted, the DSU states (Article 22:4), “The suspension of concessions should be equivalent to the level of nullification and impairment.” To be sure, taken at face value, this language is compatible with an interpretation that suggests full compensation for the plaintiff, but that does not happen in practice when the statement is taken to refer to trade volumes rather than welfare. Rebalancing trade volumes leads to equivalence only in a mercantilist sense of allowing the plaintiff to eliminate the “bad” of imports. If politicians are concerned only about the trade balance, they may feel compensated, but in terms of economic welfare, only (external) harm done to the plaintiff country due to the decline in its terms of trade can be taken care of by suspension. If it retaliates, the plaintiff country will not be compensated for the (internal) efficiency losses that will result from its higher tariffs. Instead of being in the position it would have been had the agreement been implemented, the plaintiff will find itself in the position it was before the negotiation that resulted in the agreement.

It is therefore not certain that the outcome is efficient in the sense that the benefits to the defendant exceed the costs to the plaintiff. The basic problem with drawing the analogy between the WTO and a system of contracts is that with contracts, payments for damages are assumed to involve “lump sum” financial payments that have no additional impact on resource allocation. By contrast, tariff changes do.

The analogy also fails to capture how tariffs actually work. The more money a plaintiff receives in compensation, the better off she will be. But any increase in tariff beyond optimal levels could actually make the plaintiff worse off. There is an inherent limit therefore to the amount of compensation that can be provided through higher tariffs alone. If the WTO were to institute an efficient system based on “expectation damages” that operated through suspension, in addition to permitting the plaintiff to

49. They state that “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 the breach is inefficient.”

50. Actually it entails restitution—i.e., the equivalence of getting the promisee’s money back in the event there is a breach.

retaliate, the system would have to require additional compensation equal to the efficiency costs the plaintiff incurs when retaliating by raising its border barriers. In such a system, the defendant would supplement such action with financial considerations or some other side-payment.⁵¹ Alternatively, if the WTO did not allow for suspension of concessions and could compel defendants to compensate plaintiffs with reductions in other barriers to their exports, the system could come closer to achieving expectation damages.

It is also noteworthy that the rebalancing the WTO allows is *prospective*. Plaintiffs receive no consideration for the losses due to violations that they may have suffered before authorization. The obligation violated by the foreign sales corporation (FSC) came into effect in 1995, yet Europe was authorized to retaliate only in 2002. Clearly, this practice could render the compensation seriously inadequate. What explains this provision? It could be the desire to attract participants to make agreements in the first place. Indeed, the best way to think about this WTO approach is analogous to a money-back guarantee. "Buy our product," proclaims the advertisement, "and if you're not satisfied, you can return it at any time and you will get your money back." Suppose you buy the product, find its performance wanting, return it, and get your money back. You have been compensated in the sense that you are back to the status quo ante. But you are not left in as good a position as you would have been had the product actually worked, and you are not compensated for the transaction costs, inconvenience, and forgone interest you might have earned on your money. Similarly, when a country is given permission to suspend concessions in response to nullification of an agreement, it is able to restore the status quo before the agreement, but it is not left in as good a position as it would have been had the agreement been fulfilled.

This analogy also points to a consideration that could detract from the inducement to sign an initial agreement. If transaction costs are significant, you will be less impressed by a simple money-back guarantee. If your trading partner is unreliable, and these costs are significant, you will be less likely to sign in the first place. Therefore, when these costs are large, compensation will have to be larger.

Let's illustrate some of these points with the example of the United States and Poland. Assume that initially the United States and Poland each set their tariff levels independently. (Assume both are large countries.) Taking American tariffs as given, Poland did its best when setting its computer tariffs, and the United States likewise set its tariffs on chemicals at their optimal levels, taking Polish tariffs as given. At that time, each

51. If the system required punitive tariffs in retaliation, it could force the defendant to take into account the additional efficiency losses the plaintiff incurred. However, such tariffs would not be optimal from the standpoint of the plaintiff.

was prevented from reducing its tariffs further by the adverse effect on its terms of trade. By negotiating a reciprocal agreement, however, they managed to eliminate this effect, and the United States agreed to reduce its tariffs on chemicals in return for a reduction of Polish tariffs on computers. They were now able to move to a new situation in which both were made better off.

However, assume circumstances changed for the United States (or it misunderstood the terms of the original agreement), and it chose to nullify the agreement, in effect raising its protective barriers back to their original levels. Assume that after winning its case at the WTO and being granted authorization, Poland retaliates. Trade barriers have now reverted to the original situation. The United States persists with the violation. Apparently, circumstances have changed, and the United States no longer wants the deal.⁵² By revealed preference, therefore, the United States is better off without an agreement. If circumstances have not changed for Poland, however, it is now worse off than it was with the agreement. Poland has clearly not been fully compensated for the breakdown in the agreement, but its retaliation has at least allowed it to offset the worsening in its terms of trade that resulted from the US violation. Poland is no worse off, though, than had there been no agreement in the first place.

Strictly speaking, therefore, allowing for suspension removes the additional incentive that the terms-of-trade improvement would otherwise provide for violations. It also allows Poland to offset the loss it would otherwise experience from the decline in its terms of trade due to the violation. But Poland is not as well off as it would have been had the United States stuck to the agreement. Therefore rebalancing through suspension expunges the impact of the agreement and allows the system to revert to the status quo ante. Benefits are “rebalanced” in the sense that neither side benefits from the agreement.

In sum, breaches that occur through this mechanism will not necessarily be socially (or globally) efficient, since the gains to the violator could be less than the costs to the plaintiff. Nonetheless, setting compensation at a relatively low level could have the desirable effect of encouraging those members that fear that they may have to breach, to sign agreements in the first place. However, members who believe that breach by the other side is likely may be less willing to sign than if they believed the other side would suffer more serious consequences. In particular, if the transaction costs of signing agreements are substantial, the failure to provide adequate compensation could also serve to discourage agreements by some members.

52. In this case, the original Nash equilibrium is also Pareto optimal and the best deal that can be sustained without making one of the parties worse off.

Safety Valve?

WTO remedies have weak effects in inducing compliance and do not require full compensation of loss for the plaintiff. Paradoxically, however, these features actually enhance the system's ability to provide a safety-valve mechanism.

Safety valves are useful. Modification of Schedules (under GATT Article XXVIII) and escape clause actions under GATT Article XIX (Emergency Action on Imports) are generally viewed as desirable attributes of the WTO system. Both are designed to assuage the legitimate fears of members that circumstances could change after they have signed an agreement. Rescheduling (GATT Article XXVIII) permits members permanently to opt out of tariff concessions that they no longer view to be in their interests. Escape clause actions (Article XIX) permit a member to obtain a temporary respite from its obligations in the face of unforeseen developments. By providing these mechanisms, the system encourages members to make concessions they would not otherwise make and to avoid commitments that no longer serve their interests. Yet similar provisions for "escaping" from other types of WTO obligations appear to be conspicuously absent. For example, what if a country finds its adherence to the provisions of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs) unexpectedly causes serious injury? What if a member is genuinely surprised to find that its understanding of a particular provision is different from the ruling given by the WTO DSB? Under these circumstances, a member may find that its preferred option is to violate the agreement and deal with those who complain through the dispute settlement system.

It has already been noted that the suspension of concessions allowed by the WTO in response to violations is substantially similar to that allowed in response to escape clause actions and rescheduling under Article XXVIII. The resemblance between suspensions under (DSU) Article 22 and (GATT) Article XIX is particularly strong, since both are expected to be temporary. This suggests that de facto, accepting retaliation in response to findings of violations may provide an escape clause route for circumstances not covered by Article XIX. More controversial is the question of whether accepting such retaliation could serve as a more permanent opt-out mechanism similar to Article XXVIII.

Legal scholars are divided over the precise nature of the obligation to comply with the rulings of the DSB. The DSU language clearly states that compliance is "preferred." It also states that rebalancing through either compensation or suspension of concessions is meant to be temporary. Unlike tariff reschedulings, the suspension of concessions under Article 22 of the DSU is undertaken bilaterally rather than on an MFN basis, reinforcing the idea that this departure from a basic WTO norm is a *temporary* measure. In addition, provisions in DSU Article 22 require

the DSB to keep such cases under surveillance. According to John Jackson (1997), ultimately the obligation to comply is the same as any other binding obligation under international law.⁵³ Scholars such as Judith Bello, Warren Schwartz, and Alan Sykes provide an alternative view (Bello 1996, Schwartz and Sykes 2002). They place more significance in the phrase “compliance is preferred” and argue that it is significant that the phrase “compliance is required” is never used. They also emphasize that although rebalancing is meant to be temporary, no specific time limits are placed in its exercise.

As an economist who lacks legal training, it lies beyond my competence to choose between the eminent legal authorities on this issue. However, I would note that the practical significance of the argument about an ultimate requirement to comply is relatively small. Although the application of remedies is “intended to be temporary,” the practical effect of not providing for a firm date for compliance could be to allow the defendant legally to sustain the practice indefinitely.⁵⁴

The potential benefit of such a mechanism for indefinite breach has increased as the dispute settlement system has become increasingly powerful and the scope of WTO rules increasingly broad. It is no surprise that the agreements contain gaps, ambiguities, and inconsistencies. Yet it is not easy to change them. Negotiating new agreements has become increasingly time consuming. Even if the Doha Round ends as scheduled in 2005, for example, 12 years would have passed since the end of the Uruguay Round. The DSB, by contrast, now acts fairly rapidly. Its cases must meet time guidelines, and members cannot prevent the appeals body from issuing binding rulings.

In a recent study, Claude Barfield (2001) argues that a mismatch now inevitably produces excessive judicial activism. “The internal flaw is between the WTO’s consensus-plagued inefficient rule-making procedures and its highly efficient dispute settlement system—an imbalance that creates pressure to ‘legislate’ new rules through adjudication.”⁵⁵ The ability to reverse rulings, which in domestic settings is provided to the legislature, is not effectively present in the WTO. Barfield is concerned, therefore, that the DSB interpretations will override preferences of sovereign

53. DSU Article 22:9 also states, “when the DSB has ruled that a provision of a covered agreement has not been observed, the responsible Member shall take such reasonable measures as may be available to it to ensure its observance.” This seems to reinforce the principle of a legal obligation to comply. But it also states, “the provisions . . . relating to compensation and suspension of concessions or other obligations apply in cases where it has not been possible to secure such observance.”

54. To be sure, the DSB does monitor uncorrected violations permanently.

55. Article IX:2 of the WTO agreement states that “only the Ministerial Conference and General Council are empowered to adopt definitive interpretations of GATT-WTO texts.”

national governments. The result of this imbalance will inevitably be an erosion of the WTO's legitimacy.

Barfield cites five cases to bolster his arguments. While his judgments about whether panels have exceeded their mandates in individual cases can be questioned, what is indisputable is that many WTO cases are extremely controversial. Passions are bound to be inflamed when the WTO makes rulings on hot topics such as the relationship between the trade rules and the environment (tuna-dolphin and shrimp-turtle), food safety (e.g., hormone-treated beef), nontransparent barriers (Kodak-Fuji), US antidumping provisions (see Leibowitz 2001), tax subsidies (FSC/ Extraterritorial Income Exclusion Act), and now genetically modified organisms.

How then should the WTO respond? One approach would be to weaken the dispute settlement system. Indeed, Barfield's solution to this problem is to create mechanisms for short-circuiting and/or blocking the dispute settlement process. Specifically, he advocates two changes to create "a less rigid, more flexible dispute settlement system" (Barfield 2001, 13).

But it is by no means clear that the system would be improved by adopting Barfield's recommendations. They could actually prove counterproductive. Imagine the cries of foul if, as Barfield advocates, the director-general of the WTO (or a subset of WTO members) unilaterally made the decision that a dispute was too controversial or political, or that the language agreed upon by a consensus of WTO members was too vague for the DSB to render a judgment. One can well imagine those who believe they have the upper hand in the dispute complaining vehemently that their rights were being overruled by an unrepresentative group of members or even worse a bureaucrat like the director-general of the WTO. Likewise, it is hard to imagine that a victorious litigant would look favorably at his second recommendation—a blocking minority of members overruling a decision. Adopting these suggestions might not only *reduce* the WTO's legitimacy but also impair its efficacy.

Moreover, the proposition that it is better to have agreements that are never authoritatively interpreted is highly questionable. Panelists perform invaluable services both in providing interpretations and in signaling where clarification and amendment are called for. If countries could no longer be sure that the dispute settlement system would be able to make rulings on violations, they would be less likely to rely on it. This lack of faith could produce less liberalization. It would also lead to an increase in extra-WTO retaliation and thus reverse one of the great accomplishments of the Uruguay Round agreement.

The argument here suggests that the WTO already has in place a mechanism to deal with the concerns raised by Barfield. Countries that lose cases—as with the European Union and hormone-treated beef—need not give up the sovereign right to determine their own laws. If they do not wish to comply with DSB rulings, they can either pay compensation or

accept retaliation. They can then try to have the rule changed or clarified in the next trade round.

De facto, therefore, the responses to violations under the DSU operate as an escape clause, although this is not generally acknowledged. As long as retaliation is restrained, perhaps it is sufficient to simply allow this mechanism to serve tacitly as an escape clause mechanism. However, if responses are increasingly allowed to follow the more punitive approach and reasoning used in the FSC case discussed in chapter 3, there may be pressures for a provision, analogous to Article XXVIII for rescheduling tariffs, that could be explicitly invoked in the event particular rules become untenable. In both cases, other members could respond by rebalancing their concessions.

Rebalancing?

This review of three possible goals of WTO remedies suggests some role for each. However, these attributes appear as by-products of a system that is best explained as an effort to maintain reciprocity through rebalancing. John Jackson (1969), in his study of the preparatory work done for the original GATT and ITO, quoted one of the drafters as stating, "What we have really provided, in the last analysis, is not that retaliation should be invited or sanctions invoked, but that a balance of interests, once established, shall be maintained."⁵⁶ Similarly, in writing about the original GATT agreement and its responses to violations, Robert Hudec (1993, 7) observed, "The key value underlying this rather odd legal design was reciprocity. The legal procedures were not there to enforce obligations for the sake of enforcement. They were there to correct imbalances that might arise in the benefits governments were actually receiving from the agreement. It was a diplomat's concept of legal order."

This approach appears desirable when viewed from a mercantilist perspective. If a country reneges on its promise to provide the "good" of more exports for its trading partner, the partner should either be compensated with other exports or be able to protect itself by reducing the "bad" of its imports from that country by an equal amount.

Rebalancing also accords with the ideas of trade theory. By rebalancing, the plaintiff is able to eliminate any deterioration in its terms of trade that might have resulted from the violation of the agreement. Rebalancing therefore ensures that countries make their trade policy based on the internal consequences of their decisions and denies them the ability to shift their costs onto their trading partners.

Rebalancing also serves some important political functions by maintaining reciprocity. If producers of exports are hurt as a result of foreign

56. Quoted in Charnovitz (2001, 801).

violations, at least import-competing producers can be rewarded by the retaliation. Rebalancing also allows the claimant nation to prevent foreign countries from reaping extra gains from a trade agreement that it violates (an important consideration for realists). Finally, it ensures that the terms of the bargain cannot be abridged unilaterally without consequences. This assurance gives agreements greater credibility. All these elements are important in maintaining political support for the system.

Agreement and Enforcement: Interdependent Arrangements

It is well recognized in economics that barriers to exit can create barriers to entry. Just as restrictions on firing workers, for example, can discourage firms from hiring workers, so could penalties discourage countries from agreeing to WTO disciplines at all. This barrier is important because one role of the WTO system is to entice the parties to sign as many agreements as possible. As Kenneth Dam (1970) noted in his classic study of the GATT,

The GATT has a special interest in seeing that as many agreements for the reduction of tariffs as possible are made. Enforcement of tariff bindings is important . . . but a system that made the withdrawal of concessions impossible would tend to discourage the making of concessions in the first place. It is better, for example, that 100 commitments should be made and that 10 should be withdrawn than that only 50 commitments should be made and all of them kept.

My analysis of WTO remedies suggests that the WTO system operates by offering parties close to a “no loss” or “get your money back” opportunity. If the deal looks beneficial today, you can sign it knowing that at worst, if circumstances change, you cannot be worse off.⁵⁷ After all, you will only defect if you think defection makes you better off. However, if your trading partner defects, you will be able to rebalance concessions and at worst revert to the status quo ante, raising your tariffs to their previous level, which is the best you can do given the other side’s violation. Even in this case, assuming there are no fixed costs associated with signing the agreement (and no delays in enforcing it), you will not be worse off than you would be without an agreement because you will have enjoyed benefits until the violation was committed.

On the other hand, suppose that, instead of simply allowing the plaintiff to return to the status quo ante, the rules allowed it to exact an additional penalty.⁵⁸ This *could* make the breaching party worse off than it

57. This assumes there are no fixed costs associated with signing the agreement.

58. Assume that the original negotiations began with a Nash equilibrium. Both sides were not cooperating and each side set its tariffs at the optimal level given the tariffs in

was before the original agreement. Its terms of trade would end up lower than before the agreement and thus the “externality” from breach would now be negative rather than positive. Assume that before the agreement, the plaintiff had set its tariffs at their optimal level, given the absence of an agreement. If it were then *required* to impose a penalty, it too would be worse off from participating in the agreement. While such penalties might be more effective in inducing compliance, if they failed to do so, *both* parties could be worse off and hence discouraged from entering into the agreement in the first place.

In the discussion of compensation it was noted that once an agreement is signed, the provision allowing retaliation to be equivalent to nullification and impairment does not lead to optimal breach. However, it can be shown that the provision can lead to optimal agreements being signed at the outset. Wilfred Ethier (2001b) has argued persuasively that rebalancing with commensurate responses is an optimal approach when countries negotiating trade agreements are subject to considerable uncertainty about whether or not they could find themselves out of compliance. “Each country knows that it might turn out to be either the accuser or the accused. Thus it is in no country’s interest, *ex ante*, to agree that, *ex post*, either the accuser should be unconstrained in its ability to punish or the accused should be unconstrained in its ability to proceed without punishment” (Ethier 2001b, 5). He argues, therefore, that central to multilateral liberalization is an implicit agreement among countries to violate agreed commitments if the violation implies no retreat from reciprocity.⁵⁹ He also demonstrates rigorously that a system with commensurate responses to violations leads to an optimal degree of liberalization.

In sum, the ability of plaintiffs to suspend concessions in response to WTO violations accomplishes several goals simultaneously. Above all, it permits the plaintiff to maintain reciprocity. In addition, it reduces incentives for noncompliance, it offsets *some* of the losses the plaintiff incurs from the violation, and it may provide a (temporary) safety-valve mechanism that permits breach. The system also encourages members to enter into new agreements—indeed, it will lead to the optimal amount of liberalization (Ethier 2001b).

The failure to appreciate that each of these considerations plays an important role in the system can lead to misplaced criticism and poor

the other country. In this framework, if the violation restored the defendant’s tariff to its original level, in the absence of its WTO commitments, the complainant would respond by raising its tariff back to its original level as well. It would *not* want to respond with an even higher tariff. Thus rebalancing accords with the response the defendant would choose on its own.

59. An important assumption in proving this result is that verdicts are rendered rapidly so no weight is given to the adjudication phase. Lengthening this phase will lead to less liberalization. Ethier (2001b, 20) also derives conditions in which commensurate responses remain optimal.

advice. For example, those who claim that WTO retaliations are sanctions forget where WTO agreements come from in the first place. WTO agreements reflect concessions. When countries are found in violation, they have in principle reneged on something for which they received consideration. When plaintiffs retaliate, they are not applying a punishment but returning to the status quo before the original agreement. Similarly, it is true that smaller countries have greater difficulties in applying retaliation. But in part this is the mirror image of the advantages they could enjoy from free riding in a negotiating system based on the MFN principle.

The mechanisms used by the dispute settlement system will influence what can be negotiated. For example, the WTO implicitly assumes that members have acted in good faith. When findings of violations are made, remedies are always only prospective. This certainly weakens compliance incentives because it implies that there might be no consequences for temporary violations, but it may also help facilitate agreement by giving countries assurances that they will not be punished if they misunderstand the agreement.

The system of negotiation also helps drive compliance. Members comply in part because maintaining a reputation for compliance is important not simply to preserve existing agreements but also to extract concessions in the future.

When critics propose reforming the system, they often fail to consider the impact of their recommendations on the other roles. It is certainly possible, for example, to improve compliance with stronger penalties but this could come at the expense of other goals. Stronger penalties would disturb the balance of reciprocal concessions given by members, inhibit members from exercising breach, and (in mercantilist terms) provide “excessive” compensation. Similarly, more lenient responses might allow members to breach their agreements more easily, but this would again disturb reciprocity and reduce incentives for compliance.

Conclusion

Three sets of preconceptions that observers bring to the WTO were laid out at the start of this chapter. They concerned the legal nature of the system, the purpose of WTO dispute settlement remedies, and the links between the systems for obtaining and enforcing agreements. This discussion of the system’s guiding principles and rules suggests that nuanced judgments of the degree to which the system actually accords with these preconceptions are warranted.

The system has several attributes that resemble a contract. Agreements reflect deals struck by members on the basis of mutual interests rather than global statutes enacted through legislative decisions. Responses to violations are treated as a matter between the affected parties and not as

infractions against the organization as a whole. Members rather than the body as a whole bring cases. If participants in a dispute are able to strike a bilateral deal by offering compensation, which is acceptable to those affected by an infraction, members can maintain their noncompliance without further economic consequences. If they are not able to obtain satisfactory compensation, plaintiffs are not entitled to impose punitive responses and do not even receive expectation damages. Nonetheless, the analogy to a contract remains imperfect. Those who violate a contract may be able to obtain permanent relief from their obligations by providing compensation. However, even when compensation is provided, WTO members cannot permanently evade their (legal) obligations to comply. Ultimately WTO obligations are the same as those of any legally binding international agreement. "Suspension of concessions" or compensation are in principle only temporary measures. However, because no time limits are placed on breach, the infraction could be maintained indefinitely without further consequence. Moreover, members can buy off plaintiffs with compensation before a judgment and avoid incurring such an obligation in the first place.

While WTO remedies achieve several goals simultaneously, the goal achieved most precisely is maintaining reciprocity. The suspension of concessions does operate to create an added inducement for compliance. It may provide larger members with the ability to offset losses in their terms of trade. In practice, by accepting retaliation, a defendant may escape from its obligations, temporarily in principle and indefinitely in practice.

Finally, a key feature of the system is that agreements and enforcement are integrally linked. Conceptually, the principle of reciprocity guides both the manner in which agreements are reached and the remedies that result from violations, although in practice reciprocity is enforced more precisely when concessions are suspended than when they are granted. Operationally, responses to violations reinforce the ability to obtain agreements by providing recourse but avoiding punishment. Conversely, the need to maintain credibility in ongoing negotiations reinforces compliance with dispute settlement rulings.

Nonetheless, while the paradigm of reciprocity fits neatly into a framework in which concessions relate to border barriers, it does less well when relating to rules. In particular, WTO members are obliged to adhere to rules, and violation of rules creates a presumption of "nullification and impairment." However, when faced with violations, other members cannot seek compensation or authorization to retaliate unless they can demonstrate that they have experienced nullification or impairment of benefits—that is, trade effects. Thus the system is actually a strange hybrid in which the response to violations of obligations to adhere to rules that have no impact on trade flows differs from the response to violations that affect trade.