

A Survey of the Literature on the WTO Dispute Settlement System

Henrik Horn and Petros C. Mavroidis

A Survey of the Literature on the WTO Dispute Settlement System

by

Henrik Horn

The Research Institute of Industrial Economics, Stockholm
Bruegel, Brussels
Centre for Economic Policy Research, London

and

Petros C. Mavroidis

Edwin B. Parker Professor of Law, Columbia University
University of Neuchâtel
Centre for Economic Policy Research, London

November 20, 2006

JEL no: F13, K33

Keywords: WTO, dispute settlement, DSU

* We are grateful to Kyle Bagwell, Alan Guzman and Alan Sykes for helpful comments on an earlier version of this paper. A slightly different version of this paper is forthcoming as “International Trade: Dispute Settlement”, ch. 5 in Guzman, Alan and Alan Sykes (eds) *Research Handbook in International Economic Law*, Edward Elgar Publishing Ltd, 2006.

I. INTRODUCTION

A considerable amount of academic interest has been devoted to the dispute settlement (DS) mechanisms of the GATT and the WTO, in particular during the last decade. The purpose of this paper is to present and discuss what we see as main themes in this body of research, and in particular the empirical work on the functioning of the DS system.

The large volume of research in the field makes it necessary to restrict attention to a few areas. We will concentrate on a couple of themes that we find particularly interesting, omitting many other (and surely to others more interesting) issues. A noticeable omission is that we will not deal with enforcement issues, this literature being too extensive to be covered in this survey.

The paper is structured as follows. Sections II and III provides, respectively, legal and economic backgrounds to the ensuing discussion. Section II highlights quintessential elements of the DS mechanism in the WTO – the *Understanding on Rules and Procedures Governing the Settlement of Disputes* (DSU). Section III then briefly discusses the role of dispute settlement from the angle of economic theory. The more policy-oriented literature often views the objective of dispute settlement as to resolve conflicts, to enhance transparency and predictability, or to implement the agreed liberalization of trade. But such explanations are not satisfactory from the point of view of economic theory, which seeks to explain *how* the DS mechanism may help achieve the various aims that have been suggested, and how these aims may interact with on another. For example, the desire to ease the resolution of disputes may conflict with the desire to maintain system integrity, and while transparency may increase the predictability of the system, it can make settlement, and trade liberalization, more difficult. Economic theory does certainly not offer satisfactory explanations for all these matters, but provides at least a few insights into the functioning of dispute settlement.

Section IV discusses two main themes in the empirical literature on dispute settlement, the first being the determinants of the participation in dispute settlement processes, and the second the role of the DS system for the settling of disputes. In this section we also point to a number of aspects of this body of research that we believe are in need of improvement. Some of these are quite clearly “researchable”, while it is less clear in other cases how to they could be tackled.

Section V, finally, points to some additional areas that we believe are important for understanding how the DSU operates, and which believe deserve more attention.

Before turning to the literature review, we would like to mention that we have assembled an extensive data set that is available for free downloading from the website of the World Bank (found under www.worldbank.org/trade). The data set currently covers all 311 WTO disputes initiated through the official filing of a *Request for Consultations* at the WTO, from 1 January 1995 until July 31, 2004, and for these disputes it includes events occurring until February 28, 2005. The data set covers exhaustively all stages of dispute settlement proceedings, from the moment when consultations are being requested to the eventual implementation of the rulings. It contains several hundred variables, providing information on various aspects of the legal procedure. The intention is to update this data set early in 2007 to include all disputes and events up to the end of 2006. The data set is described in Horn and Mavroidis [2006a].

II. THE WTO MECHANISM FOR DISPUTE RESOLUTION: THE DSU

We will in this section provide a very brief introduction to salient features of the DSU. A more detailed description can be found in Palmeter and Mavroidis [2004].

The DSU is one of many annexes to the agreement establishing the WTO. The administration of disputes is entrusted to the Dispute Settlement Body (DSB), comprised of representatives of all WTO Members. The system is highly decentralized: disputes

cannot be initiated *ex officio*; there is no authority assigned to a supra-national entity (a watchdog) to initiate complaints against WTO Members; and disputes are launched at the initiative of a WTO Member.

Adjudication in the WTO system has two phases: one in principle bilateral, and one multilateral. The bilateral phase consists of consultations between the complainant and the defendant.

Few disputes are of a purely bilateral nature however, and even if this is the case, other WTO Members might have an interest in the interpretation of the rules pertinent to this particular transaction since, arguably, such interpretation might be influential in interpreting their own commitments in the future. To this effect, when requesting consultations, the complainant has to notify the WTO as to the subject-matter of the dispute. Other WTO Members wishing to join as co-complainants can do so, provided that the defendant accepts their request (Art. 4.11 DSU).

The subject-matter of a particular dispute can range from disagreement over a particular transaction and its consistency with the relevant WTO law (e.g., A believes that B imposed antidumping duties without having demonstrated any injury resulting from dumped imports), to disagreements over the consistency of a legislation with the WTO rules (e.g., A believes that B, by enacting legislation which precludes its investigating authorities from conducting injury-analysis in the context of an antidumping investigation, is violating its obligations under the WTO). The standard of review however, is more demanding in the latter case.

Assuming that the parties reach no solution during the consultations-stage, the complainant can request the establishment of a panel to adjudicate the dispute. Such a request leads to the second, multilateral phase, consisting of two parts: the first is the panel procedures, the panel being the analog to a first instance court; the second part is the procedure before the Appellate Body, the last instance court. Whereas the three-person panels are *ad hoc* adjudicating bodies, the composition of which depends, in

principle, on the will of the parties to the dispute (and, in case of disagreement between them, panelists will be appointed by the Director General of the WTO), the Appellate Body is composed of seven judges appointed on four years terms, renewable once. Panels have competence to review the factual record and the legal issues before them, whereas the Appellate Body's review is limited to the latter.¹

Following a request for establishment of a panel, other WTO Members have a limited time within which they can request to appear before the panel as third parties and present their views on a particular case. Panels are assisted by members of the WTO secretariat (usually lawyers, and occasionally economists as well). The role of the secretariat should not be under-estimated: it is normally the secretariat that drafts the reports, even though panelists of course have the last word. Importantly, with the exception of secretariat members working for the Appellate Body, members of the secretariat, besides advising the members of the panel they are assigned to, also advise WTO Members on issues of their competence, and are thus not assigned exclusively to service panels.

Assuming that the Appellate Body accepts the original complaint,² the defendant will be requested to implement the judgment. Implementation should, if possible, occur immediately, although this is hardly ever the case. Instead, defendants are granted an implementation period, that is either agreed bilaterally, or through recourse to binding arbitration.

Defendants do not always have sufficient guidance to implement a judgment against them. The lack of guidance is a direct consequence of adjudicating body reports that merely recommends the losing respondent to bring its measures into compliance with its obligations. In the unusual instance where adjudicating body reports suggest ways to implement the final judgment, the addressees of suggestions are nevertheless free to choose their way of complying, since suggestions are not legally binding. As a result, disagreements as to whether implementation has truly occurred easily arise. In such case,

¹ However, the distinction between a factual and a legal issue is often over-stated in legal literature, since many disagreements about facts can be formulated in terms of a legal issue.

² If the Appellate Body rejects the original complaint, the case will, of course, be closed.

the dispute is referred to a compliance panel whose task it is to determine whether implementation occurred; the panel's report can be appealed.

If the complainant believes that no implementation has occurred, the complainant can request authorization to suspend concessions. If granted, the complainant will have the right to raise its bound duties to the level necessary to inflict on the defendant damage equal in value to the damage the complainant suffered as a result of the practice that was found to be illegal.

The entire adjudication process can, in practice, take up to three years. Nothing prohibits however WTO Members to reach a Mutually Agreed Solution (MAS) at any stage throughout the process. A MAS must be notified to the DSB, where any WTO Member can raise questions as to its consistency with the WTO rules.

Having presented some salient features of the DSU, we will now turn to studies that may help shed light on how the DSU works. In the following section, we will briefly set out some ideas concerning the possible economic role of such a system in a trade agreement. In section IV we then turn to empirical work that seeks to shed light on the actual working of the system.

III. THE ROLE OF THE WTO DISPUTE SETTLEMENT MECHANISM

An important (implicit or explicit) aim of the empirical literature discussed below, is to evaluate whether the DS mechanism in the GATT and the WTO have fulfilled their purposes. The task obviously requires an understanding of what these purposes are and how they are meant to be achieved, and it is natural to start from objectives expressed by legislators in the agreement, or elsewhere. But, such stated objectives are often expressed in very imprecise language. To appreciate the role of the DS mechanism, it is therefore necessary to view it from a more theoretical point of view.³

³ We will here be extremely brief. For a fuller survey of the economic theory literature on trade agreements,

The question of what role a formal DS system might serve may at first glance seem superfluous, the answer being seemingly obvious: to help adjudicate, and possibly also avoid, disputes. At closer scrutiny, however, matters are not as simple, in particular not once account is taken of the possibility for members to let informal mechanisms solve the same problem. But one can identify several possible rationales for a formal settlement system. In present these, we need first to highlight two central, and in the literature often emphasized, constraints on the scope of any international trade agreement of such a broad nature as the WTO.

III.1 Two inevitable features of trade agreements

There is a fundamental difference between the circumstances under which a trade agreement operates, and those under which parties contract under domestic law. In the latter case, contracts can be enforced by third parties, such as courts, who have at their disposal the ability to issue physical action, such as a intervention. This allows the contracting parties to include provisions that for certain contingencies specify courses of actions that would not be voluntarily undertaken by all parties to the contract

In the case of a trade agreement there is no outside party who can ensure that the members to the agreement abide by their obligations, and as a result, it must be *self-enforcing*.⁴ The crucial implication of this is that the agreement, and its dispute resolution mechanism, must be such that it is *always in the interest of each member to behave so as to preserve the integrity of the agreement*. In other words, members must not be put in positions where they would prefer to sacrifice the collaboration for short-run gains. The essential mechanism that allows the parties to make a meaningful agreement is thus the threat of withdrawal from the agreement by an adversely affected party. Countries are largely stuck with each other, and typically expect to remain so for the foreseeable future, and contract breach by a country is likely to be observed not only by directly affected partners, but also by the membership as a whole.

and on dispute settlement, see Staiger [1995].

⁴ We are not aware of any legal/economic analysis of the extent to which members to a trade agreement may contract third countries to help in the enforcement of the agreement.

The second constraint on an agreement with a scope such as the GATT/ WTO, stems from the fact that such an agreement must for practical reasons be *incomplete*: it will contain relatively few explicit, detailed specifications obligations (such as tariff bindings specified at a rather disaggregate product level), but much or most of its ambit will be contained in vaguely specified provisions, such as that of National Treatment. As a result, the determination of the exact ambit is left to be decided in the future, when a conflict arises.

If the agreement contains a dispute resolution mechanism, it is likely to carry a heavy burden in two respects. First, it is likely to have to administer a large number of disputes due to the fact that so much in the contract is left unspecified. And, second, a significant responsibility rests on this mechanism, since it is likely to play a very prominent role in shaping the practical ambit of the agreement.

III.2 The pros and cons of an explicit dispute settlement mechanism

What has been seen so far is thus that agreements must, out of necessity, be incomplete as well as self-enforcing. Clearly, this will limit the impact of the agreement but the threat of future punishment for violations by the respective parties (signatories?) may deter 'asocial' behavior. It is necessary to explain however, why such an understanding may benefit from the inclusion of an *explicit* dispute settlement mechanism, such as the DSU. We must also ask why the parties cannot rely only on an implicit understanding of the agreement.

One might assume that the fairly extensive economic literature on trade agreements should provide explanations for the rationale for such agreements. However, the bulk of this literature focuses on the determinants of the tariff levels that a cooperative outcome would yield, rather than on the legal form that such an outcome would be packaged in. Indeed, the formal structure of many of the repeated game models employed in this literature, is basically identical to the structure of those employed to analyze collusion in product markets. As often emphasized in the latter literature, there need not be any

difference from an economic point of view between a collusive outcome resulting from implicitly coordinated price setting, and an outright cartel. That is, in these models there is typically no particular economic role played by the fact that the agreement between the firms is explicit, rather than implicit. For the same reason, economic models of trade agreements do typically not distinguish between tariff reductions resulting from an implicit understanding between trading partners, and resulting from an explicit agreement. From an economic point of view, the question remains as to why one form is chosen over the other.

The two central aspects of trade agreements mentioned above provide them with different possible roles. A first role stems from the fact that in order for the agreement to be self-enforcing, there must be a common understanding of what constitutes a cooperative behavior, and what amounts to a violation of the implicit agreement. However, for very much the same reasons that gave rise to the incompleteness of the contract, it is extremely difficult for the trading partners to implicitly form a common understanding. An explicitly agreed format for the interpretation of the agreement, and for the resolution of disputes, may therefore help achieve coordination in at least two ways. First, the case law it produces may gradually fill the “gaps” in the agreement, thus slowly making the agreement less incomplete. However, in practice, adjudicating bodies seem skeptical about playing an active role in this regard, viewing it as impermissible judicial activity, whereby the judge takes the role of the legislator. Also, there seems to be a fear that when creating case law, future decisions may be constrained in unforeseeable ways. The second way in which an explicit agreement can ease the coordination problem, is by specifying an agreed-upon *procedure* for adjudication. The legislators thus avoid the difficulty of working out the details of the contract by agreeing to let a third party - a judge, for example - adjudicate on their behalf; the legislators accept the outcome of the ruling as the outcome of the unfinished negotiation.^{5 6}

⁵ The “tie-breaking” task could actually be performed by any method that is agreed upon in advance. For instance, an agreement to flip a coin would serve this purpose. But there would then be no relationship between the characteristics of the case at hand, and its outcome. It is thus not likely to have particularly beneficial within-dispute impact, even if it would help keep the agreement together.

⁶ Another coordination role a DS system could serve would arise if members that are not participants in a dispute could participate in the withdrawal of concessions. The advantages of such systems are examined

The second reason why there is scope for an explicit DS mechanism also stems from the complexity of the issues at stake: There will be a very large number of measures that fall in a grey zone between what is clearly allowed and what is clearly illegal. The decision of where to draw the boundary cannot be done with any degree of scientific precision and it would therefore be tempting for the parties to interpret the scope of the agreement with an eye to only their own interests. The potential for this kind of moral hazard problem suggests that adjudication should be compulsory, and that the adjudication process should be performed by a disinterested party, so that the adjudication gains a *reputation for transparency and impartiality*. An explicit agreement greatly eases this task for third parties.

A third reason for an explicit agreement stems from the fact that members can be expected to be both complainants and respondents. For instance, a simple way of avoiding gaps in the coverage of the contract would be to let e.g. the complainant unilaterally decide the course of action whenever the agreement left this open. However, members know *ex ante* the signing of the agreement that due to the unclear undertakings in the contract, they will over time sometimes find themselves in the complainant's position, and sometimes in that of the respondent. As noted by Ethier [2001a], it is therefore likely that each member has an interest in ensuring that while there are punishments for violations of the agreement, that these are not too strong.⁷ The DSU has indeed chosen an intermediate form of punishment, with remedies that should be (at most) commensurate with committed illegalities. Given the incentives for complainants to meter out very strong countermeasures, there are advantages in letting a third party determine what is

by Maggi [1999], and by Bagwell *et al* [2004]. See Lawrence [2003] for a comprehensive discussion of the role of retaliation (countermeasures) in the WTO.

⁷ It could possibly be argued that it is not necessary for countries to potentially find themselves in the roles of both complainants and respondents, in order for them to agree on limited punishments. Even if their roles as complainants and respondents would be predetermined, they might in an *ex ante* negotiation over the rules of the DS mechanism agree on limited punishments, realizing that draconian punishments, as desired by complainants, as well as very limited punishments, as desired by respondents, may deter countries from liberalizing trade. Whether or not this is the case is likely to depend on the exact bargaining format.

commensurate in any particular situation – this may be a role for a formal dispute settlement mechanism to play.

Finally, a formal DS system may ease the problem of enforcement problem by helping to instill in members of the agreement a sense of “international obligation”. This notion has considerable support particularly in the international relations literature, but has been much less commonly examined in the economics literature (a noticeable exception being the work of Kovenock and Thursby [1992]). Similarly, a third-party evaluation of alleged breaches of the agreement is likely to have a stronger name-and-shame effect, than if the parties themselves were to determine inconsistencies and adjudicate conflicts, as emphasized by e.g. Schwartz and Sykes [2002].

The above-mentioned benefits of a DS mechanism must of course be set against its costs. There are the obvious costs in terms of administration, even though it is far from clear that a centralized mechanism is more costly than a decentralized system; if anything, the opposite seems more likely. But there are also less obvious drawbacks associated with an explicit dispute resolution mechanism in a self-enforcing agreement, due to the fact that it may weaken the forces maintaining system integrity.

The weakening of the incentives to retaliate may come about for at least two reasons. One is illustrated by Hungerford [1991], who adapts a “Green and Porter model” to a trade agreement context.⁸ In Hungerford’s model, countries have access to tariffs as well as non-tariff barriers (NTBs). Tariffs are assumed to be readily observable by members, as are their terms-of-trade. However, members cannot directly observe other members’ NTBs, nor can they directly observe other exogenous random events that may affect their terms-of trade, such as changes in world market prices, consumer incomes, etc. Hence, an exporting country hence cannot tell whether bad terms-of-trade are due to unfortunate random events, or to NTBs imposed by partners. Countries cannot form an explicit agreement to reduce these barriers, since NTBs cannot be directly monitored. But Hungerford [1991] shows that absent a DS system, there is nevertheless an equilibrium

⁸ Similar frameworks have also been employed by Riezman [1991] and Furusawa [2003].

where members abstain from invoking NTBs, due to the repeated nature of the interaction. This equilibrium has the undesirable feature however, that when a member suffers a sufficiently severe terms-of-trade shock, the source of which could hence either be unobservable NTBs or unobservable demand shocks, it will have to punish the partner by withdrawing a concession, and increasing thus its observable tariff temporarily, regardless of its origin. The reason is that if members did not systematically do so, there would be incentives for partners to cheat on the agreement. In this equilibrium one would thus over time witness periods where countries trade “peacefully”, interspersed with temporary “trade wars” resulting from adverse external events for importing countries.

What effect would a DS system then have? An essential feature of a DS system is that it typically requires members (as well as adjudicating bodies) to investigate the reasons for the shift in terms-of-trade, before withdrawing concessions. If these investigations fully reveal these reasons, all unnecessary punishments could be avoided. Hence, *by ensuring that information is gathered, the DS system reduces the need for misplaced retaliatory action designed to maintain the integrity of the agreement.* This would be a desirable consequence of a well-functioning DS system, closely related to the one mentioned above, related to the benefit of a transparent third-party adjudicator.

But a DS system is not only an information gathering device, it typically also *imposes restrictions on members’ rights to take unilateral action* when they perceive to have been cheated. This maybe necessary to induce countries to enter an agreement, as noted above. , as noted above. However, it also implies that the force for maintaining the integrity of the agreement is weakened or removed. To see why this may have severe consequences, consider in the context of the Hungerford [1991] model, the extreme situation where the investigation required by the DS system is completely *un*-informative. In this case the existence of a (compulsory) DS system will only serve to weaken the trade agreement, and would thus reduce the joint welfare of the parties.

The conclusion is hence that it is crucial that if a DS system imposes restrictions on members’ ability to retaliate, that it also offers a reliable mechanism for information

extraction. More generally, as also pointed out by e.g. Staiger [1995], *there is an inherent conflict between the desire to maintain rule integrity and the desire to facilitate the resolution of disputes.*

A second special feature of the DSU is that it encourages a *negotiated settlement*:

A solution mutually acceptable to the parties to a dispute and consistent with the covered agreements is clearly to be preferred. (3.7 DSU)

Laudable as it may seem to seek to cushion conflicts between the parties, it provides another reason why an explicit DS system may threaten system integrity. This aspect is demonstrated by Ludema [2001], who reasons (very roughly) as follows: The DS system opens a door for bilateral negotiations, should a conflict arise. Once members find themselves in such a situation, they tend to have incentives to resolve the dispute without unnecessary delay, forgetting past grievances.⁹ But this softening of the repercussions of violations of the agreement is foreseen at an earlier stage. That is, by forcing members to try to resolve conflicts in a “civilized manner”, the threat of such conflicts loses some of their deterrence, and thus reduces the incentives for members to avoid conflicts.¹⁰

Yet another drawback of an explicit DS mechanism is highlighted in Guzman [2002], who attempts to explain why not all international agreements have mandatory dispute resolution mechanisms. The core notion in Guzman’s [2002] theory is the assumption that naming-and-shaming causes net joint costs to the members of the agreement: the reputational loss to the losing respondent is not compensated for through an equal gain to for the other party. Against the gains from improved system integrity in situations where members abide by the agreement where they absent the naming-and-shaming otherwise would not, must thus set the costs caused in cases where there anyway is no compliance.

⁹ One can here note the fact that the majority of disputes that are notified to the WTO never reach the panel stage.

¹⁰ Without drifting into the issue of remedies, let us just note that several authors, such as Schwartz and Sykes [2002], and Ethier [2001b], suggest that the purpose of the DSU is indeed to *limit* the bite of remedies.

IV. THE TWO MAIN THEMES IN THE EMPIRICAL LITERATURE

We now turn to empirical research on dispute settlement in the GATT/WTO. There is a very large body of writings on this general topic, far too large to be surveyed in a single paper. We will in this section discuss papers focusing on what we see as the two main themes in this area: the determinants of participation in disputes, and the impact of DS settlement in the GATT/WTO on the process (and particularly likelihood) of settlement. Due to space limitations, we cannot provide detailed descriptions of studies on these topics, but only say a few words about the main issues addressed, the methods employed, and the results obtained.

Due to the complexity of the object of study, and the resulting inadequacy of theory, it is inevitable that any study on these issues can be criticized on a number of grounds. We therefore tend to see the results to be reported more as suggestive of empirical relationships, than as “proofs” of the existence (or non-existence) of these relationships. For this reason we refrain from criticizing individual papers, and instead at the end of the section state what we see as desirable directions in which to develop the literature more generally.¹¹

IV.1 What determines participation in disputes?

A frequent allegation in the policy debate over the working of the DSU is that participation in the DS mechanism is biased to the disadvantage of poorer/smaller countries. These claims take various forms: for instance, it is argued that developing countries do not launch complaints as frequently as they should, or that they are targeted more frequently by richer countries than they should be. There is an empirical literature that seeks to examine the correctness of these claims, and that attempts to highlight the determinants of participation in the DS system more generally.

¹¹ There are many similarities between this paper and the survey by Busch and Reinhardt [2002].

IV.1.1 A basic issue: what is the unit of account?

Intuitive as these claims may seem, the literature confronts several severe conceptual problems. We will return to some of these below, but let us first briefly examine the issue of choosing the *unit of account*. In order for a claim such as “developing countries do not complain as often as they should” to be meaningful, there must be a way of counting participation, and the manner in which this is done may have an important effect on the outcome of the investigation. A simple solution is of course to say that whenever there is a consultation request registered by the WTO Secretariat (as indicated by a new Dispute Settlement number being assigned by the Secretariat), there is a dispute. This is indeed the path that has been followed in much of the descriptive, quantitative legal literature. Such an approach would be based on a number of (typically implicit) assumptions. For instance, the *Bananas* dispute DS27, which involved five countries as complainants, and in which four additional countries requested to join consultations, would count as just one dispute. An alternative might be to consider this case as involving five (or nine) *bilateral* disputes.

There are also other complications. For instance, what essentially seems as the same dispute as DS27 was filed earlier by four of the five countries, as DS16. Should this count as independent dispute(s)? Furthermore, a request for consultation may involve a very specific aspect of a very specific measure, or it may address a number of aspects of a number of different measures. However, in both cases the investigation would count as only one dispute. It is easy to identify questions for which these features of the unit of account would be clearly undesirable, but it is typically much more difficult to determine a satisfactory definition of “a” dispute. There is no generally correct way of defining a dispute. Rather, what is appropriate depends on the exact question at stake. It should also be noted that the problem of choosing unit of account is not restricted to studies of participation, but it affects any study that seeks to draw inferences across disputes.

IV.1.2 Trade interests as a determinant of participation

As already mentioned, the determinants of participation in the WTO DS system has been examined in a number of papers. Horn *et al.* [1999] focus on the question of whether

participation as complainant in the WTO DS system is biased to the disadvantage of smaller and poorer members, in the sense that they complain less often than they “should”. Clearly, in order to address this issue, there is a need for a definition of an unbiased benchmark. For instance, it is highly likely that a country that exports many products to many markets will encounter more illegalities than a country that mainly exports one product to one market. Consequently, the former should have more issues to complain about, but how much more? Furthermore, larger traders are more likely to trade products in such large volumes as to make it profitable to carry the partly fixed costs of litigation, and should therefore be expected to be more often featuring as complainants. Should this be taken into account in the definition of the unbiased benchmark? As can be seen, the definition of the unbiased benchmark is far from trivial.

Horn *et al* [1999] start from the premise that the unbiased benchmark should allow members to complain proportionally to the number of questionable trade measures they encounter. Lacking a convincing theoretical prediction for the number of illegalities committed per country, and hence faced by trading partners, they assume that countries commit illegalities for each imported product with the same frequency, regardless of the nature of the exporting country or of the product. Using data for the first four years of the WTO DS system, and with products defined at the 4-digit HS level, Horn *et al* [1999] show that the actual distribution of bilateral disputes across members are fairly well predicted by their suggested non-biased benchmark, in particular when the latter is adjusted in order to exclude exports with smaller values (assuming that such values are not worthwhile to litigate about).

Bown [2005] substantially refines this analysis. As noted, countries can choose to pursue disputes by themselves, they can participate as co-complainants or as third parties. Or they can decide to not participate at all, free-riding on the efforts of other countries, or perhaps, on the contrary, being hurt by the litigation. Bown [2005] considers the determinants of these choices on the basis of the 116 disputes during 1995-2001 in which importing countries were found to illegally restrict imports. Disputes are divided into two separate sets, depending on whether they concern discriminatory measures, or non-

discriminatory measures. For each of the disputes involving discriminatory measures, exporters to the market (defined at the 6-digit HS level) are divided into two groups: those that are harmed by the measure, and those that benefit (by being exempted, for example). For approximately half of the group of disputes involving discriminatory measures, Bown [2005] further identifies other countries that were also harmed by the measure, but who did not participate in the legal process. This data is then employed in a multinomial logit model in order to examine the impact on the propensity to complain, to be a third party or to free ride, of various factors that may plausibly affect participation.

With regard to disputes over measures that adversely affect many trading partners, it is shown that size of exports is positively related to the propensity to complain, in line with the finding of Horn *et al* [1999]. It is also positively related to participation as a third party, and negatively related to the propensity to free ride.

IV.1.3 “Legal capacity” and “power” as determinants of participation

Some notion of trade interest can go quite far in explaining the distribution of disputes across countries, but it is clear that it cannot provide the whole picture. Two intuitively appealing hypothesis have therefore been examined in the literature. According to the “Legal Capacity Hypothesis”, the lack of legal capacity of prevents poorer countries from participating as complainants as much as they “should”. The second hypothesis, dubbed the “Power Hypothesis”, holds that smaller/poorer countries complain less against larger/richer countries that they “should” due to their lack of “power”. Various reasons have been suggested, such that they do not expect to be able to enforce rulings, or that they fear a back-lash in other forms, such as loss of preferential treatment in trade, or some form of non-trade retaliation such as reduced foreign aid, or military assistance.

Several studies have highlighted the role of these suggested explanations for participation in the DSU, in particular. The two explanations have often been juxtaposed, even though they are by no means mutually exclusive. Horn *et al* [1999] make a simplistic examination of the two popular claims. Using the size of countries’ WTO delegations in Geneva as a proxy for countries’ legal capacity, Horn *et al* [1999] find some, but weak,

support for the notion that countries with more legal capacity litigate more, controlling for trade interests. To shed some light on the Power Hypothesis, Horn *et al* [1999] aggregate WTO members into four groups – G4, other OECD countries, developing countries other than LDCs, and LDCs – and consider whether the pattern of litigation between these groups suggested any bias, using the above-described definition of a non-biased benchmark. Contrary to what the Power Hypothesis would seem to suggest, they find that developing countries other than LDCs are *over*-represented as complainants against both G4 countries and against other OECD countries, and that they are *under*-represented as respondents against G4 countries, while the pattern as respondents against other OECD countries is more mixed. It is difficult to draw any strong conclusions concerning LDCs. The crude measure employed suggests that they are underrepresented both as complainants and respondents against developed countries, but the numbers involved are so small that that this finding is hard to interpret.

Bown [2005] also examines the role of the Legal and Power Hypotheses. Bown [2005] finds some, albeit weak, support for the Legal Capacity Hypotheses, in that the coefficients for the variables GDP per capita and the size of WTO delegations have the expected signs, even though not being highly significant. But Bown [2005] also finds strong evidence for some form of Power Hypothesis, showing that the nature of the bilateral relationship between the importing country and exporters plays an import role in determining whether to complain, act as a third party, or abstain, in disputes involving illegalities that have an adverse effect on a number of countries. It is thus shown that a high share of the respondent's exports going to a certain country makes it more likely that this country will be a complainant (and less likely that it will free ride). A possible interpretation here is that "power" matters in the decision to complain, since such a high share makes the enforcement possibilities stronger. However, this relationship holds also when considering only a subset of fairly large exporters, where there would intuitively seem to be less role for "power" to be at play. The interpretation thus seems to be either that this intuition is flawed, and that "power" is important also in the relationship between more developed countries, or that the relationship captures something else than what associate with "power".

Guzman and Simmons [2005] shed further light on the Legal Capacity- and Power Hypotheses- explanations of the determinants of participation. Guzman and Simmons [2005] interpret the Power Hypothesis to refer to the amount of power a member can exert outside the system (such as the withdrawal of aid), and do not include in this concept power exerted within the multilateral system (such as the amount of concessions that a complainant can credibly threaten to withdraw). Their data set is based on bilateral disputes in the WTO between 1995 and April 2004, as defined by requests for consultations. The study regresses the GDP of the defendant against a number of explanatory variables, and controls. The GDP of the defendant is interpreted as a measure of both its market size and its political power. A main explanatory variable is the GDP of the complainant. GDP is a natural measure of the (absolute) power of the complainant, and to the extent that this is an important factor driving the decision of whether to complain, one would expect there to be a positive relation to the GDP of the respondent: only economically large countries would challenge other large countries. But it should also measure a country's capacity to pursue disputes: a large country is likely to have more resources to use for disputes. If it plays an important role in this respect, one would expect to see a negative relationship between this variable and the GDP of the defendant: a capacity constrained exporter will concentrate on the disputes that involve the largest stakes, and they typically concern the large markets. Countries with more resources – a higher GDP – can also go after the smaller fry, and will thus be more likely to litigate against small countries.

Guzman and Simmons [2005] also include other proxies for legal capacity. In addition to the commonly employed variable capturing the size of countries' Geneva delegations, they include the number of embassies abroad, countries' non-military government expenditures, and an index for the quality of government bureaucracies drawn from International Country Risk Guide. They also use a number of controls.

Several OLS model specifications are run. One specification includes the size of complainants' GDP, which is found to be negatively and significantly correlated with the

GDP of respondents, thus supporting the Legal Capacity Hypothesis and refuting the Power Hypothesis. The remaining specifications exclude complainant GDP, but include GDP per capita, which is negatively related to the GDP of the respondent (although not always significantly so). They also include one of the other above-mentioned measures of legal capacity. In each case, the included variable comes out highly significant and with the expected sign, except for the measure of the quality of the government bureaucracy.

Positively and significantly related to the GDP of respondents is also the value of the imports by the complainant from the respondent. One interpretation of this finding would be that it supports the Power Hypothesis. However, Guzman and Simmons [2005] prefer to see this variable as a control, and interpret the finding to merely indicate that large respondents tend to export a lot. There are also a number of other controls that are significantly correlated with respondent GDP.

Overall, Guzman and Simmons [2005] see their results as supporting the primacy of the Legal Capacity Hypothesis over the Power Hypothesis as an explanation of the choice of respondents. More generally, they conclude that even though it is very difficult to determine a non-biased benchmark for developing country participation, these countries seem constrained by limited legal resources to sue the system as frequently as richer countries. Because of such constraints, developing countries are more selective as to which cases they challenge before the WTO. However, lack of “power” does not seem to be an important explanatory factor.

IV.1.4 Membership of a preferential trading arrangement as a determinant of (non)-participation

Several authors find that countries tend to complain less against other members of the same preferential trade agreement to which they themselves belong. For instance, Bown [2005] finds that a highly significant, and also quantitatively important factor is whether the potential complainant belongs to the same preferential trading arrangement as the country with the illegal import measure: exporters are much less likely to complain against countries belonging to the same preferential trade agreement, or act as third

parties in such disputes. Furthermore, it is shown that the importance of the importing country as a donor of foreign aid to potential complainants tend to vary positively with the propensity of the latter to abstain from participating in disputes, and negatively with the propensity to act as third party.

IV.1.5. The form of political governance as a determinant of participation

A different perspective on participation in the DS system is provided by Reinhardt [2000], who considers the role of the form of political governance for participation. The decision by a country to launch a dispute is the result of a domestic political process, and one should expect the nature of such a process to be highly dependent on the political institutions in the country, in particular since private parties do not have standing before the WTO, and the selection of conflicts to be brought to the WTO thus is made by government institutions or politicians. Reinhardt [2000] examines a number of aspects of this issue, one being whether democracies are more or less likely to complain before the WTO. A number of theoretical arguments can be made in either direction, so while it seems plausible that the political system may affect the propensity complain, the direction is unclear.

Reinhardt [2000] uses a rich data set comprising all 604 “bilateral” disputes that occurred during the period 1948-1998. The statistical models employ, in addition to indices for democracy, a number of explanatory variables capturing aspects of GATT/WTO members, and use various probit specifications.

A main finding is that the more democratic a state is, the *more* it will initiate disputes, controlling for the trading countries’ relative size, and for one country’s dependence on trade with the other. Furthermore, this effect is very strong, quantitatively speaking. Reinhardt [2000] also shows that not only are democracies more likely to initiate disputes, there is also a strong tendency for democracies to be targeted more often. The offered explanation is that democratic governments will be more susceptible to domestic pressure for protection, and will as a result be more prone to implement illegal measures, or not to implement agreed-upon liberalization, and thus be the targets of litigation.

Furthermore, a country is more likely to initiate disputes against other countries that account for a large share of the first country's imports and exports, and also against countries that depend on it for their imports and exports, partly in line with the finding of Bown [2005].

Reinhardt [2000] also considers the impact of the creation of the DSU, revealing that it had no significant impact for probability of dispute initiation between developed countries. The creation of the DSU did however significantly lower the probability of disputes being filed by developing countries, and it significantly increased the probability of a developing country being targeted. Reinhardt [2000] concludes that the rise in the number of disputes in the multilateral trading system is not the result of the introduction of the DSU, but stems from the underlying increased dependence on foreign trade, in line with the trade interest explanation of participation, and on the general democratization of the world.

IV.1.6 Retaliation as a determinant of participation

Casual observation suggests that countries occasionally complain in retaliation for previously being the target of complaints. This is a special case of the situation where complaints are not only based on the merits of the case, but also on the characteristics of the potential adversary. The literature contains a few tests of the prevalence of this type of behavior. For example, Reinhardt [2000] includes for this purpose a binary variables capturing whether in the previous year the respondent initiated a dispute against the complainant side. This variables is highly significant, indicating that a dispute in the previous year increases the probability of a dispute in the opposite direction the year thereafter with a factor of 55.

Bown discusses the role of retaliation in several of his papers. In Bown [2002] he explores the possibility of violation of the agreement by a country facing and not facing retaliation (or threat thereof). Bown [2004a] finds substantial evidence showing that threat of retaliation by the victorious complainant yields credibility so as to allow defendants to honor their commitments. In a related paper, Bown [2004b] finds that

developing countries have recognized the importance of retaliatory threats, and have responded by changing their pattern of initiation of disputes, so as to take better advantage of the instances where they have leverage to threaten retaliation and thus induce compliance.

At a more disaggregated level, Blonigen and Bown [2003] find that the threat of a retaliatory antidumping investigation makes it less likely that a WTO Member will name the country that will likely retaliate among the countries that will be investigated for alleged dumping practices. They also find that the prospect that a particular WTO member might launch a complaint (in any field of the WTO), makes it less likely that a member investigating dumping practices might end up with a positive finding of injurious dumping against companies originating in the country threatening with a complaint.

IV.2 What determines the duration of disputes?

The majority of the requests for consultation that are filed with the WTO are not determined by WTO adjudicating bodies, but by the parties alone. The majority of disputes thus lead to either an officially announced (as requested per the DSU) Mutually Agreed Solution, or simply remain inactive indefinitely, and is thus presumably solved. As mentioned above, the DSU sees a MAS as the preferred mode of resolving disputes, so a settlement is from this point of view desirable; this view is understandable in that a decision to move the dispute to a panel stage constitutes an escalation of the dispute. Such a move is likely to increase the stakes both by increasing the direct costs of administering the proceedings, and the associated opportunity cost from the use of legal and administrative resources for the particular dispute, as well as indirect costs (and possibly also benefits) associated with its impact on the reputation of the participants. In this regard, avoiding an escalation of the dispute is desirable.

But there are also other aspects to take account of. To start with, the problem of defining the unit of account is as pervasive here as in studies of participation. Another aspect is the selection of conflicts that lead to requests for consultations. Presumably they are not a random selection, but tend to be special in some sense. What are these special

characteristics? How come that the conflict could be solved once it reached the consultation stage, but not before? Why does the interaction between the parties change, as the dispute is filed with the WTO? From a theoretical point of view, the answers to these questions are far from obvious, even though one might intuitively identify some explanations.¹² There are also other aspects of settlements that one may wonder about, such as the terms of the solution when these are not made public, for instance, to what extent do they abide by the Most-Favored Nation provision?

Several empirical papers have examined various aspects of the time profile disputes, and in particular, the propensity for settlement, during the GATT and the WTO periods.

IV.2.1 Does the DS mechanism ease settlement?

A first issue of interest is *whether the DS mechanism actually eases settlement?* One way to approach the issue is to examine whether the introduction of changes to the DS system that can be expected to enhance the system can be shown to actually lead to more settlement? Busch [2000] takes such an approach when studying the impact of the 1989 Dispute Settlement Procedures Improvement reform, which provided for the right to a panel. The study estimates several logit models employing data on bilateral disputes for the whole GATT period. Three separate binary dependent variables are included, indicating whether partial or full remedies are agreed during the consultation phase, whether the dispute was paneled, and whether there are concessions during the panel stage. The independent variables include a dummy for whether the year is before or after the 1989 Improvement, and a variable capturing the dyad's joint democracy score. A number of additional explanatory variables are used, indicating for instance, the number of complainants joining the dispute, whether it is brought by a developing country against a developed country, the degree of trade dependence, and the trade openness of the parties.

¹² For instance, Busch and Reinhardt [2001], Reinhardt [2001], and Guzman and Simmons [2002], discuss theoretical aspect of settlements. Busch and Reinhardt [2006] argue that third-party participation undermines pre-trial negotiations. They test their hypothesis using their data set and find that third party participation lowers the prospect for early settlement.

A main finding by Busch [2000] is that the 1989 Improvement, which allegedly sharpened the DS mechanism, did *not* foster more concessions under either the consultation or panel stage. The study also found that respondents with a larger share of trade in GDP tend to settle less under the consultation stage, contrary to what might perhaps be expected.

Busch and Reinhardt [2003b] present empirical evidence suggesting that the advent of the DSU significantly improved the propensity of respondents to concede, when considering aggregate numbers, even though this depiction is not valid for developing country complainants, or for disputes between the EC and the US. However, the more favorable picture during the WTO years does not stem from an increase in early settlement due to the introduction of the DSU, but is instead argued to be the result of the expanded scope of actionable cases, and more rich country complaints against developing countries. During the WTO era, richer complainants (in terms of GDP per capita) have been more likely to induce settlement than poorer countries, controlling for differences in GDP. But contrary to what one might assume however, the authors argue that this is not because richer complainants find it easier to induce compliance, nor is that poorer countries disproportionately lose disputes, but instead because they are less successful at inducing other countries to settle.

The WTO impact on the lifespan of disputes are highlighted by Grinols and Perrelli [2006]. They develop a theoretical model that predicts that the DSU should lead to more, and to shorter, disputes before the WTO. In order to empirically investigate these predictions, the authors examine three types of disputes, all of which involving the US, by means of various forms of duration analysis¹³: USTR Section 301 disputes 1975-2000, GATT disputes 1975-1994, and WTO disputes 1995-2000. An initial finding shows that the increase in the number of disputes during the WTO era cannot be readily explained by

¹³ The main application of duration analysis in economics has been to labor market issues, such as the movement in and out of unemployment, but have recently found applicability to a number of economic phenomena where there is movement in and out of different groups. But these methods have not been employed elsewhere to the context of dispute settlement, as far as we are aware. Grinols and Perrelli [2006] provide brief explanations of a number of the parametric, semi-parametric, and non-parametric duration analysis methods they employ.

the expansion of membership that occurred during this period – most respondents that the US has litigated against during the WTO era were also members of the GATT. Nor can the trend be explained by increased trade volumes; something else must lie behind it, such as the change in the nature of dispute resolution under the DSU.

Grinols and Perrelli [2006] indeed find that the advent of the WTO had a positive and significant impact on the number of multilateral disputes (still involving the US), but that it did not significantly affect the number of Section 301 cases. The reason seems to be an increased propensity by the USTR to use the GATT/WTO dispute resolution mechanism. Grinols and Perrelli [2006] also shows that the average lifespan of disputes are significantly shorter since the advent of the WTO, even though the more exact impact is rather complex and depends on the nature of the dispute.

IV.2.2 When is settlement most likely?

A second issue that has attracted interest in the literature is *when during the process is settlement most likely?* Reinhardt [2001] uses two ordered probit models, applied to data on GATT disputes initiated between 1948 and 1994. The dependent variable is an ordinal measure indicating whether the respondent conceded fully, partially or not at all, to the demands by the complainant.¹⁴ One model runs this variable against a few variables capturing mainly when the panel was established, and the direction in which the panel outcome went, whereas the other model includes these variables and a large number of other variables capturing various characteristics of the countries involved, the measure at stake, and the dispute itself (most of which turn out to be insignificant).

A striking feature of both these models is that the establishment of a panel makes concessions significantly more likely. As expected, a ruling in favor of the defendant makes it less likely with a concession, but even more surprisingly is that this is also true in case it goes *against* the respondent. The publishing of the panel verdict as such tends to make it less likely with concessions (this effect is significant in the first model only). It is

¹⁴ Much of this classification stems from Hudec [1993], but it has also been revised and updated by Reinhardt [1996], Reinhardt [2000], and Busch [2000].

the *threat* of an adverse ruling for the respondent, rather than the verdict itself, that induces a settlement.¹⁵

The study by Busch [2000], while not focusing directly on this issue, is consistent with the findings by Reinhardt [2001]. Busch and Reinhardt [2003b], considering concluded GATT/WTO disputes between 1980 and [2000], bring further evidence to bear on this issue. They show that the reason why these countries are less prone extract concessions is not that they lose dispute more frequently compared to richer members, but their inability to take advantage of the pre-panel publication stage for settlement. This in turn is explained by lack of legal capacity, rather than power for enforcement.

IV.2.3 The role of political governance for the propensity to settle

A third theme in the literature is *whether there are differences between democratic and less democratic states in their propensity to settle?* A number of arguments can be made in either direction. Examining this question empirically, Busch [2000] finds that during the GATT period, disputes between democracies were more likely to be settled during the consultation stage, compared to when either of the parties to the dispute were less democratic. The pattern did not persist however once a panel had been constituted. Busch [2000] also shows that countries with a large trade to GDP ratio were less likely to settle both before and after the constitution.

Guzman and Simmons [2002] examine the role of democracy for the propensity to settle, but from a different angle. Their argument is based on the assumption that transfer payments among states are costly, and that trade measures that are of an all-or-nothing nature, are likely to be harder to settle on than those concerning measures of a more continuous nature. A natural solution to this indivisibility problem is to introduce some form of side payment or broaden the negotiation by introducing additional issues. Guzman and Simmons [2002] argues however, that democracies will find it more

¹⁵ Reinhardt [2001] also develops an incomplete information model that shows, among other things, why the threat of a ruling by an adjudicator, rather than the ruling itself, may induce compliance.

difficult to do this, since the opposition to the broadening that will come from the exposed sectors will be much harder to withstand in a democracy.

In order to examine the empirical validity of their theory, the authors the extent to which the “lumpiness” of the issues involved in disputes can explain whether disputes lead to panels or not. To this end, all WTO disputes are classified into one of two groups: those addressing continuous measures (tariffs, non-zero quotas, or subsidies), and those involving discontinuous measures (bans, health and safety regulations, product classification issues, and absence of required laws). A number of controls are also introduced. A first set is intended to capture the relative power of the parties to the dispute, and include dummies for whether it is a developing country against a developed country, etc, various measures based on GDP, and the complainant’s dependence on the respondent’s market (hypothesizing that the larger this dependence, the more prone will the complainant be to accept a proposed solution). There are also controls for institutional factors, such as whether the complainant and respondent are both parliamentary rather than presidential, the idea being that parliamentary governments tend to be less exposed to protectionist legislative pressure. Another institutional control captures whether the countries involved are democracies. Since democratic governments are likely to find it more difficult to withstand protectionist pressure, they should be expected to be more prone to leave decisions for panels, rather than settle themselves. Finally, Guzman and Simmons [2002] also control for general trade dependence, arguing that more trade dependent states should be more prone to take a dispute to a panel in order to obtain a clear ruling.

The above-mentioned variables are used in a number of logit models. Generally, speaking, the results support the notion that the lumpiness of the issue affects paneling decisions, but the relationship is more complex than was hypothesized, since it interacts with the nature of the political structure in a complex fashion. Both lumpiness, and the degree of democracy, tends to *reduce* the propensity to panel, thus contradicting what was expected to be the case. However, the combination of a lumpy issue and a democratic pair tends to have a positive impact on the paneling propensity. There is also

a certain tendency for large complainants to settle more often, but this seems unrelated to the relative size of the complainant and the respondent as such.

V. AREAS IN NEED OF FURTHER RESEARCH

The discussion so far sought to portray the main themes in the theoretical and in particular the empirical literature. We will end by pointing to some of the weaknesses/lacunae we see in this literature.

V.1 Weaknesses in the empirical work

Although it is difficult to synthesize the empirical literature highlighted here, the general picture emerges as follows. Export values, or the diversity of exports, go a long way toward explaining the distribution of the number of disputes that have reached the WTO. But developing countries still seem to be at a disadvantage in their propensity to act as complainants. The Legal Capacity Hypothesis finds support, while the picture is more mixed with regard to the Power Hypothesis. Finally, the nature of the mode of government is important, with more democratic countries seeming more prone to be involved in disputes on either side.

So what remains to be done? As already mentioned, the issues under study in this literature are highly complex, and it uses a very large brush when painting its (normally non-formalized) theory. There are therefore a number of problems in this literature that need to be resolved in order for the findings to become more than just suggestive. In what follows we will briefly point to some of the problems we see (the order is not meant to indicate the relative importance we attach to these issues).

(i) A first issue is the choice of *unit of account*. There is indeed an awareness in the literature of the importance of this matter. In particular, several studies use other definitions, capturing various aspects of the bilateral nature of disputes. However, we are not aware of any study that seriously contemplates what is “one” issue in a complaint. Was the *Banana* dispute about one issue – such as the EC banana import regime – or was

it about several issues, such as the distribution system, quantitative import restrictions, etc? Or, to take the *EC - Sardines* dispute: was it about the labeling of sardines, or about the role of international standards, or both? More generally, we are not aware of any attempt to derive the definition of “one” dispute from any underlying theory. At the same time, we add up number and seek to draw inferences on the basis of these numbers. In our view, this problem is sufficiently severe to lead us to seriously question the meaningfulness of the whole literature on “bias” in participation. At the very least, one would like to see much more systematic analysis of the sensitivity of the findings to the choice of unit of account.

(ii) Another critical issue for studies of biases in participation is obviously the *definition of the non-biased benchmark*. For instance, the notion employed by Horn *et al* [1999] that in an unbiased situation, each country would complain in proportion to how often it encounters illegalities, as long as the trade values involved exceed some lower threshold, is highly dubious in that it ignores the fact that the a market with a \$1 million turnover may be as important to a small country, as is a market with a \$100 million turnover to a 100 times larger country. Here, there is an urgent need for formulations that are much better grounded in theory.

(iii) A somewhat related problem in this context is how to deal with the *indivisibility of disputes*. For instance, if a group of countries according to some measure should have .4 disputes, but has none, is this to be treated symmetrically to the situation where a country should have 9.4 disputes, but has 9?

(iv) The proxies for legal capacity, in particular, seem very crude. There is a need to identify exactly what type of capacity is needed, and when. For instance, is it the capacity to detect illegalities, or to litigate once they are detected? How do we take account of the fact that even developed countries tend to hire private counsel when litigating? Perhaps this suggests that it is not legal capacity, but the lack of budgetary resources that is missing? Thus, there is also a need for clearer conceptual view of what the differences might be between rich and poor countries, as well as better proxies for whatever is

decided to be the appropriate variables to capture. There is also a need to refine the proxies employed to measure power, even though we see less of a problem here. In particular, one would want to see measures that are more sensitive to the nature of the bilateral relationship between countries, than those commonly employed.

(v) Closely related to the question of how to define an unbiased benchmark, is that of why and when countries commit illegalities? For instance, is it typically done in order to defuse domestic political pressure that otherwise would pose a more serious threat to the country's ability to maintain its commitments, are illegalities mainly the result of obscure agreements, or of aggressive pursuit of national (or interest group) interests? Understanding such questions is crucial to the formulation of an unbiased benchmark for the propensity to litigate, among many other questions: for instance, if small countries face proportionally more illegalities than larger countries, a non-biased benchmark should possibly require that small countries use the DS system more than proportionally. There are a few papers that provide theoretical explanations for when illegalities may be committed, such as Bown [2002], Büttler and Hauser [2000], Grinols and Perrelli [2003], and Guzman [2003]. However, there is little empirical work that sheds light on this issue. An interesting first step has been taken by Bown [2004b], who examines the determinants of countries' choices of whether to violate or adhere to GATT rules when making trade policy changes during rounds. But more theoretical and empirical work on these issues is highly desirable.

(vi) A more general issue is *how to interpret the selection of requests for consultations?* To date, a little more than 300 hundred such requests have been filed in the WTO, but it is inconceivable that this would represent the totality of grievances that WTO members have had with other members during the 10 years since the advent of the WTO. On the contrary, what has been registered with the WTO Secretariat is not just the tip, but the tip of the tip of the iceberg. This raises the extremely important questions for this literature of whether we can by studying these relatively few disputes, draw *any* inferences about the working of the DS system, for all those markets and trades where no complaint is filed? If we believe this to be the case, then how do we explain the fact that

these particular conflicts ended up as formal disputes at the WTO, while other conflicts did not? That is, what determines the *selection* of disputes that appear before the WTO? Perhaps there is something special about them, and it is for precisely this reason that they are brought to Geneva? There are reasons to suspect that the disputes in the WTO are not representative of conflicts in general. While a member can always decide unilaterally to litigate, in most cases such a decision will be preceded by contacts between the two sides to the conflict. It is only when they both decide not to give in that there will be a consultation request. Similarly, the DSU requests a period of consultation before proceeding to the panel stage. Since litigation uses substantial resources on both sides, one would normally expect the parties to settle before this stage, unless their subjective probability distributions over outcome of the litigation diverge significantly. The disputes we see in the DS system, at least those that reach the panel stage, are hence those where both sides find it worthwhile to participate in costly litigation. Indeed, it seems implausible that this would be the case for the bulk of trade conflicts. The registered disputes are thus most likely different from other conflicts. The question of *how* they differ remains unanswered.

(vii) A closely related issue is how to interpret the observation that *a certain group of countries* have (relative to some benchmark) launched few complaints? This may be due to the fact that the mere threat of complaints from this group has sufficed to keep its trading partners from invoking policies that the group would complain about? Or maybe, on the contrary, this group distrust the system to such a degree that it does not find it worthwhile to pursue disputes, or maybe it lacks the resources to identify issues that would be worthwhile to contest, or to litigate about issues it has identified? Our intuition may suggest one answer or another to these questions, but this does not suffice if we want to make methodologically more well-founded claims concerning the effect of the system.

(viii) Yet another problem is the fact that what is ultimately of interest is presumably the extent to which the *whole DS system* provides sufficient benefits to, for instance, developing countries. It is possible, at least in theory, that, countries that are not active complainants (or third parties) still benefit from the efforts of more active members. If a

member A successfully attacks an import measure maintained by member B, member C who happens to export the same product may also benefit from A's victory. Hence, if we are to evaluate the beneficiaries of the system in general, we have to take into consideration these indirect effects. Differently put, it seems likely that there will often be positive externalities from complainants to other exporters, at least as long as rulings are implemented respecting the Most-Favored Nation principle. One should therefore also expect there to be problems of free-riding between members. And it is not inconceivable that there may also occasionally be negative spillovers. Regardless of the direction of these externalities, they need to be taken account of in an assessment of the pros and cons of the DSU relative to some benchmark.

V.2 Why an explicit agreement in the context of an incomplete agreement?

An obvious question concerns the role of an explicit (as opposed to an implicit) mechanism. We have so far tried to focus on work that examines explicit mechanisms, but it is not always clear why the highlighted insights would not *also* apply to informal mechanisms. Intuitively, it would clearly be impossible to coordinate on an implicitly agreed DS mechanism with as much detail as DSU. What is less clear however, are the features of the situation that would arise absent the explicit agreement. There is here very little formal literature to lean against.

A second area that we feel is poorly understood is the role of dispute settlement in the context of an incomplete contract. Intuition would suggest that the problems facing a DS mechanism, and its role, would be very different if the agreement only contained easily verifiable commitments, such as tariff bindings, compared to the situation where there are a number of other provisions, such as that of National Treatment. We would thus like to see more analysis of the optimal design of dispute settlement in the context of incomplete contracts.

V.3 Does the DS mechanism serve its purpose?

A difficult but crucial issue is whether dispute settlement mechanisms in the GATT and the WTO have served their purposes? As was seen above, the literature on settlements

can be viewed as an attempt to answer this question, assuming that the purpose is to foster settlement. But as we have seen, there may be a conflict between fostering settlement, and maintaining system integrity. Consequently, it is possible, at least as a theoretical proposition, that by enhancing settlement, the DS system in the e.g. the WTO has reduced trade liberalization.

This leads us to the basic question of what is the real purpose of WTO, and for whom? Economists, at least, would almost by reflex assume that the purpose is to liberalize trade, and the Preamble to the *Agreement Establishing the WTO* also mentions substantial tariff liberalization (and non-discrimination) as a purpose. At the same time, there is a real issue whether trade is much more liberal today than it would be absent a multilateral agreement. For instance, would trade barriers be significantly higher between say the EU and the US absent the WTO? If it could not be shown that the WTO have had a significant liberalizing effect for its main players, could the purpose of the agreement be better understood as being something else than trade liberalization? Indeed, in the policy discussion, proponents of the WTO often mention stability of rules, transparency, etc, as main objectives of the agreement. Others may argue that a main purpose of the GATT, at least, was more political than narrowly economic.

The reason for mentioning this issue in this context is that the role of the DS mechanism might be very different if the purpose is to improve transparency, or the stability of economic or political relationships rather than to foster trade liberalization. Whenever we evaluate the achievements of the DSU, we cannot avoid to openly or implicitly take a stance on the question of what the agreement is to achieve in the first place.

V.4 The quality of adjudication

The discussion so far has dealt primarily with quantitative aspects of dispute settlement in the GATT and the WTO, even though we have occasionally touched more qualitative question such as the possibility for countries to settle before paneling. As noted above, the resolution of disputes may indeed be the main purpose of a dispute settlement mechanism. However, it seems reasonable to assume that its purpose is not only to induce

settlement, but also to promote a desirable form of implementation of the agreement. This raises a new set of issues concerning both the terms at which settlement occurs, and in particular about the *qualitative nature of determinations by the adjudicating bodies*.

Of course, decisions from the adjudicating bodies of the GATT/WTO have been subject to much analysis in legal literature. There seem to be fewer attempts to discuss these issues from the point of view of a joint legal and economic perspective, however. But an assessment of the case law would be very partial if performed from a legal perspective only: it can hardly be denied that a main purpose of the GATT, and perhaps even more the WTO, is to achieve various economic aims, as is also stipulated in the Preambles of the Agreements. Whether decisions by the adjudicators contribute to achieving these aims cannot be evaluated without an economic analysis of the case law.

The literature does contain some joint economic and legal analyses. For instance, Sykes [2003a] discusses the WTO case law on safeguards and concludes that inherent problems with these provisions have been exacerbated by a lack of acknowledgement, by panels and the Appellate Body alike, of the problem before them. Sykes [2003b] examines the case law on the necessity-test, the notion that when deviating from their obligations, WTO members have to ensure that they choose the least restrictive (in terms of impact on international trade transactions) means to reach their ends. Sykes [2003b] argues that it is typically hard to detect any systematic criteria being employed by adjudicating bodies. Horn and Mavroidis [2004a] discuss the case law on tax discrimination cases. They come to a similar qualitative conclusion, being unable to discern what method of analysis the adjudicating bodies employ.¹⁶

¹⁶ The studies from the ongoing American Law Institute project *Principles of International Trade Law: The World Trade Organization* also largely support this critique of the quality of the case law. In this project, a group of economists and lawyers have been scrutinizing all Appellate Body reports, as well as un-appealed panel reports, issued since January 1, 2001, from a joint law and economics perspective. While not always disputing the outcome of the determinations, a very common finding is that decisions are extremely deferential to the words used in the WTO agreement, which are often read in clinical isolation of their context, that is, without WTO judges asking, and answering, the question of what function any given legal instrument has been assigned to play. Rulings are also very often criticized for lacking economic logic. These studies are published in Horn and Mavroidis [2004c], [2005] and [2006]. See www.ali.org.

To conclude, the few attempts of putting the adjudication in the WTO into a joint economic and legal perspective, tend to be rather critical of the methodological side of rulings. However, most of the studies of this type provide *ad hoc* analyses of either specific rulings, or less frequently, the case law under certain provisions. There is therefore a need for much more systematic analyses of the quality of the case law.

V.5 Evidentiary standards

A special aspect of the quality of the case law are the evidentiary standards employed by the adjudicating bodies. Horn and Mavroidis [2004b] discuss one special aspect: the role of the burden of proof in disputes involving the interpretation of National Treatment. In their view, the burden has too easily been shifted over to respondents, with the consequence of leading to too many “convictions”. Grossman and Sykes [2006] consider some principle aspects of another facet of evidentiary standards – the rule of waiver. They examine how such rules may affect the number of claims brought before the adjudicators, and thus also litigation costs.¹⁷

The distribution of the burden of proof (production), as well as the associated required level of persuasion, are critical for determining the ambit of the provisions of many of the central provisions of the WTO. This will become increasingly the case since, as a result of the continuing reduced relevance of protection through classic trade instruments, disputes arise more and more frequently between one informed (the regulating party) and one uninformed (the challenging) party. This process will become even more pronounced if the Appellate Body starts practicing what it says it will: to treat regulatory intent as part of the standard of review in such cases.

V.6 The role of the judge

As a final area in a far from exhaustive list of lacunae in the literature, we would like to mention the question of the role of the judge. For instance, the European Community has recently submitted a proposal to replace the existing regime of *ad hoc* panels with one of

¹⁷ Yet another aspect of the evidentiary standards concerns the appropriate role of scientific expertise.

permanent panelists. This proposal has been suggested to have a dramatic effect if implemented. However, we are not aware of any research from a law and economics perspective of how the composition of the adjudicating bodies may affect systematically affect rulings.

REFERENCES

Bagwell, Kyle, Petros C. Mavroidis and Robert W. Staiger [2003]. The case for auctioning countermeasures in the WTO, *NBER Working Paper Series, No 9920*.

Bloningen, Bruce A. and Chad P. Bown. [2003]. Antidumping and retaliation threats, *Journal of International Economics*, 60: 249 – 273.

Bown, Chad P. [2002]. The economics of trade disputes, the GATT's Article XXIII and the WTO's dispute settlement understanding, *Economics and Politics*, 14: 283 – 323.

Bown, Chad P. [2004a]. On the economic success of GATT/WTO dispute settlement, *The Review of Economics and Statistics*, 37: 678 – 720.

Bown, Chad P. [2004b]. Developing countries as plaintiffs and defendants in GATT/WTO trade disputes, *The World Economy*, 27: 59 – 80.

Bown, Chad P. [2005]. Participation in WTO dispute settlement: complainants, interested parties and free riders. *World Bank Economic Review* 287-310.

Busch, Marc L. [2000]. Democracy, consultations, and the paneling of disputes under GATT, *Journal of Conflict Resolution*, 44: 425 – 446.

Busch, Marc L. and Eric Reinhardt [2001]. Bargaining in the shadow of law: early settlements in GATT/WTO disputes, *Fordham International Law Journal*, 24: 158 – 172.

Busch, Marc L. and Eric Reinhardt [2002]. Testing international trade law: empirical studies of GATT/WTO dispute settlement, pp. 457 – 481 in *Daniel Kennedy and James Southwick (eds.), The political economy of international trade law, essays in honour of Robert E. Hudec*, Cambridge Un. Press, Cambridge, Mass,

Busch, Marc L. and Eric Reinhardt [2006]. Three's a Crowd: Third Parties and WTO Dispute Settlement, *World Politics* (forthcoming).

Bütler, Monika and Heinz Hauser [2000]. The WTO dispute settlement system: A first assessment from an economic perspective. *Journal of Law, Economics & Organization*, 503,504.

Ethier, Wilfred J. [2001a]. Punishments and dispute settlement in trade agreements, Copenhagen: EPRU.

Ethier, Wilfred J. [2001b]. Theoretical problems in negotiating trade liberalization, *European Journal of Political Economy*, 17, 209-232

Furusawa, Taiji [2003]. The role of the WTO Dispute Settlement Procedure on international cooperation. Mimeo.

Grinols, Earl L. and Roberto Perrelli [2003]. The WTO impact on international trade disputes: an event history analysis. *The Review of Economics and Statistics* (forthcoming).

Grossman, Gene M. and Alan O. Sykes [2006]. European Communities – Anti-Dumping Duties on Imports of Cotton-Type Bed Linen from India Recourse to Article 21.5 of the DSU by India. Forthcoming in Horn and Mavroidis [2006b].

Guzman, Andrew T. [2002]. The cost of credibility: explaining resistance to interstate dispute resolution mechanisms, *Journal of Legal Studies*, 31: 303-326.

Guzman, Andrew T. [2003]. The political economy of litigation and settlement at the WTO. Mimeo.

Guzman, Andrew T. and Beth Simmons [2002]. To settle or empanel? An empirical analysis of litigation and settlement at the WTO, *Journal of Legal Studies* 31, S205 – S235.

Guzman, Andrew T. and Beth Simmons [2005]. Power plays and capacity constraints: the selection of defendants in WTO disputes, *Journal of Legal Studies* 34, 557 - 598.

Hoekman, Bernard and Petros C. Mavroidis [1994]. Competition, competition policy and the GATT, *World Economy* 13:

Horn, Henrik and Petros C. Mavroidis. [2004a]. Still hazy after all these years: the interpretation of National Treatment in the GATT/WTO case law on tax discrimination, *European Journal of International Law*, 15: 39-69.

Horn, Henrik and Petros C. Mavroidis. [2004b]. The allocation of burden of proof in non discrimination cases before the WTO. *Mimeo*.

Horn, Henrik and Petros C. Mavroidis (eds) [2004c]. *The WTO case law of 2001; The American Law Institute Reporters' studies*, Cambridge University Press: Cambridge, UK.

Horn, Henrik and Petros C. Mavroidis (eds) [2005]. *The WTO case law of 2002; The American Law Institute Reporters' studies*, Cambridge University Press: Cambridge, UK.

Horn, H. and P. Mavroidis [2006a] “The WTO Dispute Settlement Data Set: User’s Guide”. Available under www.worldbank.org/trade.

Horn, Henrik and Petros C. Mavroidis (eds) [2006b]. *The WTO case law of 2003; The American Law Institute Reporters' studies*, Cambridge University Press: Cambridge, UK.

Horn, Henrik, Mavroidis, Petros C. and Hakan Nordstrom [1999]. Is the use of the WTO dispute settlement system biased? CEPR Discussion Paper Series No 2340. Also in *The*

Law and Economics of the WTO Dispute Settlement System, Petros C. Mavroidis and Alan O. Sykes (eds), Edward Elgar Publishing, 2005.

Hudec, Robert [1993]. *Enforcing international trade law: The evolution of the modern GATT legal system*, Salem, New Haven: Butterworth Legal.

Hungerford, Thomas L. [1991]. GATT: A cooperative equilibrium in a non-cooperative trading regime, *Journal of International Economics*, 31.

Kovenock, Dan and Marie Thursby [1992]. GATT: Dispute settlement and cooperation, *Economics and Politics*, 4(2), 151-170

Lawrence, Robert Z. [2003]. *Crimes and punishment in the WTO*, Institute of International Economics: Washington D.C.

Ludema, Rodney D. [2001]. Optimal international trade agreements and dispute settlement procedures, *European Journal of Political Economy*, 17(2), 355-376

Maggi, Giovanni [1999]. The role of multilateral institutions in international trade cooperation, *American Economic Review*, 89(1), 190-214

Mattoo, Aaditya and Petros C. Mavroidis [1997]. Trade, environment and the WTO: the dispute settlement practice relating to Art. XX of the GATT, pp. 325-344 in Ernst-Ulrich Petersmann (ed.), *International trade law and the GATT/WTO dispute settlement system*, Kluwer: Amsterdam.

Palmer, David and Petros C. Mavroidis [2004]. *Dispute settlement in the WTO, practice and procedure*, Cambridge University. Press: Cambridge, UK.

Reinhardt, Eric [1996]. *Posturing parliaments: Ratification, uncertainty, and international bargaining*, PhD dissertation, Columbia University, Department of Political Science.

Reinhardt, Eric [2000]. Aggressive multilateralism: the determinants of GATT/WTO dispute initiation 1948 – 1998. Mimeo.

Reinhardt, Eric [2001]. Adjudication without enforcement, *Journal of Conflict Resolution*, 45, 174-195.

Riezman, Raymond [1991]. Dynamic tariffs with asymmetric information, *Journal of International Economics*, 30(3-4), 267-284

Schwarz, Warren and Alan O. Sykes [2002]. The economic structure of renegotiation and dispute resolution in WTO/GATT system, *Journal of Legal Studies*, 31: 179 -

Staiger, Robert [1995]. International rules and institutions for cooperative trade policy, Ch. 29 in Grossman, Gene M. and Kenneth Rogoff (eds), *Handbook of international economics, Volume 3*, 1495-1551, Amsterdam: Elsevier.

Sykes, Alan O. [2003a]. The safeguards mess: a critique of the WTO jurisprudence, *World Trade Review*, 2: 261 – 304.

Sykes, Alan O. [2003b]. The least restrictive means, *University of Chicago Law Review*, 70: 403 – 416.