Trade Dispute Diversion: The Economics of Conflicting Dispute Settlement Procedures between Regional Trade Agreements and the WTO

James H. Cassing
Department of Economics
University of Pittsburgh

Abstract

A somewhat underappreciated aspect of the burgeoning rush to regional trade agreements is a discrepancy between the dispute settlement procedure (DSP) embodied in the original WTO Dispute Settlement Understanding and that found in the language of many Regional Trade Agreements (RTA). This paper explores the issue in the context of a dynamic repeated game of tariff agreements. As is well known, the institutional alternatives available in negotiating multilateral freer trade agreements – regional agreements, side agreements, trade dispute settlement punishments, and so on – can prescribe the limits and shape the nature of self-enforcing trade agreements. Here we suggest the extent to which deviations from the WTO DSP embodied in RTAs – e.g., “private interest access,” “third party procedures,” and “choice of forum” – can not only work against the interests of “weaker parties,” but can furthermore undermine multilateral agreements closer to free trade.

Contact Information:

James H. Cassing
Professor of Economics
4908 Wesley W. Posvar Hall
University of Pittsburgh
Pittsburgh PA 15260 USA

e-mail: jcassing@pitt.edu
phone: 412 648 8746
fax: 412 648 1793
Trade Dispute Diversion: The Economics of Conflicting Dispute Settlement Procedures between Regional Trade Agreements and the WTO

1. Introduction

Disagreements between sovereign parties to trade agreements ("trade disputes") are inevitable as economic interests change and commitments require clarification or interpretation. Of course, the attendant clarifications of commitments also require credible mechanisms of enforcement, or "punishment." The World Trade Organization (WTO) foresaw this need and embodied a dispute resolution mechanism or Dispute Settlement Procedure (DSP) in the Dispute Settlement Understanding (DSU) of the original Agreement. Similarly, the recent explosion of regional trade agreements (RTA) creates new avenues for trade disputes and so has itself been accompanied by the incorporation and extension of particular dispute resolution mechanisms. However, a somewhat underappreciated aspect of the burgeoning rush to regional trade agreements is a discrepancy between the dispute settlement procedure embodied in the original DSU and that found in the language of the RTAs. This is true of both United States free trade agreements such as NAFTA or the various bilateral agreements, and European Union Partnership agreements. As it happens, this can matter substantively for trade and welfare as parties attempt to tailor DSP rules in line with particular interests and to divert disputes to their preferred forum.1

This paper explores the issue in the context of a dynamic repeated game of trade agreements. As is well known, the institutional alternatives available in negotiating multilateral freer trade agreements – regional agreements, side agreements, trade dispute settlement punishments, and so on – can proscribe the limits and shape the nature of self-enforcing trade agreements. Here we suggest the extent to which deviations from the WTO DSP embodied in RTAs – e.g., "private interest access," "third party procedures,"

---

1 Woolcock (2006, p.5) has observed that “there is thus nothing new in rules being devised or evolving from a multilevel process. What is new is the emergence of the regional level as one of the most important....This multilevel nature of trade and investment policy has been recognized in policy circles. For example, the OECD Ministerial Declaration of 2001 called for efforts to ensure that regional and multilateral regimes are complements and not substitutes for one another....”
and “choice of forum” – can not only work against the interests of “weaker parties,” but can furthermore undermine multilateral agreements closer to free trade.²

The paper proceeds as follows. In the next section we review briefly the essence of a DSP and its role in shaping and supporting trade agreements. We also highlight some of the key elements in the DSP of the WTO and of some RTAs, emphasizing the discrepancies. Section 3 reviews some of the economics and political science literature on why discrepancies in DSP arise and on how these discrepancies might matter. In section 4 we present a more formal model of the DSP and derive some conditions under which the options available in a regional DSP can have the perverse effect of weakening support for a stronger WTO DSP. Since it is the harshness of the potential punishment that underwrites a more liberal trade regime, we will argue that any such weakening of support will in turn weaken support for a more cooperative freer multilateral trade regime. Indeed, a number of authors have portrayed RTAs as “stumbling blocks rather than building blocks” and lamented the distraction away from the multilateral process. Section 5 discusses nuances and offers some conclusions.

2. The Role of the DSP in Shaping Trade Agreements

The Procedural Mechanisms

On the surface a DSP is needed as part of any trade agreement simply to codify, however loosely, what are the procedures of redress should one of the parties to the agreement be accused of deviation. Also it is recognized that a forum for clarification of obligations will invariably be required on occasion in the future. After all, even a trade agreement as simple as one to lower tariffs still involves numerous definitions, formulae, and

² A number of papers have investigated the impact of RTAs on multilateral cooperation, but the mechanisms driving the impact, e.g. “tariff complementarity,” are very different from the mechanism explored here. See especially Saggi (2006), Saggi and Yildiz (2007), and Maggi (1999) and the references therein for a nice review of this literature along with some interesting new results. For a more “political” model dealing with trade agreements and time inconsistency see Maggi and Rodriguez-Clare (forthcoming) and references to some of their earlier papers.
modalities. In fact, the reach of the WTO agreements goes far beyond that – covering non-tariff barriers and so on – and many RTAs go even further, entering the uncharted waters of labor and environmental concerns, for example, while going well beyond the WTO in dealing with intellectual property rights (IPR).

As a practical matter, when structuring a DSP there is an inevitable tension between a more or less legalistic approach, the latter deferring instead to more reliance on negotiation of disputes. Typically the contrast is made as in Jackson (2000) between the “rule-oriented approach” versus the “power-oriented or diplomacy approach.” Roughly, the WTO leans toward the former and various RTAs toward the latter, presumably because the WTO is a multilateral forum where power is somewhat diffused whereas the most important RTAs tend to be between a large developed country – especially the US or EU – and a smaller less developed country pursuing deeper integration than is offered by the WTO but compelled by the larger country to accept less legalism, or at least legalism more favorable to the larger country such as “TRIPS-plus,” in the DSP (Drahos, 2007).³ Cognizant of the power asymmetry accompanying many RTAs, the less powerful partner has often insisted on a “choice of forum” provision which permits the complainant in a dispute to select either the DSP of the RTA or of the WTO. At the same time, while the less powerful country typically has in mind shifting a trade dispute away from the RTA forum to the WTO, the more powerful country often welcomes “choice of forum” to move the dispute in the other direction. This in effect links a country’s attitude toward the DSP negotiated in the WTO with that negotiated in an RTA.

The WTO DSP follows guidelines laid out in the Understanding on the Rules and Procedures Governing the Settlement of Disputes (“Dispute Settlement Understanding”) which addresses jurisdiction, procedures, and substantive content. (See for example Mavroidis and Sykes (2005) and the references to a vast literature therein.) In sum, the procedure mandates consultation among parties in a dispute; investigation, ruling and

³ Interestingly, it was the US that pressed for the more rule-oriented approach in the WTO while it is at the same time notorious for exerting its superior bargaining position and rule-bending in RTAs. Presumably the US negotiators were well aware of the more limited bargaining power advantage in adjudicating disputes in the more multilateral WTO forum. See Drahos (2007) for a characterization of the US approach to trade negotiations. Also of interest for a historical perspective is Petersmann (1997).
recommendation by a panel; and, potentially authorization for the aggrieved parties to suspend obligations (“retaliate”) against the violator. In practice, retaliation has not been so common, suggesting that the “threat of punishment” has itself been sufficient to deter deviations from the trade agreement – that is, the agreement is “self-enforcing” and the level of freer trade cooperation is “supported” by the punishment threat inherent in the DSP.

The DSP accompanying various RTAs – for example, as outlined in Chapter 20 of the North American Free Trade Agreement (NAFTA) -- is also nominally legalistic and follows the DSU in spirit (WTO, 1995; Gantz, 1999). However, there are some important differences. First, the initiation of a trade complaint is government to government in the WTO; The private sector and non-governmental organizations (NGO) do not have standing to force actions by the WTO. However, according to the rules of the EU, NAFTA and various bilaterals, it happens that unions, business groups, and activists have the ability to press complaints. For example, in the NAFTA, labor unions and environmental groups have initiated complaints, and there are special dispute settlement procedures afforded private parties with investment disputes against governments (NAFTA, supra note 1, chapter 9). Such access to the trade dispute apparatus cannot help but politicize the DSP and give an advantage to the larger, more powerful country’s narrower interests. (See Levy and Srinivasan (1996); Robertson (2006).)

Second, the consultation and adjudication stage forums for a trade dispute have different structures. In the WTO, non-disputant third party countries have access to the process. This enhances the ability of disputant parties to build coalitions in pressing their cases. Also, only third parties – representatives of three non-disputant countries -- comprise the panel if that stage is reached in the dispute. And, while there is a right of appeal, the rulings and recommendations of the panel cannot be vetoed by either disputant as in the pre-1995 GATT. If an offending trade action persists, retaliation (suspended obligations)

---

4 See Smith (2000) for an integrated summary of the key features of most of the post-1957 RTAs. His Table 3 (p. 156-7) indicates levels of legalism in dispute settlement design using the characteristics of third-party review, third-party ruling, judges, standing, remedy, and implied level of legalism.
is proscribed within the context of the DSU and of Article XXVIII (“Compensation”). In contrast, for most RTAs, non-members are not privy to the consultations even if they have important interests. Furthermore, as in NAFTA Chapter 19 (“Antidumping and Countervailing Duty Matters”), national rather than regional or international law is sometimes given precedence. Furthermore, the variety of mechanisms for resolving disputes can itself be confusing and open to interpretation, which admits a far larger role for the “power-oriented approach” over the “rules-oriented approach” in resolving trade disputes (Gantz, 1999). As for arbitration, the analogue of the WTO panel in an RTA is typically comprised of representatives of the RTA member countries which, as with NAFTA’s “Free Trade Commission” consisting of the trade ministers of the three NAFTA countries, issues and recommends implementation of an arbitral report. Thus, the adjudication of disputes potentially gives more weight to self-interest and power politics. Once again, the DSP is likely biased in favor of the larger, more powerful partner to the regional agreement.

Finally, there is the option in most RTAs -- including NAFTA, the US bilaterals, and the EU Partnerships -- of “choice of forum.” While a complainant in an RTA has access to the regional DSP, there is also the option of taking the complaint to the WTO where appropriate. (Some RTAs limit choice of forum on some issues; E. g. NAFTA requires that certain environmental disputes listed in Article 104 be settled under the NAFTA’s Chapter 20 rather than under the GATT/WTO mechanism. For more detail on the legal requirements controlling choice of forum, see Gantz (1999).) In light of the differences in the procedural structure discussed above, this is an important option since third party access, coalition building, panel composition, and so on are all affected by the choice. Roughly, in a trade dispute, a weaker bargaining power country would prefer the WTO forum while the country with more bargaining power would prefer the regional forum where that power cannot be diluted (Drahos, 2007). For example, in a case involving agricultural products tariffication (U.S. v. Can.), the US chose the NAFTA forum essentially because it would have been difficult to argue that Canada had violated the WTO Agreement on Agriculture even though it was Canada’s implementation of its
obligations under the WTO that led to the charges of NAFTA violations in the first place (Gantz, 1999).

Suggestively at least, these discrepancies between the multilateral and regional DSP are substantive in that they influence complainant access to the resolution mechanism, potential for more or less coalition formation, and discretion as to which forum to exploit. Consequently, nations have an interest in DSP design. And, although multilateral and regional trade negotiations occur in distinct and separate policy venues, the character of the multilateral (WTO) DSP sometimes has force in both venues due to “choice of forum.” As we next discuss, this linkage is important because it is the DSP that shapes how much trade cooperation can be achieved. Therefore, more powerful nations potentially pursuing future regional trade agreements and hoping to bolster the use of the more favorable (for them) regional DSP are given an incentive to weaken the multilateral DSP which in turn weakens the extent of possible cooperation within the WTO. That is, what a country agrees to regarding the WTO DSP could potentially erode its bargaining power in an RTA and so have an effect on what can be achieved in regional trade negotiations.5

We now turn to review informally the logic of recent economic (game theoretic) explanations of the DSP as an institution which in fact shapes and limits cooperative trade agreements. Roughly, the notion is that a DSP creates institutional friction which dissipates the incentives for countries simply first to break trade agreements and then, when detected, just quickly renegotiate back to where they were before the violation. But then there would be no “punishment” to deter deviation from the agreement in the first place – and so no agreement. In subsequent sections we will formalize this notion and introduce conflicting regional and multilateral DSPs.

5 Woolcock (2006, p.3) notes that incompatible rules at the regional and multilateral levels can be problematic and observes that “difficulties may arise…when regional or bilateral enforcement results in interpretations of general norms that diverge from those of the WTO or other global bodies.” See also Woolcock (1996) for a discussion of rules competition.
The Self-Enforcement Underpinnings

While on the surface the DSP is a quasi-legalistic mechanism to clarify, adjudicate, and reconcile trade disputes, economists have recently focused on the linchpin role played by a DSP in supporting credible cooperative trade agreements (Ludema, 2001; Bagwell and Staiger, 2002; Klimenko et al., 2008).\(^6\) From this perspective, the world is portrayed as a collection of sovereign nations which can gain from cross-border transactions – trade and investment. Thus there is an incentive for countries to cooperate in order to create an economic environment absent trade barriers or other commercial impediments and so conducive to the globally first-best state of free trade. But, for any particular country, there is an incentive to unilaterally impede commerce as through an optimum tariff/non-tariff barrier which moves the terms of trade in a favorable direction or, more subtly for example, lax enforcement of patents or copyrights which allows the local citizenry to obtain valuable property by distributing “pirated” material at very low prices. However, as all countries start unilaterally to deviate from more cooperative free trade, the end result is trade restriction and reduced welfare for all. Typically this is portrayed as reversion to the non-cooperative “Nash equilibrium” wherein all nations erect “optimal response” trade barriers which maximize national welfare given the (“optimal response”) trade barriers of the other nations. Although all nations would benefit by cooperatively restoring free trade, there is now no incentive for any one country to lower its barriers given the trade barriers of the others.

So the question emphasized in the economics literature is how sovereign nations can obtain maximum cooperation – free trade -- while discouraging narrowly self-motivated defections. At first face the answer is of course a free trade agreement. But in order to discourage deviation from the agreement there must be some sort of credible punishment to discourage the deviation. This credible punishment which imposes discipline on

\(^6\) See especially Keck and Schropp (2007) for a thoughtful critical review of this literature.
agreement member nations is said to “support” the agreement. While this will be discussed more formally below, in sum the DSP can be viewed as the “punishment” for violating a trade agreement in that the violation results in a retreat toward the Nash equilibrium and reduced welfare for some period of time which could be prolonged as the trade dispute is resolved within the institutional context of the DSP. It is precisely this restriction on instantaneously restoring the initial agreement conditions that constitutes the extent of the punishment. In a sense, the DSP is an institution that makes history (behavior) matter. And so, any nation cognizant that unilateral deviation results in retribution that cannot be quickly negotiated away will not deviate in the first place so long as the short run gain does not outweigh the discounted present value of the long term loss. Or, to use the terminology of game theory, at every point in time the agreement must be “subgame perfect” or “self-enforcing.”

Importantly, just how much cooperation can be supported – how close to free trade the world can get -- depends on the nature of the DSP since this is what governs how quickly a violator will be caught and how severely punished. A very vigilant, streamlined DSP (rapid detection of violations) entailing very harsh punishment (lengthy loss of potential cooperative gains) supports the most cooperation, and in the limit free trade. In what follows we will treat the character of the DSP as the sole determinant of the trade agreement. Of course, it is the trade agreement itself that is negotiated and then the DSP put into place. But, “rule-making” is totally contingent on enforcing the rules, and this is what the DSP does. As Woolcock (2006, p. 13) observes:

“Rules mean little unless they are enforced, and all trade and investment agreements contain some degree of ambiguity that requires interpretation. Dispute settlement provisions, review procedures and the scope for non-state actors to bring cases of non-compliance are therefore important elements in rule-making...Regional interpretations of rules may, however, set precedents for wider multilateral rules and thus challenge WTO rules.”
3. The Causes and Consequences of Differing DSPs

In this section we review some of the sources of variation in DSPs across trade agreements and then reflect on some of the consequences for how these differences temper trade agreements. At the heart of the differences, of course, are the ambitions and the relative bargaining power of the would-be member countries. Then, in the next section, we explore how the differences in DSPs in conjunction with “choice of forum” may limit the extent of multilateral cooperation.

Different Rules for Settling Disputes

A number of studies have noted substantive differences in the rules and adjudication procedures applied in various trade agreements (Hudec, 1990; Smith, 2000; Woolcock, 2006). Since the goals of various trade agreements as well as the bargaining power of potential member countries differ, it is not surprising that no one DSP is institutionalized for, nor suits, all trade agreements (Schott, 2004; Aggarwal and Fogarty, 2004). For example, the regional goals of the Southern African Development Community (SADC) with a focus on transport logistics and freer trade corridors between relatively poor countries contain potential trade disputes far different from, say, the US-Jordan Free Trade Agreement with a focus on a deeper integration for credibility reasons from one side and for securing substantial intellectual property guarantees regarding pharmaceuticals from the other. Or, in the context of the original Canada-US Free Trade Agreement there was clearly concern by Canada about application of anti-dumping (AD) and countervailing duty (CVD) laws that led the Canadians to press for what is now Chapter 19 of NAFTA providing (then) for a binational panel review of disputes of AD/CVD policies. In particular, national law must be applied in AD/CVD cases within NAFTA and parties are prevented from interfering with the panel process. (This is in contrast to the more “power oriented” approach of Chapter 20.) Similarly, the EU integration of its internal market was bolstered only after the move to qualified-majority voting that restricted individual countries’ prerogative to veto unfavored decisions
(Yarbrough and Yarbrough, 1997; Garrett, 1992; Moravcsik, 1991; Sandholtz and Zysman, 1989). And, in EU-centered RTAs, investment is typically excluded since it does not fall under European Community competence (Reiter, 2006).

In their seminal study of dispute settlement mechanisms, Yarbrough and Yarbrough (1997) offer a good framework for thinking about alternative DSPs and when each might emerge. First, they characterize alternatives as running from least forceful (“self-help”) to most forceful (“pure third-party enforcement”). Specifically, they identify four sorts of DSPs:

1. Third-party information gathering and dissemination, but no external enforcement other than unilateral retaliation;

2. Non-binding third-party adjudication and recommended remedy, but no external enforcement other than retaliation;

3. Binding third-party adjudication, no retaliation allowed when no violation found, and enforcement through permissible retaliation by aggrieved party;

4. Binding third-party adjudication and enforcement, as happens for example in a national legal system.

The distinction between the first two and last two DSPs is essentially the possibility of legal non-compliance in the first two. DSP 2 characterizes NAFTA and many RTAs, as well as the GATT DSP before the Uruguay Round, as the only recourse to any disagreement with a report is retaliation since any panel finding can be overridden by a disputant and compliance is not mandatory. In fact, in the pre-1995 GATT the defendant in a trade dispute could refuse to permit panel formation, could veto adoption of a panel finding which required consensus, and could block authorization of retaliation. Thus, the

---

7 Smith (2000) offers a similar taxonomy recounting the key features of institutional design that make a dispute settlement procedure more or less legalistic. See also Smith (1995).
DSP which characterizes many RTAs is typically somewhat skewed toward the “power approach” discussed above where unequal bargaining power can be exploited more easily.

In contrast, DSP 3 is closer to the current WTO panel system wherein the panel formed in a trade dispute retains some independence and compliance with the recommendations is mandatory. This approach is generally more “rules oriented” and would somewhat blunt the advantages and disadvantages of asymmetric bargaining power. Also, this mechanism must result in an end to a dispute, although of course the resolution might entail a required remedy and punishment.

So there is a spectrum of potential DSPs, and which will be chosen depends on the nature of the set of countries negotiating a trade agreement. Yarbrough and Yarbrough cast the outcome as the result of a “coordination game” wherein agreement negotiators’ preferences differ with respect to the nature of the dispute settlement mechanism and of the accompanying enforcement. In such games, the DSP that emerges depends on the participants’ perceived costs of coordination failure and costs of compromise. Thus, what emerges in an RTA between a large nation, for whom the cost of failure is low and the cost of compromise is high, and a smaller nation, for whom the opposite holds, is a DSP like DSP 2 more amenable to the larger country exploiting its power or ameliorating its particular domestic special interests. And, in the context of a broader, more balanced set of negotiating countries as with the WTO, what emerges is a DSP like DSP 3 which is more even-handed and favorable to less powerful interests.

Smith (2000) explores the issue econometrically matching 63 post-1957 RTA treaty texts with characteristics of the RTA partners to predict DSP design outcomes. He concludes (p. 173): “Thanks to their market size and lesser dependence on trade, relatively large countries tend to prefer less legalism than their smaller counterparts...In almost every

---

8 Schelling (1960) discusses communication, announcements, and solutions in so-called “pure coordination” games.
pact with high asymmetry, legalism is absent.”\textsuperscript{9} Drahos (2007) seems to concur with this view and further motivates the next section when he notes (p. 192) that “what the United States does in trade dispute resolution matters to the evolution of the trade regime.”

\textit{The Main Differences and How They Matter}

While the discrepancies in the wording and coverage between the WTO and various RTA DSPs can be subtle and complex, generically there are three classes of differences that matter:

- private party access (“complaint initiation”)
- non-disputant representation (“third-party procedures”)
- choice of forum

In this section we discuss the first two differences and in the next section show how, in light of the first two differences, the third difference might create an undesirable link to the WTO. In essence, the argument will be that the first two differences encourage more powerful countries (in terms of bargaining power) to divert trade disputes to the regional forum and away from the multilateral forum if they can. The opposite holds for the less powerful countries. At the same time, the third difference weakens the ability of the more powerful country, when it is the alleged transgressor, to shift a dispute to its preferred regional forum and so creates an incentive to undermine the less preferred multilateral forum.

Private party access:

\textsuperscript{9} See Rosendorf (2005) for a discussion of this and other empirical studies of DSP design, but with the somewhat different focus of agreement stability.
A number of legal scholars have noted that whereas the DSP of the WTO gives only official government entities direct access to the dispute settlement mechanism, many RTAs – e.g., NAFTA with respect to the side agreements on labor and environmental issues -- and the rules of the EU go further in giving private interest groups some ability to press complaints (Levy and Srinivasan, 1996; Trebilcock and Howse, 1999). For example, recently a leading U.S. labor coalition and an industry group have filed a complaint with the U.S. Trade Representative asking that it formally sanction the government of Jordan for "gross workers' rights violations" under the free trade agreement with the United States and Israel. Specifically, the case by the AFL-CIO, the largest trade union confederation in the United States, and the National Textile Association (NTA) marks the first time a business association has formally joined in seeking a workers' rights complaint under a trade agreement (IPS, August 11, 2008).

Whether such access is desirable is not clear, but seems doubtful. On the one hand, Kahler (1995) et al suggest that domestic political interests’ linkage to political institutions is a strength of the EU as connections to policy become more substantial. Others, however, fear that narrow business interests and aggressive NGOs are driving the new regionalism in a direction not necessarily in the interests of poorer countries or even sometimes of the larger, more powerful countries’ other diplomatic objectives. (See Robertson (2006) and some of the references therein.) The argument here is that the poorer countries have little power to resist the demands of powerful interest groups such as pharmaceutical, labor or environmental interests and direct private access to the DSP removes the firewall which a larger country government might well want to preserve in pursuing its broader agenda in contrast to the narrowly defined self-interest of the businesses, unions, and NGOs.

Politically, such access by private interests creates a constituency with a bias in favor of RTAs, where they can have direct influence on behavior in the other country, and against the WTO, where they have only more muted, arms-length ability to participate in disputes. Consequently, some governments may feel politically compelled to entertain
trade disputes in the regional forum rather than in the WTO (Levy and Srinivasin, 1996). Drahos (2007, p. 209) summarizes:

“...(T)he current proliferation of FTAs represents something of a crossroads for the DSU. By constituting many possible trade courts, the United States and the EU are creating a system in which their respective influential domestic trade interests will lobby them to go to the trade court in which those interests are most likely to obtain satisfaction. How much such a system will help to improve trade liberalization that leads to real economic gains is an open question.”

Non-disputant representation:

Another difference between the DSP of the WTO and that of most RTAs concerns the treatment of interested parties to a trade dispute which have no standing in the RTA, so-called “third-party procedures.” As outlined in the DSU of the WTO, certain procedural aspects of dispute resolution allow for third party participation in a dispute between two WTO members (DSU, supra note 2 arts. 9.1, 10.1, 10.2). At the initial consultation stage, so long as the third party has substantial trade interest, that party may be enjoined in the consultations based on Article XXII of the GATT. More substantively, once the panel stage is reached, third parties have a right to be heard, to make written submissions, to receive written submissions, and even to bring their own actions (Article 10 of the DSU). Importantly, this means that weaker parties in a dispute are afforded the opportunity to enlist support through cooperation and collective action in both defending and bringing trade actions (Gantz, 1999; Drahos, 2007). In an RTA, such coalition building is simply not available when the forum is the RTA DSP because participation is limited to RTA members. Of course, for the US or EU with increasingly many bilateral agreements, this means no coalition building opportunity at all for the partner country in many trade disputes.
Drahos (2007) has argued forcefully that the procedures of the DSU that facilitate coalition building are especially important to weaker states. In particular, a coalition (p. 201) “signals to a panel that there are aggregated interests of the WTO membership at work in the dispute rather than just the individual interests of one strong State and one weak State.” This in turn can suggest to the WTO panel interpretations more considerate of the entire WTO membership (Smith, 2003). In contrast, as Drahos writes, in an RTA trade dispute, particularly within a bilateral agreement, “the capacities and power of the two parties are what they are.” Several case studies and other empirical evidence seem to support this view (Brown, 2005; Busch and Reinhardt, 2005; Shaffer, 2003).

In sum, the DSP of the RTA once again works against the interests of the weaker party in a dispute. Of course, as discussed above, the weaker party apparently agrees to the somewhat disadvantageous procedures of the RTA DSP because the agreement itself has redeeming value. And, if possible, the weaker state would prefer to fight all of its trade dispute battles in the more favorable multilateral WTO forum. In fact, a feature of most RTAs, “choice of forum,” makes this possible for many trade disputes. However, if the regional agreement DSP favors the more powerful country’s interests – either due to catering to private party self-interest or because the process favors the stronger party procedurally – then that country has an incentive to weaken any alternative adjudication forum that may be available to a trade dispute adversary. We now turn to this issue and its implications for the WTO.

4. “Choice of Forum” and the Link to the WTO

One counterbalance to the asymmetric large country advantage of an RTA DSP would seem to be the option of taking a regional trade dispute to the WTO. A recurring feature of all US RTAs and some other RTAs, including the EU and implicit at least in the Partnership Agreements, is the so-called “choice of forum.”10 This feature gives the

---

10 See Gantz (1999) for a detailed discussion of the jurisprudence involved and for a number of case studies drawn from NAFTA and from the WTO.
complaining party in a trade dispute wherein both regional and multilateral agreement obligations have been breached the choice as to which DSP is to be applied, that of the RTA or of the WTO. For reasons discussed above, there would appear to be an advantage to the stronger party channeling complaints to the regional forum and the weaker party to the WTO forum. When the government or special interests in the more powerful country are the complainants, and when the country has in place a number of RTAs, the dispute can in fact often be moved to the more favorable forum. In fact, it has been suggested that this is precisely why a more powerful country might seek a web of regional trade agreements (Bhagwati, 1994; Drahos, 2007). As a practical matter, it is really only the US and the EU that have extensive networks of bilateral RTAs and whose complainant interests stand to benefit most from diverting trade disputes away from the WTO DSP. Indeed, it has been argued that the choice provision is largely the creation of the “highly organized intellectual property lobby groups in the United States” (Drahos, 2007, p. 192).

Of course, on the surface at least, smaller, weaker partners to RTAs can themselves use the provision when they are the complainants in order to shift the trade dispute to their favored forum, presumably the WTO. This, however, links the RTA to the DSP of the WTO and creates an incentive for the stronger member of an RTA to weaken this option: That is, to weaken the WTO DSP. We next discuss the consequences of this incentive in a more formal model of trade disputes and the DSP.

The Model

Recently, economists have focused on the theoretical underpinnings of dispute settlement procedures and the connection with the overlying trade regime. The basic idea is that

---

11 Or, as Reiter (2006) has bluntly observed: “Of course, given the very asymmetric character of many agreements and the US interest as a capital exporter in securing access and protection for investment in its RTA partners’ markets (and not the reverse), the United States has everything to gain from applying the open-ended disciplines of the NAFTA model in combination with investor-state dispute settlement. In asymmetric agreements such stringent rules pose no real threat to US interests in terms of the risk of legal challenges from foreign investors. NAFTA-type investment rules are therefore typically part and parcel of the US FTAs (and BITS).”
credible commitments to mutually beneficial trade agreements must be self-enforcing and so require that any contemplated deviations confront potential punishment sufficiently harsh to discourage violations aimed to secure short-term gains.

Formally, suppose initially that there are two ("large") countries, Home (H) and Foreign (F), engaged in free international trade as in a standard competitive trade model. While the simple formal model presented here couches trade agreements in terms of tariffs, and so countries are taken to be "large" so that they can have an incentive to deviate from the agreement by imposing an optimum tariff, clearly the spirit extends easily to investment guarantees, intellectual property rights protection, and so on. In this sense virtually every country is "large" in that each can affect world prices of some things that it imports. Indeed, many trade disputes recently, mostly informal rather than formally notified, have been between economically quite disparate countries and, in the case of the US, have revolved around IPR ("pirating and patenting") and investment access issues.

Using capital letters for home country variables, denote by \( M \) and \( X \) (\( m \) and \( x \)) the free trade quantities of imports and exports, respectively, of H (F). In equilibrium, \( M = x \) and \( X = m \). Both countries have the option of imposing import tariffs, \( T \) and \( \tau \). Social welfare for each country is assumed to be a continuously differentiable, quasi-concave function defined over consumption of the goods. Welfare can be written in implicit form as a continuous function of the tariff levels, \( W = W(T, \tau) \) and \( w = w(T, \tau) \), with welfare in autarky normalized to zero. (See, for example, Mayer (1984).)

Each country now has a best response tariff conditional on the tariff level of the other country. In a one period game, given the assumptions made, there exists an interior Nash equilibrium and we denote the solution \( T_n \) and \( \tau_n \). Note that prohibitive tariffs that take trade to autarky are also trivially a Nash equilibrium.

Now suppose that the game is an infinitely repeated one which unfolds over a sequence of discrete intervals \( t, t = 0, 1, 2, 3, \ldots \). A (cooperative) trade agreement beginning in period \( t = 0 \) is represented by a pre-set level of negotiated tariffs \( T_t \) and \( \tau_t \) at each period.
(stage) yielding welfare levels of $W(T_t, \tau_t)$ and $w(T_t, \tau_t)$ for H and F respectively.

Suppose that an agreement sets tariff levels at a constant $T^*$ and $\tau^*$ for all time. In a cooperative agreement with $T^* < T_n$ and $\tau^* < \tau_n$, welfare $W$ and $w$ are higher than in the non-cooperative Nash equilibrium.

Denote by $\omega_t$ the one period gain to H from deviating from the agreement in period $t$ -- e.g., by imposing an optimum tariff in $t$ -- and by $W_{pt}$ the value of the “punishment payoff” which would ensue beginning in the next period -- e.g., the Nash response by F to the H tariff. Note that “no punishment” in a recurrent trade agreement with the possibility of renegotiation would mean that the agreement is simply reinstated after the violation is detected. But then $W_p$ would just coincide with the initial agreement payoff after period $t=0$ no matter what, and so it would always pay to “cheat” until caught. As it turns out, $W_p$ will entail reversion to a Nash equilibrium for some finite period of time.

Then, using $\delta$ to denote the discount factor at each stage and normalized by $(1-\delta)$, a necessary and sufficient condition for the agreement to be a subgame perfect equilibrium (SPE) in period $t=t'$ is given for H by:

$$
(1) \quad \sum W(T, \tau) > (1-\delta) \omega_t + \delta W_{pt}
$$

where, $\sum W(T, \tau) = (1-\delta) \sum_{t=t'}^{t=\infty} \delta^t W(T, \tau)$.

That is, the LHS which represents the discounted present value at period $t'$ of adhering to the agreement must exceed the RHS, which is comprised of the immediate value of deviating from the agreement (as with an optimum tariff) and the discounted value of the punishment payoff that would ensue upon detection the next period.\(^\text{12}\)

The trade agreement as represented here is then the jointly determined solution to the H problem $\max W$ subject to (1) and the analog for F. In much of the literature it is

\(^{12}\text{The normalization (1-\delta) is introduced to facilitate comparisons across time and } W_{pt} \text{ is discounted by } \delta \text{ because punishment takes place in the next period (Ludema, 2001).}\)
assumed that $H$ and $F$ are symmetric in order to simplify the mathematics involved, but this is not necessary. Nonetheless, it is convenient to maintain the assumption when we refer to the WTO trade agreement. Thus, $T$ will be the same for all members and we can suppress $\tau$ in the notation. Also, we can then allow that there are more than two countries in the WTO agreement.

*The Solution With and Without Renegotiation*

Since $W_p$ decreases the more severe is the punishment, the most severe punishment which is an SPE would seem to be a permanently prohibitive tariff resulting in autarky forever. That is, $W_p = 0$ as the violator loses the gains from trade for the rest of history. This, in turn, would support the trade agreement closest to free trade since the temptation to deviate for a short-run gain comes at a very high cost. Or, in terms of condition (1), $LHS > RHS$ would hold for the relatively lowest tariff levels given any $\delta$. The level of tariff would itself be the lowest consistent with the inequality holding (weakly). However, in a recurrent trade agreement, such a punishment path is not itself credible in that it is not “renegotiation proof.” That is, if one country violated the trade agreement, the other country could then impose autarky forever: But why would they? Once in the no trade state, it would behoove both nations to renegotiate at least some mutually beneficial trade and both would be better off by simply moving to the interior Nash equilibrium, $T_n$ and $\tau_n$.

So the question addressed in the literature (Ludema, 2001; Klimenko et al, 2008; and others) is to find the feasible (renegotiation proof) punishment path which in turn then determines the character of the trade agreement that is supportable. Following Ludema, such a punishment path that would support the tariff agreement with tariffs at the level $T^*$ entails, upon detection of a violation, reverting to the Nash equilibrium for some length of time and then reestablishing the initial agreement. The punishment in the symmetric game for $H$ and $F$ is given by a path through time:
\{(T_n, T_n) \text{ for } t^* - 1 \text{ periods, } Q_t^* \text{ for } 1 \text{ period, and } (T^*, T^*) \text{ beyond } t^* + 1\}

where \(Q_t^*\) is a randomized tariff pair and serves a technical convexifying role.\(^\text{13}\)

Importantly, it is precisely the friction created by the DSP – detection, hearings, adjudication – which takes \(t^*\) periods to run its course that creates the punishment (loss of freer trade benefits) that in turn supports a trade agreement. That is, if renegotiation after a trade agreement violation were costless, then countries would continually violate the agreement and the only equilibrium to emerge would be the Nash equilibrium forever (Klimenko et al, 2008). The more severe is the punishment – the longer is \(t^*\) -- the closer to free trade will be the trade agreement.

So we can think of the trade agreement that might emerge as a search for \(t^*\) which in turn determines the lowest tariff level \(T\) which is supportable or consistent with the SPE condition (1). For what follows, it is convenient to adopt the approach of Ludema wherein the optimal trade agreement values – tariffs and length of reversion to the Nash equilibrium – is reduced to maximizing welfare of the punishment path. And this, in turn, can be simplified to the problem:

\[
\text{(2) } \max_{T, \Lambda} [(1 - \delta^\Lambda) W(T_n) + \delta^\Lambda W(T)]
\]

subject to the SPE condition (1), and where \(\Lambda\) is the “effective punishment length” and approximately \(t^*\), the number of periods before the dispute is resolved and the initial agreement tariff levels reestablished. There is only one argument in the welfare function reflecting the symmetry assumption in the WTO forum. The solution to problem (2) obtains at \(\Lambda^*\) and \(T^* = T(\Lambda^*)\).

\(^{13}\) Ludema’s seminal portrayal of the DSP in this context (Ludema, 2001) proves the existence of this punishment path to support a tariff agreement and discusses the convexification assumption. Klimenko et al (2008) derive a similar punishment path characterization of the DSP and argue that it will only exist if it can be externally imposed. Both papers also provide an excellent discussion of the institutional underpinnings for characterizing a DSP as a reversion to a Nash equilibrium for some finite period of time.
Geometrically, the solution is illustrated as in Ludema (2001), but with different notation, in Figure 1.

**FIGURE 1 GOES HERE**

For convenience, call the country represented the home country H. In Figure 1, \( W(T_n) \) represents welfare in H associated with Nash reversion. \( W(T(\Lambda)) \) reflects the welfare associated with the trade agreement which depends on the tariff level, \( T \), which in turn depends on the severity of punishment for deviation from the agreement, i.e. the length of reversion to the Nash equilibrium, \( \Lambda \). The higher is \( \Lambda \), the higher will be welfare since lower tariffs can be supported in the cooperative equilibrium. The welfare of the punishment path \( W(\Lambda) \) is itself the maximand of (2) which is a convex combination of Nash reversion for all time, \( W(T_n) \), and welfare of no deviation from the agreement, \( W(T) \).

Now consider the shape of the “punishment path” welfare \( W(\Lambda) \) as it relates to the “length of punishment” \( \Lambda \). Note that at \( \Lambda=0 \), there is no punishment for deviation and so the Nash equilibrium obtains. As \( \Lambda \) approaches infinity, \( \delta^{(\Lambda)} \) tends to zero and so again \( W(\Lambda) \) coincides with the Nash equilibrium value \( W(T_n) \). Somewhere in between these extremes, \( W(\Lambda) \) achieves a maximum -- \( \Lambda^* \) -- and this supports the tariff agreement \( T^*=T(\Lambda^*) \) with welfare \( W^*(T^*(\Lambda^*)) \).

**The Impact of Choice of Forum**

Suppose now that there is a multi-country WTO Agreement fashioned as above by more powerful states like H and F along with weaker states who are also members of the WTO. Furthermore, suppose that H is negotiating or has negotiated an RTA with a fellow WTO
member country that is a relatively weaker state, f. For example, we might think of H as the US and f as Colombia. While H is taken to be more powerful than f in a bargaining sense, we assume that the bargaining power of H is relatively weaker in the multilateral setting and stronger in the bilateral regional setting. This follows from the above discussion of the possibilities for coalition building available to weaker countries in the more global forum of the WTO and the presence of counter-balances represented by other powerful countries to any one powerful nation. Suppose also, for clarity and following Maggi (1999), that we can treat the spillover economic (trade) consequences of the trade negotiation outcomes in the RTA as independent of the WTO negotiations. That is, essentially we rule out significant trade diversion owing to trade policies adopted in the regional setting on the wider group and vice versa. Then we will suppose that welfare from the WTO Agreement and the RTA is additively separable so that we can use the above analysis as a starting point. (For example, similar to the three country model of Maggi with separable bilateral relationships, assume that each pair of countries trade two distinct goods that are neither supplied nor demanded by the third country.)

Now in the absence of choice of forum, the trade agreements would not be linked directly in that each country, say H and f, would negotiate for the most favorable agreement in the forum in question. Again, we assume that H is a large country endowed with stronger bargaining power relative to f, but that H has less relative leverage in the multilateral forum, where the bargain is among many nations, and more in the regional forum, where the bargain is just with f. Since it is the DSP that supports the agreement in this model, the DSP negotiated in the WTO will look different from the DSP of the RTA, and so too will the agreements, or supportable levels of obligations in the WTO versus the RTA. In particular, in the regional agreement H will be able to fashion a DSP which is more favorable to its export trade-opening or IPR interests and more tolerant of its self-motivated or protectionist trade agreement violations than would be the case in the multilateral forum. For example, in both the US-Morocco FTA and the Jordan-US FTA,
the US was able to obtain with “TRIPS-plus” concessions regarding patent protection and pharmaceutical generics which go far beyond any WTO TRIPS obligations.¹⁴

Next, suppose that there is a choice of forum provision in the RTA. In modeling the impact of choice of forum, we will maintain the assumption of symmetry of countries in the WTO agreement negotiations but not in the RTA negotiations where we assume H to have a bargaining advantage over f. Besides being mathematically convenient, this seems plausible in that in the WTO forum, narrow country self-interest is more difficult to advance given the possibilities of coalition building. For example, in the Doha Round, many developing countries were able to hold negotiating positions credibly that simply would not have been possible in a bilateral agreement with the EU or the US. Roughly, for example, Colombia is less malleable when negotiating in the WTO than it is in negotiating the US-Colombia Free Trade Agreement. And, this is also consistent with the incorporation into the US bilaterals of many features championed by US business interests and NGOs which go well beyond what is achievable in the WTO – specifically, some clauses on market access, investment guarantees, IPR, labor, and environmental conduct. (Alternatively, an equivalent for the model below, we could think of the WTO Agreement as being shaped by a few strong states – the US and EU – who are also engaged in RTAs with a number of weaker states without much influence in the WTO beyond membership.)

Now, with a choice of forum provision, although H would by assumption be able to negotiate a more favorable DSP with the weaker f in the RTA, any DSP negotiated in the WTO can in some sense be extended at least partially to the RTA, thus weakening H’s bargaining position – that is, imposing more balanced discipline on H -- in the RTA. Consequently, H has an incentive to incorporate this downside into its WTO negotiations and attitude toward that forum’s DSP.

¹⁴ While we do not model the RTA explicitly, technically it would emerge in a model similar to the WTO portrayal above except the “threat point” in the bargaining when an agreement breaks down would be an asymmetric Nash bargaining game solution. Klimenko et al discuss this in terms of “frontier punishment” wherein the punishment reversion is to the frontier of supported repeated game payoffs (worse than cooperation) and just where along the frontier is determined by the relative bargaining power of countries.
Mathematically, the trade agreement welfare maximization problem now must be amended to incorporate a term $Z(\Lambda)$, which will enter negatively, meant to capture the reduction in H welfare when the WTO DSP is strengthened because this undermines H’s ability when it is the trade dispute defendant to divert disputes into the more favorable (for H) DSP of the RTA due to choice of forum. Although we will not model the RTA negotiations explicitly, the term $Z(\Lambda)$ should be considered as a reduced form representation of the extent to which access to the WTO DSP benefits F (at the expense of H) in any RTA trade dispute where F is the complainant. By construction, we write $Z(\Lambda) > 0$ with the derivative $dZ/d\Lambda > 0$ to reflect that a harsher WTO DSP -- i.e., larger $\Lambda$ -- gives the weaker RTA member F the ability to hold H more accountable in a trade dispute over an agreement violation. Presumably $Z(0) = 0$.

Thus, the WTO trade agreement will emerge as before from the maximization of the modified problem:

$$\max \quad W - Z$$

subject to the SPE condition.

As before, in the WTO negotiations the effective punishment length and tariff levels are determined by solving the now modified problem:

$$\max_{T,\Lambda} \quad [(1 - \delta^\Lambda) W(T_n) + \delta^\Lambda \{W(T) - Z(\Lambda)\}]$$

subject to the SPE condition (1).

The problem is the same as (2) above except for the presence of $Z(\Lambda)$ meant to capture that a harsher (harder to manipulate for stronger parties) DSP in the WTO, i.e. a higher
value for $\Lambda$, reduces H welfare by weakening its power-oriented negotiation ability in the other (RTA) forum. The Z term arises in the modified problem (3) because the punishment path welfare is a weighted sum of the welfare of the trade agreement and the reversion to the Nash equilibrium.

Now denote the maximand of the initial problem (2) above as $B = [(1 - \delta^\Lambda) W(T_n) + \delta^\Lambda W(T)]$ and write the maximand in (3) as:

(4) \[ B - \delta^\Lambda Z(\Lambda) . \]

At the maximum, and noting that $T = T(\Lambda)$, we have:

(5) \[ dB/d\Lambda - [\Lambda \delta^{\Lambda-1}Z(\Lambda) + \delta^\Lambda dZ/d\Lambda] = 0 \ . \]

Evaluating the expression at the original equilibrium value for $\Lambda$ that solves problem (2), $\Lambda^*$, reveals:

(6) \[ dB/d\Lambda - [\Lambda \delta^{\Lambda-1}Z(\Lambda) + \delta^\Lambda dZ/d\Lambda] < 0 \]

since $dB/d\Lambda = 0$ at $\Lambda^*$ and $[\Lambda \delta^{\Lambda-1}Z(\Lambda) + \delta^\Lambda dZ/d\Lambda] = d(\delta^\Lambda Z)/d\Lambda > 0$.

Intuitively, of course, this is because the original agreement ignores the deleterious indirect effect on H welfare of a stronger WTO DSP. Now, as before, B is increasing in $\Lambda$ up to the maximum and then decreasing, as in Figure 1. However, since $dB/d\Lambda = 0$ at the maximum in the original problem, as in Figure 1, and since $d(\delta^\Lambda Z)/d\Lambda > 0$, it follows that, at the new maximum welfare solution, $\Lambda$ must be smaller, noting that in the vicinity of the original solution value $dB/d\Lambda = 0 < d(\delta^\Lambda Z)/d\Lambda$.

\[ ^{15} \text{In general, if a function f(x) achieves an interior maximum at x}', and if g(x) is increasing in X, then f(x) - g(x) achieves a maximum at x}' < x'. \]
The new equilibrium, \( \Lambda^{**} \) and \( T^{**} \), along with the initial equilibrium, are illustrated in Figure 2.

FIGURE 2 GOES HERE

Total welfare of the trade agreement is now given by \( W(T(\Lambda)) - Z(\Lambda) \). The punishment path is again a convex combination of agreement compliance total welfare and the Nash equilibrium (reversion) welfare, which together can be written as \( W_p(\Lambda) - \delta^\Lambda Z(\Lambda) \). As before, at \( \Lambda=0 \) the Nash equilibrium obtains, recalling that \( Z(0) = 0 \). And, as \( \Lambda \) approaches infinity all of the weight in the convex combination is put on the Nash equilibrium.\(^{16}\) Hence, \( W_p(\Lambda) - \delta^\Lambda Z(\Lambda) \) achieves a maximum somewhere in between at \( \Lambda^{**} \) which supports the tariff agreement \( T^{**} \). As discussed above, \( \Lambda^{**} < \Lambda^* \) because of the presence of the \( Z \) term reflecting the spillover effects into an RTA.

Two points are of interest. First, at \( \Lambda^{**} \) -- representing the new WTO DSP -- the supportable WTO tariff is higher than in the absence of an RTA. This is due to the “choice of forum” provision which H negotiators realize allows f to impose stricter enforcement in an RTA than H could otherwise negotiate by using its superior bargaining power in an RTA. That is, the dilution of H’s bargaining power in the WTO forum seeps into the RTA forum. So, H dilutes the WTO DSP in response.

\(^{16}\) A natural assumption is that \( W(\Lambda) - Z(\Lambda) > 0 \) everywhere, but this is not necessary. All that is required in order to avoid a corner solution at \( T^{**} = T_n \) is that \( W(\Lambda) - Z(\Lambda) > 0 \) from the right in the vicinity of \( \Lambda = 0 \). The corner solution would represent the extreme case where the RTA is so important and advantageous for H that any enforcement at all in the WTO, given “choice of forum”, is too destructive in the RTA.
Second, since it is the DSP that underwrites the self-enforcing tariff level in the WTO, the resulting WTO tariff is now higher -- $T^* < T^{**}$. Note that H welfare arising from the WTO agreement is now lower, $W^{**}$, for two reasons: 1) The WTO DSP is now weakened and so $\Lambda$ is lower and cooperative tariffs higher, and 2) At the lower $\Lambda$, H welfare is further reduced by $Z(\Lambda)$ because any WTO DSP at all can be used against H in the RTA due to “choice of forum.” (This would not happen if the WTO DSP and the RTA DSP were exactly the same institution, but then H would not be securing more favorable treatment in the RTA despite enhanced bargaining power.)

Naturally H welfare will be reduced on account of higher WTO tariffs, but this ignores that there are offsetting gains in the RTA where H can now obtain a more favorable regional agreement. Of course, overall welfare for H, including welfare obtained from the RTA, will be higher. As in Figure 2, the gain in the RTA must exceed the loss on account of H taking the RTA into account, which is given by $W^*-W^{**}$. Therefore, the extent to which RTA negotiations erode the multilateral process depends on the magnitude of the spillover costs $Z(\Lambda)$. Apparently these costs depend on the extent of H’s asymmetric power in an RTA and on how much gain can be achieved by exploiting this power in the regional forum.

_Some Preliminary Evidence and Indications_

Empirically, the extent to which the choice of forum option erodes multilateral cooperation is difficult to assess. Case studies in Drahos (2007) suggest it might be extensive, although this was not the focus of that study. He does conclude though that “… it is worth noting that these choice-of-forum provisions to be found in US FTAs do not sit very comfortably with the goal of strengthening the multilateral trading system.”

Gantz (1999) and others certainly document well that choice of forum both matters and diverts trade disputes to the forum viewed as more favorable by the complainant. But, these studies do not address directly the forum diversions’ potential role in weakening the
WTO. Nonetheless, Drahos (2007, p. 209) implies as much in concluding his case studies when he writes:

“One priority that developing countries should be thinking about is to argue for provisions in these dispute settlement chapters that require the parties in the case of double breach to take the matter to the WTO. That would also be consistent with the aim of the DSU to strengthen the multilateral trading system. The effect of creating a web of bilateral dispute resolution fora will be to give the United States and the EU more opportunities to play for rules.”

Brown (2005) investigates developing country participation in the WTO and finds that the existence of RTAs negatively affects the participation of these countries in the WTO. In light of this, Drahos (2007) observes that as the US and the EU enter into more FTAs with developing countries, the coalitional possibilities of the WTO DSP will be weakened for developing countries as a group. Also, developing countries in the WTO but with an RTA with the US or the EU may become reluctant to join a coalition in a trade dispute against either for fear that this would create other costs in the RTA. Again, the WTO DSP would be weakened. Whether these adverse consequences for the WTO are an explicit aim of US or EU policy motivated by choice of forum, of course, is not proven by these observations.

Finally, since the welfare gain in the RTA must offset the loss in the WTO on account of weakening the latter DSP, it is at least possible to entertain reasonable conjectures. Roughly, the US would probably not be willing to give up cooperative gains in the WTO in order to enhance its bargaining position with a small country where there is little to gain – say, Singapore – but might be more sensitive when there is a potentially large market RTA like Korea, or where the RTA is with an important trade partner eager to make some substantial concessions for domestic political reasons – say, Mexico. The model and logic presented above suggest that these latter agreements could indeed ultimately impair the WTO DSP and so the extent of multilateral cooperation.
5. Summary and Conclusions

International trade agreements give rise to trade disputes. Dispute settlement procedures proscribe the mechanisms by which disputes are to be resolved. The WTO DSP is broadly outlined in the Dispute Settlement Understanding of the Uruguay Round. Recently, however, the increasing presence of regional trade agreements in international commerce – now counted at over 250 by the WTO – has given rise to a host of parallel dispute settlement procedures which are subtly different from each other and that of the WTO.

At one level, this paper has described some of these differences in DSPs and suggested that some countries are generally treated more favorably by the regional DSP and other countries by the WTO DSP. The reasons largely have to do with the disparate extent of bargaining power or political/economic leverage available in a regional versus a more multilateral forum. There is some evidence that the regional forum favors the stronger State because it reduces the possibility of coalition building for the weaker State. And certainly the regional forum favors certain private interests – business, labor, NGOs -- in the stronger State because they are given standing and the ability to initiate complaints according to many regional DSP.

Now, in a number of trade disputes, nations which hold concurrent membership in the WTO and an RTA can choose which DSP to use in resolving a complaint. Since the DSPs differ, this creates an incentive to divert a complaint to the more favorable DSP. There is considerable evidence that this in fact occurs. But this also gives an incentive for nations to weaken the DSP to which potential adversaries in a complaint might try to divert a trade dispute. Since stronger states, like the US and EU, might tend to prefer the regional forum, these states gain an incentive to weaken the WTO DSP.

The issues involved now become deeper and potentially more important in light of the current economic view of the role of a DSP portrayed in the international trade and political science literature. There the DSP is cast as the linchpin in free trade agreements,
determining the extent to which freer trade can be achieved. This view builds on contract theory.

Central to any mutually beneficial cooperative contract is the belief by the contractors that the agreement will be fully and faithfully executed. Within nations, courts and cops enforce such behavior and so serve to underwrite the mutual gains which accrue to the contracting parties. Deviation from the terms of an agreement can be punishable offenses. Between sovereign nations, there are no such external forces to ensure compliance with the terms of international agreements. Generally, it is only the parties to the agreement themselves that have the wherewithal to punish cheaters. (By analogy, it is the victim that must punish the criminal.)

In the realm of commerce, unimpeded international trade and investment is mutually beneficial for all nations, but any one nation can do even better by discriminating against foreigners in favor of its own citizens. Of course, when all nations then revert to such discrimination – optimum tariffs/NTBs, pirating, international fraud, etc. – the gains from trade are reduced and could be enhanced through international agreement not to discriminate so, or at least to discriminate less. But, in the absence of credible courts and cops, international cooperative agreements need to be self-enforcing; That is, the perception of punishment that could credibly be meted out by partners to the agreement must outweigh any gains derived from cheating. And, the harsher is the perceived punishment, the more is the cooperation that can be supported – i.e., the closer the world comes to free trade.

In international cooperative trade agreements, the dispute settlement procedures can play an important punishment role. When an agreement violation occurs, counterparties not only withdraw trade benefits (retaliate) but the DSP requires that the trade dispute be adjudicated in a proscribed manner that takes time. Thus, the violator is institutionally prevented from apologizing and returning immediately to the initial cooperative agreement on the previous more favorable terms. If it were not for the DSP, it would
otherwise be in everyone’s interest to let the violator do so, thus inadvertently undermining any cooperative agreement in the first place.

But now return to the “choice of forum” feature of many RTAs. States are given an incentive to weaken the least favorable forum. For stronger states, this forum is often that of the WTO. So there is an incentive to weaken the WTO DSP. But, since it is the DSP that underpins the extent of trade liberalization in the first place, the result is less cooperation in pursuing freer trade and investment. Thus, the explosion of RTAs may undermine the multilateral process in yet one more way.
References


IPS-Inter Press Service International Association


Smith, James McCall. 2003. WTO Dispute Settlement: The Politics of Procedure in Appellate Body Rulings. World Trade Review:


Figure 1

Figure 2