CHAPTER 12

The Theory of Dispute Resolution with Application to Intellectual Property Rights

Mostafa Beshkar and Eric W. Bond

Department of Economics, Vanderbilt University, VU Station B#351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA
E-mail address: mostafa.beshkar@vanderbilt.edu
E-mail address: eric.bond@vanderbilt.edu

Abstract

We survey several of the theoretical models that have been applied to the analysis of the GATT/WTO dispute settlement process. These include repeated game models, which emphasize the punishment aspect of dispute settlement, and incomplete contracting models, which emphasize the “gap-filling” aspect. Our analysis emphasizes the implications of these models for the strengthening of the dispute settlement process under the WTO and for its application to the TRIPS agreement. We also discuss how models of settlement bargaining can be applied to obtain empirical predictions about which cases will actually proceed to an actual finding by the dispute panel.

Keywords: Trade agreements, WTO dispute settlement, TRIPS

JEL classifications: F13, F53, F55

1. Introduction

The World Trade Organization (WTO) features a dispute settlement process (DSP) that has come to play a key role in the organization’s trade agreements. A major change in the General Agreement on Tariffs and Trade (GATT) in its transition to the WTO has been the strengthening of the DSP mechanism to make it more likely that rulings of the WTO panels on disputes are carried out in a timely fashion. The new DSP has been actively utilized since its inception in 1995, with more than 350 complaints brought to the WTO by the end of 2006. The effectiveness of the DSP is one of the features that sets the WTO agreements apart from other international agreements. It is often cited as one of the factors permitting the effective inclusion of agreements on intellectual property rights (IPR) as part of the WTO.

This process has not been without its critics. One concern with the DSP is that it is still relatively ineffective at punishing countries that violate their WTO obligations,
even with its recent strengthening. Disputes often carry on for years and defendants can escape having to pay any penalties until the point at which the judgment has been made and the appeals process completed. On the other hand, it has been argued that the flexibility of the system is necessary for countries to be willing to engage in significant trade liberalization. The potential for swift and certain punishment for any country that deviates from a previous promise might cause countries to be more cautious about making significant cuts.

A second concern has to do with the impact of the dispute settlement process on the developing countries. Although participation by developing countries in the system has increased, the costs of bringing a case are disproportionately high for them. The fraction of complaints brought to the WTO by developing countries is smaller than the proportion those countries represent in the WTO, and the least-developed countries have not participated in the process. This suggests that developing countries may be underrepresented in the DSP.

Our objective in this chapter is to provide a review of the theoretical literature on trade agreements as it relates to the settlement of disputes, and to use this literature to address the questions that have been raised about the DSP. In particular, we emphasize the role of the DSP in relation to the WTO’s Agreement on trade-related aspects of intellectual property rights (TRIPS). We begin in Section 2 with a brief summary of the main features of the Dispute Settlement Understanding that was negotiated as part of the Uruguay Round. The Uruguay Round also incorporated the TRIPS Agreement, and we discuss some of the issues that are specific to the application of dispute settlement to its subject matter. In Section 3 we provide a benchmark model of an international agreement that resolves a Prisoner’s Dilemma between the member countries, and illustrate how this basic model can be used to explain the benefits of international agreements over both tariff rates and IPR protection. In particular, the motivation for bringing together agreements on trade and IPR is discussed.

The Prisoner’s-Dilemma model provides a setting in which the role of dispute settlement can be discussed. We highlight two approaches to modeling dispute settlement, repeated-game models and incomplete-contracting models, which illustrate the tension between enforcement and flexibility in trade agreements. Repeated-game models emphasize the requirement that since international agreements lack a third-party enforcer, they must be supported by threats of future trade punishment. The dispute settlement process is a natural venue in which to find this enforcement power. However, the simplest repeated-game models suggest that punishment should be swift and severe, which seems to contradict experience with the WTO process. We discuss how extensions of the repeated-game models have been used to provide explanations of why punishments may be limited.

Incomplete-contracting models, on the other hand, emphasize the fact that there may be states of the world in which the prescribed actions in a trade agreement are not efficient. The incompleteness of the contract is typically thought to arise due to the cost of including terms in the contract to cover all of the possible states that may occur. A potential role for dispute settlement in this case is to make it easier for the parties to obtain an efficient outcome. This approach supports the wisdom of the flexibility
of the DSP. We then discuss the extent to which the rules of the WTO process can be interpreted as making it easier to obtain more efficient outcomes.

In Section 4 of the paper we analyze the DSP from the point of view of settlement bargaining. The question here is to explain how the rules of the DSP will affect the incentives of the parties to initiate disputes and whether to settle them prior to the panel ruling. These models provide predictions about how the probability of settlement is related to the litigation costs of the parties and the features of the case. In particular, these models can be used to determine whether the observed behavior of developing countries as defendants and plaintiffs is consistent with the hypothesis that they are handicapped by high costs of litigation. Section 5 offers some concluding remarks.

2. The WTO dispute settlement procedure and its application to intellectual property

The current operation of the DSP is governed by the dispute settlement understanding (DSU) that was negotiated as part of the 1995 Uruguay Round Agreements that founded the World Trade Organization. Our purpose in this section is to highlight the main features that models of the DSP have sought to explain. We also provide a brief discussion of the specific application of these procedures to disputes under the TRIPS agreement.

2.1. The dispute settlement process

Article XXIII of the GATT gave member countries the right to initiate a complaint against another member that had taken actions which “nullified or impaired” the benefits resulting from the agreement. Complaints take one of two forms: a member country has taken actions which are either directly in conflict with its obligations under the WTO agreements (a “violation case”) or a member has taken actions which have the effect of undermining benefits accruing to a member country under the agreements (a “non-violation case”). The DSU encourages countries to resolve the dispute by consultations. If a mutually satisfactory agreement to the issue cannot be reached, then the complaining country can request the formation of a panel of experts to evaluate the complaint. If multiple countries make the same complaint against a member, these complaints are combined and considered by a single panel. Third parties that have an interest in a complaint may provide statements for consideration by the panel.

The dispute panel evaluates the submissions of the parties for their consistency with the WTO agreements and issues a finding of fact. The report of the panel is similar to the ruling of a court in common-law systems, in that it can cite precedent from previous cases in reaching its decision. In the event that a member is found to have taken measures that violate the agreement, the panel includes a recommendation on how these measures should be altered to bring them into compliance. The respondent has the right to appeal the ruling to a sitting Appellate Body. If the panel report is approved by the member countries and the respondent fails to comply with the report...
after a reasonable period of time, the complaining party has the right to compensation
in the form of the withdrawal of equivalent concessions. This means that the complain-
ing party can suspend its obligations to the respondent in an amount that is equivalent
to the amount of nullification or impairment it has suffered. The DSU specifies that
compensation is intended to be temporary, in that it will only last until the respondent
country’s policies are brought into line with its WTO obligations.

The DSU involves two major attempts to strengthen the dispute settlement process
above the one that had operated earlier under the GATT. The first is the elimination of
the requirement of unanimity in order for a panel report to be implemented. Under the
GATT process the formation of a panel or the acceptance of a panel report could be
blocked by any member, including the defendant. Srinivasan (2006) notes that reports
on a number of sensitive cases were blocked in the period 1987–1989, including all
five cases brought under the plurilateral subsidies code. In contrast, under the DSU
the adoption of the panel report is automatic, unless the members unanimously decide
not to adopt the report. The second major change has been the setting of specific time
requirements for each stage of the dispute settlement process in order to speed it up.
Under the GATT procedure, a case could carry on for as long as a decade. The new
WTO procedure attempts to ensure that the process proceeds more quickly by setting
specific time limits at each stage. For example, a panel has six months from the time
that it is established to complete its report, with a period of nine months being allowed
for exceptional cases.

In practice, these changes have been a mixed success. With regard to attempts to
speed up the process, Horn and Mavroidis (2006) report that these guidelines have
not always been met. The average length of time from establishment of a panel to
the issuance of its report has been over a year, with 18 of the cases taking more than
500 days. Furthermore, Srinivasan (2006) argues that new regulations concerning the
timing of negotiations have had the effect of replacing the lengthy pre-trial negotia-
tions by lengthy post-trial negotiations and arbitrations. However, the changes have
resulted in substantially more usage of the dispute process as illustrated in Figure 1.
Under the GATT, there were an average of 4.5 complaints filed per year from 1980 to
1994. From the beginning of the WTO process to 2003, an average of 38.7 complaints
per year have been filed.

To summarize, there are several features of the DSU that stand out. First is the
emphasis on getting members to comply with their obligations under the WTO agree-
ments, preferably by consultations without having to resort to establishment of a panel.
Punishment is to be used only as a last resort, and is intended to last only until the
defecting country comes into compliance. In practice, only about 40 percent of the
311 complaints brought under the DSU in its first 10 years of operation have reached
the stage of having a panel report issued. A second point is that the DSU discourages
unilateral actions. Even if a country is certain that a violation has occurred its grievance
will not be recognized by the WTO until a finding has been issued by the panel. There-
fore, a defendant can respond to such unilateral actions by initiating a counter-dispute
against the complaining party. Third, rather than facilitating the imposition of punish-
ments against the offending party, the DSU tries to limit retaliation as the outcome of
disputes among member countries. This reveals an important difference between the
conventional court systems and the WTO dispute settlement mechanism. While con-
ventional courts act as external entities for enforcement of contracts, the DSU plays
the role of an arbitrator which helps the contracting parties avoid the breakdown of
cooperation.

2.2. TRIPS and dispute settlement

Prior to the formation of the WTO, there were international agreements covering a
variety of forms of intellectual property. These included the Berne Convention (copy-
rights), the Paris Convention (patents, trademarks and industrial designs), the Wash-
ington Treaty (integrated circuits), and the Hague Agreement (industrial designs). The
World Intellectual Property Organization (WIPO) was established by treaty in 1967
and became a specialized agency of the United Nations in 1974. Its essential tasks are
to administer these various treaties on intellectual property and to foster cooperation
between member states.

The inclusion of the TRIPS Agreement in the WTO had two primary impacts on
the protection of intellectual property.\(^1\) The first was to raise the standards for in-
tellectual property protection worldwide, both by strengthening many of the rules
governing IPR in existing agreements and by expanding the application of global re-
quirements to all WTO member countries.\(^2\) For example, a minimum standard of 20

\(^1\) Among the many published descriptions of TRIPS standards and the need for policy reforms, see Maskus
(2000).

\(^2\) TRIPS is a foundational agreement of the WTO and is not within the ambit of WIPO.
years for patent life was set and the range of intellectual property covered was expanded. In addition to requiring that laws be made consistent with these minimum standards, TRIPS also required countries to have adequate enforcement procedures to allow owners of IPR to take action against infringers. The application of TRIPS to all WTO members required the rewriting of many national laws on IPR protection. These changes were particularly significant for developing countries, many of which had lax or non-existent protection of intellectual property. As a result, developing countries were given a five-year transition period to meet their TRIPS obligations and the least-developed countries were allowed a 10-year transition period.

The second important feature of the TRIPS Agreement was the application of the WTO dispute settlement mechanism to disputes involving IPR. Previous agreements either did not specify a dispute settlement mechanism, or had ineffective mechanisms that were not used in practice. The Paris Convention on patents included a provision that allowed disputes to be taken to the International Court of Justice. However, Emmert (1989) notes that a majority of the member states had not recognized the compulsory jurisdiction of the International Court of Justice for patent issues and that, in practice, there were no disputes taken before the Court. Similarly, enforcement issues were not seriously addressed in the Berne Convention on copyrights and many thought that protection offered by the majority of members was inadequate. Emmert (1989) summarized the dispute settlement process of these agreements as “effectively worthless.”

In addition to providing a structure for resolving disputes over IPR protection, the linking of TRIPS with the WTO trade agreements offers a more effective means of providing compensation for an injured party in broader trade disputes. On some occasions, it will be difficult to identify a means for making adequate compensation based solely on merchandise trade flows. For example, suppose that a complainant country exports to the respondent’s market but imports relatively little from that country. It may be impossible to find tariff concessions that can be suspended by the complaining country to provide adequate compensation. This is particularly true of IPR protection, because many developing countries are primarily importers of technology rather than exporters. If a developed nation were to suspend application of its patent or copyright rules for them there would be little effective impact.

The DSU specifies that when choosing compensation, the complainant should first consider concessions that are in the same sector as that in which the injury occurred. If this is impossible, then it may suspend concessions under another sector within the same agreement (e.g., trade in goods under the GATT or commerce in services under the General Agreement on Trade in Services, the GATS). If neither of these options yields appropriate punishment and the circumstances are serious enough, the complainant can seek to withdraw concessions under another covered agreement. In the

---

3 WIPO has a process for settling disputes over IPR claims between parties from different jurisdictions. However, this process is for private parties, which do not have standing before the WTO and, as such, is not a substitute for the WTO process.
case of a violation of TRIPS, a sector refers to a particular type of intellectual property. This would mean, for example, that a violation of patent law would be compensated first by suspension of obligations having to do with patents. If adequate compensation cannot be found under patent obligations, then compensation on another form of intellectual property (such as copyright) can be proposed. If adequate compensation under TRIPS is not available, then obligations on goods trade could be suspended. Similarly, suspension of TRIPS obligations could be used as compensation for nullification of benefits on goods trade.

In principle, this flexibility in compensation can operate effectively in both directions. First, it can provide a means for countries that export technology to enforce agreements with technology-importing countries by denying them market access for some of their exports of goods. Second, it can provide a means for a small country that imports intellectual property to obtain compensation when it is denied market access by a technology-exporting country. For example, in its dispute regarding the banana-import restrictions of the European Union, Ecuador was ultimately permitted to suspend application of certain components of its copyright regime to EU products, though it chose not to do so.

3. Modeling international agreements on trade and intellectual property

In order to understand the role of dispute settlement, we must begin with a theory of why countries pursue international agreements on trade and IPR.

3.1. Welfare externalities and the Prisoner’s Dilemma

Most theories are built around the notion that trade agreements are a means of resolving a Prisoner’s Dilemma between countries. In order to allow for the variety of trade and intellectual property policy instruments that can be covered in an international agreement, we illustrate the Prisoner’s Dilemma using a two-country model in which the home (foreign) country must each choose a level of a policy instrument \( a(a^*, s) \). The payoffs to the countries will be denoted by \( U(a, a^*, s) \) and \( U^*(a, a^*, s) \), respectively, where \( s \) is a state variable that affects the payoffs to the countries and is assumed to be observable to both of the countries. We provide specific interpretations of these actions and the state variable in terms of tariffs and intellectual property protection below.

The optimal policy choice of the home country is denoted \( \tilde{a}(a^*, s) \), which is the home-country action that maximizes its welfare given the foreign policy and the state of the world. The foreign country’s optimal policy is similarly denoted \( \tilde{a}(a^*, s) \). The Nash-equilibrium policies will be the values \( \{a_N(s), a_N^*(s)\} \) that satisfy \( a_N = \tilde{a}(a_N^*, s) \) and \( a_N^* = \tilde{a}^*(a_N, s) \). Since the Nash-equilibrium policies maximize each country’s welfare with respect to its own policy choice, this equilibrium will be inefficient as long as the policy choices of one country affect the welfare of the other. Specifically, if welfare of each country is increasing (decreasing) in the policy choice of the other country, both countries will gain by a mutual increase (decrease) in \( a \) in
the neighborhood of the Nash equilibrium. Absent this spillover, unilateral policies
will be optimal and there will be no need for a trade agreement.

In the case of trade policy, the policy instrument for the home (foreign) country
will be its tariff, $\tau(\tau^*)$. When the home country sets its tariff unilaterally, there are
two forces leading to a positive optimal tariff. First, there is the “terms of trade” effect
due to the fact that the home tariff depresses the price of the imported good when
the home country is large enough to affect world prices. The worsening of the terms
of trade has a negative impact on the rest of the world, so tariffs will be too high in
the Nash equilibrium because of their negative spillover to other countries. This is the
classic argument as to why tariff wars occur, and the implications of this behavior for
welfare in the non-cooperative equilibrium were first examined by Johnson (1953).

More recently, emphasis has shifted to introducing a “political economy effect,”
which is that politically powerful interests in the import-competing sector will influence
the government to provide protection against foreign competition. This approach
can help to explain why tariffs might be positive even in countries that are too small to
influence world prices and why tariffs may seem to protect organized interest groups.
Note however that under the political-economy models, the Prisoner’s Dilemma may
remain operative, which leaves room for mutual welfare-improving agreements. The
models of Grossman and Helpman (1995) and Bagwell and Staiger (1999) show that
the terms of trade effects continue to provide a negative spillover to the rest of the
world even when political effects are taken into account. In other words, even under
political-economy models, the use of tariffs features a beggar-thy-neighbor approach
because a country manipulates its trade policy to benefit its residents at the expense of
other societies (Jackson, 2000).

A similar Prisoner’s Dilemma applies to agreements involving intellectual prop-
erty as well. In this case we can denote the welfare of the home country as $U(T, T^*)$,
where $T(T^*)$ denotes the strictness of the intellectual property protection chosen by
the home (foreign) country. Protecting intellectual property involves a tradeoff be-
tween the static deadweight loss generated by granting a monopoly to the innovator
against the greater rate of innovation that results when innovations are more profitable.
With diminishing returns to innovation, an optimal degree of patent protection will ex-
ist at which $\partial U(T, T^*)/\partial T = 0$. Two caveats should be added to this argument in an
open economy, as recently shown by Grossman and Lai (2004). One is that each coun-
try’s national loss due to monopoly will be greater when the rights to the innovation
are owned by foreigners. The second is that patent protection in small countries might
have minimal national benefits, since the additional profits earned by innovators in
that market might be such a small share of worldwide profits that they would have
little impact on the world rate of innovation. As a result, countries that are primarily
importers of intellectual property or countries with small markets might have little or
no incentive to protect intellectual property absent an international agreement.

---

4 Bond (2005) provides examples of early national laws that failed to provide national treatment to foreign-
ners in the protection and enforcement of intellectual property rights, and also examples of how the strictness
An increase in foreign patent protection will have a favorable spillover to the home market, because it will lead to higher profits for home innovators that sell in the foreign market and will also raise the rate of innovation. Since the spillovers to overseas markets are not considered by countries in setting their patent policy, an agreement that expands protection in both countries in the neighborhood of the Nash-equilibrium level will yield a welfare improvement for both countries. Thus, the Prisoner’s Dilemma also provides an argument for international agreements to strengthen protection of intellectual property. It should be noted here that although the theory predicts benefits from mutual expansion of intellectual property protection, it does not guarantee that a harmonization of IPR rules will benefit all countries. In particular, small countries and countries that are primarily importers of intellectual property would expect to receive little benefit from agreements that required a substantial increase in their degree of protection. This would suggest that one of the main reasons for the inclusion of TRIPS in the WTO was to provide a way of inducing developing countries to adopt more stringent protection of intellectual property by trading concessions on market access, as argued in Bond (2005). For example, Helfer (2004) notes that “[d]eveloping nations agreed to include intellectual property within the newly created WTO in exchange for securing access to the markets of industrialized states for their agricultural products, textiles, and other goods.”

The above discussion identified the sources of gain from international agreements. An international agreement would be an efficient contract if it specified Pareto-efficient actions \( \{a(s), a^*(s)\} \) to be taken by each of the countries in each of the relevant states of the world. However, there are two problems with designing such contracts. The first is that since the action specified for each country is not a best response to the policy assigned to the other country, countries have an incentive to choose a different policy than the one specified under the agreement. Therefore, some form of enforcement is required to ensure compliance of all parties to the contract. The second is that if the number of potential states of the world is extremely large, then it may be too costly to write a contract that specifies the complete set of state-contingent actions. In this case the specification of actions will be incomplete and it will be necessary to identify what is intended if a set of circumstances occurs for which actions have not been anticipated.

of national laws was linked to the country’s market size and whether it was an importer or exporter of IPR. These examples are consistent with the argument that a country’s policy on IPR will be influenced by its terms of trade, since granting property rights to foreigners will raise the price of imported goods that enjoy protection. Such agreements could also be achieved if there were direct cash transfers from the technology-exporting countries to the importing countries in order to compensate them for extending and enforcing protection of intellectual property agreements. However, it seems more common for countries to trade concessions on different issues rather than to make outright transfers.

In the case of trade policy, the state-contingent actions might include safeguard protections that allow for alterations in the tariff that arise in the presence of an import surge which might fuel protectionist pressure. For intellectual property, the ability to issue compulsory licenses of pharmaceuticals in the event of a national emergency might be thought of as a state-contingent policy.
Within a particular country the role of enforcing and interpreting contracts typically is handled by courts that have the power to impose judgments on the parties in the event of a dispute. However, such an external enforcement agent is typically lacking in international agreements between sovereign governments. WTO obligations do not have the force of domestic law for many countries (e.g., the United States), so domestic courts would not automatically prohibit actions that violated such commitments. International courts could potentially rule on such disputes but do not have the jurisdiction to impose penalties. Within the WTO the primary threat against a country is the loss of future benefits under its trade agreements if it does not comply. For example, the notion of using threats of retaliatory tariffs to enforce trade agreements has been formalized using the theory of repeated games.

3.2. Enforcement of complete agreements with repeated interactions

The theory of repeated games has shown how cooperative agreements that yield payoffs that Pareto-dominate the Nash-equilibrium payoffs for both countries can be sustained using history-dependent strategies. An example of this is a trigger strategy, in which a deviation by either country triggers a reversion to a punishment phase that specifies actions \( a_P \) and \( a^*_P \) to be taken by the countries. Since we are focusing on complete contracts that specify actions for all states of nature in this section, we simplify the presentation by assuming that there is a single state. Therefore, \( s \) will be suppressed in specifying actions and punishments.

During the punishment phase, the home country receives a payoff \( U_P \equiv U(a_P, a^*_P) < U(a, a^*) \) and the foreign country receives \( U^*_P \equiv U^*(a_P, a^*_P) < U^*(a, a^*) \). Letting \( \delta < 1 \) denote the discount on payoffs received in the next period and assuming for simplicity that this punishment lasts forever, a deviating country will receive a discounted payoff of \( U(\tilde{a}(a^*_P), a^*_P) + \delta U_P / (1 - \delta) \). In order for the punishment to be an effective deterrent, it is necessary that it be credible in the sense that a deviating party would expect the punishment actions \( (a_P, a^*_P) \) to be actually carried out by all parties in the event of a deviation. Formally, this requires that the punishment payoff represents an equilibrium to the repeated game beginning at the time following a deviation. A commonly used assumption is that the punishment involves a return to the pre-agreement trade war between the countries, which is credible because it is a Nash equilibrium of the one-shot game.

The agreement will then be self-enforcing for the home country if it satisfies

\[
U(a, a^*) \geq (1 - \delta)U(\tilde{a}(a^*_P), a^*_P) + \delta U_P. \tag{1}
\]

A similar requirement applies to the foreign country. Since \( U(\tilde{a}(a^*_P), a^*_P) > U(a, a^*) > U_P \), condition (1) will be satisfied for \( \delta \) sufficiently close to 1. Let \( A(\delta, U_P, U^*_P) \) denote the set of agreements \( (a, a^*) \) for which (1) is satisfied given the discount factor and punishments. Any of the agreements in this set represents an equilibrium to the repeated game. This multiplicity of equilibria provides a theoretical role for the institutions of the WTO, for that organization can be thought of as designing a set of
rules by which countries choose an agreement from the sustainable set \( A(\delta, U_p, U_p^*) \). Presumably this would call for the WTO to design an efficient bargaining protocol, so that the agreement chosen is not Pareto-dominated within the set of enforceable agreements.\(^7\)

Kovenock and Thursby (1992) provide a stronger role for international trade agreements by assuming that a violation of a commitment made results in an additional cost being imposed because of the terms of an additional international obligation. They assume that this cost, which may be either a fixed cost or a per-period cost, results any time a country violates an agreement or the prescribed behavior during a punishment phase. They show that this additional international obligation cost makes agreements easier to sustain, but still may not allow countries to sustain free trade.

If one adopts the interpretation of the WTO as providing a coordinating mechanism for playing the repeated game, then one might also expect the WTO to choose its rules to make the set \( A \) as large as possible. For example, it is clear from Equation (1) that the set \( A \) will be non-decreasing in the value of \( \delta \), because a larger \( \delta \) makes an agreement easier to sustain by putting greater weight on the punishment phase. The WTO can increase \( \delta \) by reducing the time span, \( T \), between initiation of a dispute and imposition of necessary punishments. To see this, note that \( \delta = e^{-rT} \), where \( r \) is the market rate of interest. In fact, \( \delta \) can be interpreted either as a measure of patience of the contracting parties or the speed with which they can detect and penalize deviations. Obviously, \( \delta \) is decreasing in the length of time it takes a dispute panel to issue a report and to approve punishment against the deviator. The evidence based on experience seems mixed on this point. Since \( \delta \) is decreasing in \( T \), efforts by the WTO to accelerate the dispute settlement process should expand the set of sustainable agreements. However, the WTO requirement that countries wait until the dispute panel has ruled, rather than engaging rapidly in unilateral punishments, seems to conflict with this logic. In the case of a clear violation, it would seem that countries could adjust tariffs to punish deviators (e.g., by executive order in the United States) much more rapidly than could be accomplished by waiting for a complaint to work its way through the dispute process.

A second way for the WTO to influence the size of the set of sustainable agreements is through its choice of punishment levels. The lower the payoff received during the punishment phase (and the longer the period of the punishment), the larger will be the set of sustainable agreements. In general-equilibrium trade models there are typically two Nash equilibria: an interior Nash equilibrium with positive tariffs and trade flows and an autarkic Nash equilibrium with arbitrarily large tariffs imposed by each country. An infinite repetition of the latter outcome would be the worst sub-game perfect

7 Suppose that there are multiple states and that these are observable to the players so that specified actions can be state-contingent. In this case the no-deviation constraint will require that the future payoffs be expectations taken over the possible future states. This case is developed by Bagwell and Staiger (1990) for the case of symmetric countries where free trade is the Pareto-efficient outcome in all states. They show that the best sustainable agreement may involve contingent protection, where free trade is sustainable in some states but not in others.
Nash equilibrium. Thus, it should be the punishment chosen by an institution that is trying to support the largest set of trade agreements. However, clearly this prediction does not characterize the punishments prescribed in the DSU, since the general reluctance to punish deviators and the modest nature of the punishments available, such as withdrawal of equivalent concessions, are not consistent with the severest possible punishments.

Two approaches have been used to explain why the punishments of the dispute settlement mechanism are so mild when theory would predict strong punishments. One explanation arises from the possibility of renegotiation by the parties once a deviation has occurred. The repeated-game structure implicitly assumes that the parties choose strategies at the beginning of the game, and then mechanically follow them as the game unfolds. However, there is a certain implausibility to the assumption that the punishment could involve an infinite reversion to the autarkic Nash equilibrium if the parties can communicate. Once the punishment phase is underway, there will exist agreements that would leave both countries better off than they are during that phase. Thus, one would expect the countries to engage in renegotiation during the punishment phase.

This problem suggests that self-sustaining agreements need to render unnecessary the possibility of renegotiation. Ludema (2001) examines the types of punishments that are renegotiation-proof using a definition of that concept (due to Pearce, 1987), which requires that the punishment path be sub-game perfect and involve no punishment worse than itself. He shows that the renegotiation-proof punishment path involves play of the interior Nash equilibrium for a finite number of periods, followed by a return to the best supportable equilibrium. This punishment is less severe than the most severe sub-game-perfect punishment, which would be infinite reversion to the autarkic equilibrium.

Klimenko et al. (2006) obtain a more pessimistic conclusion about the ability of countries to use trigger strategies to sustain cooperation. They assume that the parties to the agreement meet each period and bargain over the set of feasible trade agreements using the Nash bargaining solution. With this bargaining framework no cooperation can be supported, because the unique equilibrium is the infinite repetition of the static Nash equilibrium. The reason is that since negotiations are repeated each period, the bargaining problem looks the same each time and it is not possible for the outcome to depend on the history of the bargaining game. However, an external entity, such as the WTO, can promote cooperation by labeling the history of the relationship as “cooperate” or “dispute.” This labeling ensures that the parties cannot jointly ignore a violation when it occurs. In their formulation, the role of international trade institutions is to change the initial conditions of the bargaining game at each period, based on the history of the relationship.

A second approach to explaining the mild sanctions of the dispute settlement mechanism is to argue that the WTO contains two types of punishment mechanisms corresponding to two different types of deviations. In the first type of deviation, a country abrogates most or all of its trade obligations at once. Although deviations of this type are not explicitly discussed in the WTO agreements, it is understood by members that
such a deviation would result in the breakdown of multilateralism and a return to a trade-war equilibrium. The threat of the breakdown of the system deters such deviations, but since this harsh punishment is an effective deterrent it is never actually observed as long as the system is an equilibrium.

In the second type of deviation, a country defects from some aspect of its obligations without intending to withdraw from the agreement. It is these small deviations that are handled by the WTO’s dispute settlement mechanism. One appeal of this approach is that it highlights the fact that the disputes handled by the WTO do not seem to have the features of cases in which a country deviates on all of its obligations at once.

Zissimos (2007) shows that the idea of having different types of punishments for different magnitudes of deviations can be used to explain why, historically, tariff reductions negotiated at the GATT and WTO have been gradual. He assumes that the contracting parties face a two-part punishment mechanism. If the deviation of one party from the original agreement is considered moderate or “non-abusive” it will be punished by withdrawal of equivalent concessions, whereas a large or “abusive” deviation will face punitive punishments through reversion to Nash equilibrium. He interprets an abusive deviation as one that is sufficiently large that it exceeds the previous period’s tariff binding. On the other hand, breaking the current period’s promised binding is assumed to be non-abusive and results in the withdrawal of equivalent concessions. With this punishment scheme, an immediate reduction of tariffs to the neighborhood of free trade is not sustainable because even relatively large deviations (i.e., anything less than a reversion to the initial one-shot Nash equilibrium tariffs) would be met only by modest punishment. However, low tariffs will be sustainable once the previous period’s binding has become sufficiently low, and free trade will be reached asymptotically.

To summarize, the emphasis in repeated-game models is on the effectiveness of threats of retaliation in preventing countries from reneging on promises made under a trade agreement. The DSU is a natural place to study the enforcement issue, because it explicitly links the withdrawal of concessions to violations of obligations.

An important common feature of these repeated-game models is that the punishments are set in such a way that the countries never actually deviate from the agreement. This behavior is common to both models with severe punishments and to models where punishments are restricted either by renegotiation or the withdrawal of equivalent concessions. However, the examples of disputes mentioned earlier illustrate that in practice countries choose policies that violate WTO agreements, suggesting that these models miss an important feature of trade policy.

8 Bagwell and Staiger (2002) describe the former as “off the equilibrium path” deviations and the latter as “on the path” deviations.

9 Riezman (1991) considers a model in which countries have private information about shocks, but punishments can only be imposed based on publicly observable information. In this case punishments will occur in some states of the world, but countries will never actually deviate from their prescribed strategies.
3.3. Incomplete-contracting models

Contractual incompleteness arises when the contract does not specify a complete set of policy actions for each state of the world. The WTO agreement is clearly an incomplete contract. One obvious form of incompleteness is that it places fairly limited restrictions on a government’s use of domestic policy instruments, even though they could be used to influence the volume of trade. Since tax policies, competition policies, and environmental policies affect domestic relative prices, they can potentially be used to alter the volume of trade or to reverse the effects of promises made on trade liberalization. A second form of incompleteness arises because there is limited flexibility to change the level of negotiated trade policies in response to changes in the economic environment. For example, tariffs are bound from above and cannot be raised beyond those bindings in response to a domestic recession.

If contracting is costless, then a contract will specify efficient actions \( \{a(s), a^*(s)\} \) in each state \( s \). When contracting is costly, however, it may be inefficient to describe actions that will be carried out in states that have low probability or are particularly difficult to describe. Therefore, contracting parties may prefer to write general, rather than detailed, terms regarding those states or simply ignore them, implying that the action specified for a particular state, \( s' \), may not be Pareto-efficient. If state \( s' \) occurs, there will exist actions for the countries that Pareto-dominate the action specified in the contract. It would be desirable in this state to let the countries renegotiate their actions, since one party can gain by deviating from the action specified in the contract even after compensating the other party for the deviation. This is referred to in the legal literature as an “efficient breach” of the contract, for it makes one country better off but leaves the other whole. One way to allow for efficient breaches is to adopt a compensation standard that requires the deviating party to be liable for any damages imposed on the other party to the contract.\(^{10}\)

Ethier (2001a, 2001b) argues that the dispute settlement process is a means of injecting desired flexibility into a contract in the presence of contractual incompleteness. He considers a scenario in which, prior to the signing of a contract, each of the countries recognizes that there is a possibility of a shock that will cause that party to choose a domestic policy action that has a negative spillover on the trading partner. He shows that if the countries choose contractual terms \( \text{ex ante} \), there are circumstances in which it is optimal to have an agreement in which punishment takes the form of the withdrawal of equivalent concessions and that the deviating country will choose to accept the punishment rather than reverse its policy. This result is analogous to the notion of efficient breach, since the parties are better off \( \text{ex ante} \) by allowing the deviating country to violate the agreement while compensating the injured party.

\(^{10}\) This contrasts with a specific performance standard, which would require the deviating party to carry out the actions specified in the contract.
Horn et al. (2006) also adopt an incomplete-contracting approach, although their emphasis is slightly different. They study types of government actions and states of the world that should be specified in a contract when contracting costs are related to both the number of states included in the agreement and the number of policy instruments it restricts. The potential policy instruments are both tariffs and domestic policies, such as production and consumption taxes, while there is a domestic consumption externality. If contracting were costless, an efficient agreement would equate producer prices in the two countries and use domestic policies to deal with the consumption externality. The emphasis in their results is on the interaction between discretion and rigidity in the optimal contract when contracting is costly. Rigidity refers to economizing on contracting costs by setting the same value of an included policy instrument for all states, whereas discretion refers to economizing by leaving a policy instrument out of an agreement. Discretion on a policy instrument allows a government to react to states that are not specified in the contract, but this discretion might also allow a government to indirectly undo restrictions imposed by the contract.

In their model, the Prisoner’s Dilemma between countries is driven by terms of trade externalities. The benefits of the agreement arise from restrictions on the use of policies that influence the terms of trade. Therefore, an agreement that restricts only domestic policies cannot be optimal because it is ineffective in dealing with the terms of trade problem. When contracting is costly, parties should focus on restricting policies with larger influence on the terms of trade and leave other policies to the discretion of individual governments. They also show that rigidity and discretion tend to be complements in the agreement when there is uncertainty about the level of the consumption externality, since the government needs to use its domestic policy instruments to respond to shocks to that level. However, rigidity and discretion are substitutes when shocks are to the level of demand (and hence trade volumes).

This analysis implicitly assumes that the contract is enforced as written, so no room is allowed for the DSP to assist in the adjustment of contract terms in states where the existing terms are not efficient. An alternative means of introducing flexibility into contracts without requiring an enumeration of all states is to tie contract adjustments to the outcomes of certain key variables. For example, the escape clause in trade agreements allows for temporary tariff protection in cases where trade liberalization results in substantial harm to the domestic industry. Contract incompleteness may imply a “gap-filling” role for the courts. Thus, when a dispute arises in an unspecified contingency, the court may provide an interpretation of the contract that sets out an efficient course of action to be taken by the parties. In a discussion of partnership contracts with contractual incompleteness, Shavell (2006) argues that there is an optimal method for interpreting agreements that will yield higher expected social welfare than can be obtained by enforcing them as literally written. If the court employs efficient interpretation rules, the parties will find it unnecessary ex ante to specify actions in as many states of the world. Therefore, the prospect of optimal interpretation by the court allows parties to conserve costs by writing simpler contracts. This would suggest that DSU guidelines might be set so as to enhance the efficiency of market-access negotiations.
3.4. Enforcement, flexibility and TRIPS

The repeated-game approach to describing the WTO emphasizes the enforcement role of the DSP, whereas the incomplete-contracting approach emphasizes the potential for flexibility. We conclude our discussion of these approaches by examining some of the evidence from IPR-related disputes and the potential role for the DSP to introduce flexibility into the WTO agreements.

Positions taken by the proponents of TRIPS in the negotiations during the Uruguay Round seem consistent with a significant enforcement role for the DSP. Furthermore, the experience with TRIPS supports the effectiveness of threats of trade retaliation. The dispute cases brought before the WTO involving IPR have focused almost exclusively on the incompatibility of national laws with the requirements of the TRIPS agreement. The U.S. case against Denmark was fairly typical. The U.S. complaint was that Denmark’s law failed to satisfy its TRIPS obligations because it did not allow for provisional measures in civil litigation involving IPR. The case was resolved by mutual agreement, with Danish law being changed to allow provisional measures. There has been only one case, also settled by consultations, involving the failure of a country to enforce its intellectual property laws. Overall, 13 of the 25 IPR-related cases filed under the dispute settlement procedure were resolved by mutual agreement. Nine of the cases resulted in panel reports being filed, with three of those being appealed.

For TRIPS cases, contractual incompleteness has (at least so far) seemed to play a less significant role as a source of disputes. The case that might best be interpreted as being about contractual incompleteness is the “bars and grills” case, in which the European Union challenged a U.S. law granting exemptions from paying royalties on music that was played in public places (restaurants, bars, and retail outlets) if the establishments met a certain maximum size requirement. The EU successfully argued to the panel that this exemption violated the TRIPS agreement because it was too broad and covered more than 70 percent of the restaurants and nearly half of the retail establishments in the United States. The issue in this case was the interpretation of the meaning of exemptions for “special circumstances” that are allowed under TRIPS. After the United States failed to change its law on this point within a reasonable period of time, the EU was granted compensation. Compensation was made by a cash payment from the United States to the EU, which marked the first time in which a cash payment


12 Greece – Enforcement of Intellectual Property Rights for Motion Pictures and Television Programs (DS-125, Complainant: United States). Television stations in Greece were broadcasting movies without authorization of copyright holders, and the complaint asserted that remedies to copyright holders were not being provided or enforced. A mutually agreed solution was reached that involved additional legislation to enhance enforcement and actions that resulted in the closure of four offending television stations. See also DS-124.

13 United States – Section 110(5) of the U.S. Copyright Act (DS-160), panel report adopted 27 July 2000.
was made as compensation in a WTO case. Specifically, the United States agreed to pay $3.3 million as compensation for the lost royalties of European music-rights holders during the three-year period 1996–1998. In return, the EU agreed to withdraw its petition temporarily. However, such monetary compensation is not a common practice in the WTO to this point in time.

The analysis of incomplete contracts raised the question of whether the DSU allows a country to adjust its policy to a more efficient one by paying compensation to the injured country. The first point to note is that the DSU encourages countries to find a mutually agreed solution through negotiations. Article 2 (7) of the DSU states that “A mutually acceptable solution that is agreeable to the parties and is consistent with the covered agreements is clearly to be preferred.” This would indicate that the agreement encourages the negotiation of post Pareto-improving adjustments in the agreement, which is consistent with the notion of allowing an efficient breach. However, these negotiated adjustments must still satisfy the basic principles of the agreement. This would preclude the countries from negotiating an adjustment that involved the use of voluntary export restraints, which are now prohibited by the WTO. Furthermore, the next sentence in the agreement indicates that “[i]n the absence of a mutually agreed solution, the first objective of the dispute settlement process is usually to secure the withdrawal of the measures concerned if they are found to be inconsistent with any provisions of any of the covered agreements. The provision of compensation should be resorted to only if the immediate withdrawal of the measure is impracticable and as a temporary measure pending the withdrawal of the measure that is inconsistent with the covered agreements.” This language suggests that the objective of the dispute panel should be to ensure that the parties carry out their obligations as outlined in the agreement, which would suggest a specific performance standard.

Although a withdrawal of equivalent concessions is allowed as compensation, the emphasis on the use of compensation as a last resort and its temporary nature would suggest that the intention is not to allow countries to purchase relief from obligations under the agreement. In practice there have been very few cases in which monetary compensation has been awarded. The bars and grills case is one of the few cases in which compensation has actually been paid by a country. In a number of cases the granting of the compensation has spurred the defendant country to change its policies to be in compliance with the WTO agreement.

### 3.5. Multilateralism and the WTO

The models discussed above focused on trade relations between two countries. These two-country models are useful for simplifying the problem, but omit interesting issues that arise in a multilateral setting. One such issue is the possibility of bilateral trade imbalance. If two countries have a bilateral trade imbalance, the principle of reciprocity in trade concessions would prevent them from attaining negotiated free trade. Under this situation, a multilateral trading system may provide proper incentives for further trade liberalization. Maggi (1999) analyzes a three-country model of a tariff
Mostafa Beshkar and Eric W. Bond

game with bilateral trade imbalances but a multilateral balanced trade. He shows that under bilateral trade imbalance, threat of a multilateral punishment leads to higher levels of cooperation. The idea is as follows: in an extreme case, suppose that country A only imports from B and only exports to C, and C only exports to B and only imports from A. Then bilateral punishments are impossible and, therefore, multilateral punishments become necessary to deter deviation. According to this framework, the role of the multilateral trading system is to detect violations of the agreements in bilateral relationships and to authorize interested third parties to impose sanctions against the offending party.14

Even though Maggi (1999) suggests a plausible role for a multilateral trade institution, his framework does not represent the reality of the WTO. The DSU does not recognize the right of third parties to impose sanctions against an offending country unless this third country is also adversely affected by the same policy that is the subject of the original dispute. In practice, however, multilateralism may have been instrumental in enforcing the trade agreement in a more subtle way. Davis (2006) compares two dispute cases under different institutional settings to show that the WTO dispute settlement process can help developing countries by means of legal “band-wagoning.” The first case is a complaint by Vietnam, accusing the United States of violating its obligation under a bilateral treaty the two countries had reached. Specifically, Vietnam attempted to prevent the United States from issuing food-labeling requirements that would have excluded its catfish from the U.S. market. Since Vietnam was not a member of the WTO, the dispute was pursued outside that organization’s dispute settlement process. The second case is a WTO dispute in which Peru, along with the United States and Canada, was a complainant against food-labeling regulations issued by the EU that would have limited access to its scallops and sardines.15 Davis points out that both Peru and Vietnam were disadvantaged because of their lack of retaliation power and legal resources. However Peru was far more successful in inducing compliance by the offending party. Peru’s success can be attributed to the presence of Canada and the United States, two countries with substantial retaliation power and legal resources, as joint complainants. Vietnam did not have such advantages in its bilateral relationship with the United States and was unsuccessful in its suit.

A related problem concerns the ability of small and developing countries to obtain adequate compensation in the event that they are injured by policies of large developed countries. This can be most easily seen in the case where the small country is effectively a price taker. If the large country imposes a tariff on imports from the small country, the terms of trade for the latter will be worsened proportionally. If the large country refuses to remove the barrier in response to a panel finding, the imposition of a retaliatory tariff by the small nation would have no impact on the terms of trade of

14 Third parties may be interested in limiting the market access of the offending party because of the standard beggar-thy-neighbor nature of trade policies.

15 European Communities – Trade Description of Scallops (DS 12), mutually agreed solution circulated 19 July 1996.
The Theory of Dispute Resolution with Application to Intellectual Property Rights

1. Commitment models of trade agreements

The analysis in this section has been based on the assumption that trade agreements are intended to resolve a Prisoner’s Dilemma. Another proposed approach is based on the idea that trade agreements commit governments to permanent trade liberalization in the presence of pressure from domestic interest groups to retain high protection. An example is the model by Maggi and Rodriguez-Clare (1998), in which politicians know that domestic firms in the import-competing sector will lobby for protection in the future. By committing to a trade agreement, they prevent these interest groups from forming and demanding trade barriers. They show that this type of commitment is most valuable for weak governments that are able to extract little surplus from the interest groups in the bargaining over tariffs.

In the commitment approach, the DSP would have a slightly different role to play. A decision by the dispute settlement body can be used by the government to deflect political pressure in nations where powerful interest groups demand protection. Under this interpretation, the government might be willing to temporarily erect trade barriers in response to political pressure, but would use the authority of the dispute panel to justify eliminating the restrictions once the panel had ruled.

4. Economic models of litigation and settlement bargaining

We now turn to an analysis of the behavior of countries involved in trade disputes. One of the stated goals of the DSU is to encourage the settlement of trade disputes through

---

16 Mexico has proposed that retaliation rights be made transferable to third parties as another means of dealing with this problem. Bagwell et al. (2004) analyze this issue and point out both benefits and costs of such an amendment.

consultations, with dispute panels being utilized as a last resort when consultations fail. Engaging in the dispute settlement process can be quite costly, since attorney fees for the panel stage can easily reach $200,000 to $300,000 for a simple case. For complicated cases, costs have exceeded $10 million.\(^ {18}\) The main incentive to settle the dispute is to avoid the costs of litigation.\(^ {19}\) The demands of the parties during the settlement negotiations will be influenced by expectations about the outcome if the case goes to a panel. Thus, the rules of the dispute process will influence the distribution of gains between the parties from the settlement, as well as the incentive to file complaints.

In this section we start with a formal representation of the settlement-bargaining process. Then we will discuss several sources of inefficiency – including asymmetric information among parties, political restrictions, and the like – and identify the characteristics of the parties or the dispute itself that may influence the outcome of the process. Finally we draw on settlement-bargaining models to explain empirical observations regarding the behavior of developing and developed countries in the DSP.

4.1. The settlement-negotiation problem defined

To illustrate the settlement negotiation problem, suppose we have an action pair \(a^0 = (a^0_c, a^0_d)\) reflecting the initial policy choices of the complainant \((a_c)\) and defendant \((a_d)\) at the time the complaint is made. The payoffs to the complainant and defendant from this policy pair are denoted \(W^0_c\) and \(W^0_d\), respectively. Let \(a^1\) denote the policy pair that is expected to result if the panel rules in favor of the complainant, where \(W^0_c < W^1_c\) and \(W^0_d > W^1_d\). We will begin by focusing on the case usually considered in the literature on settlement negotiation, where \(D = W^1_c - W^0_c = W^0_d - W^1_d\) is the amount of monetary damages that the defendant would be required to pay the complainant in the event of a settlement. Let \(p_d(p_c)\) be the probability that the defendant (complainant) assigns to a win by the complainant and \(c_d(c_c)\) the cost to the defendant (complainant) of litigating the case. The expected welfare gain of the complainant from proceeding with the complaint is \(G(p_c) = p_c(W^1_c - W^0_c) - c_c\) and the expected welfare loss of the defendant from proceeding is \(L_d(p_d) = p_d(W^0_d - W^1_d) + c_d\). In order for the complaint to be credible, it must satisfy \(G(p_c) \geq 0\). If this condition fails, the defendant country will not negotiate because it will recognize that the complainant would be worse off by proceeding to the litigation stage.

\(^{18}\) Shaffer (2003) reports that in the Photographic Film and Paper dispute initiated by the United States against Japan, these charges exceeded $10 million and were paid by Kodak and Fuji.

\(^{19}\) A defendant may also want to settle without going to trial in order to avoid the precedents set by the panel ruling, which may be used by other member countries to initiate new disputes in future. Moreover, in deciding on whether to pursue a contentious case, the big players in the WTO may also consider the effect of litigation on the evolution of the multilateral trading system. The latter case will be discussed below.
A successful outcome to the negotiations can arise if there exists a monetary transfer $S$ satisfying

$$L_d(p_d) = p_d D + c_d \geq S \geq G_c(p_c) = p_c D - c_c.$$  \hfill (2)

With common beliefs about the probability of a successful complaint (i.e. $p_c = p_d$), settlement should always occur prior to litigation in order to save the costs of litigation. If we assume that the outcome of the bargaining process is described by the Nash bargaining solution, the parties will split the gains from reaching an agreement. This yields an equilibrium transfer $S^N = pD + (c_d - c_c)/2$ and payoffs of $W_c^0 + S^N$ and $W_d^0 - S^N$. The party with higher litigation costs will be at a disadvantage in bargaining and thus will receive a lower payoff, even though litigation does not actually occur.

The rules of the dispute settlement process will affect the distribution of payoffs between the complainant and the defendant, as well as the incentive of countries that believe they have been injured to file complaints. There are two aspects of the dispute settlement process that favor the defendant in the bargaining process. The first is that the WTO discourages unilateral retaliation on the part of the complainant. Retaliation during the consultation phase would raise $W_c^0$ and reduce $W_d^0$, which would raise the net payoff to the complainant and reduce that to the defendant. It should be noted, however, that there is some question about the effectiveness of the WTO restraint against unilateral retaliation. Reinhardt (1999) found that when one country is the defendant in a WTO suit filed by a partner country in the previous year, the chances of the prior defendant filing a new dispute against that partner are increased by up to 51 times. These ‘tit for tat’ filings could represent the defendants’ responses to retaliatory actions taken by the complaining countries or their attempts to impose costs on those countries. The second aspect of the DSU that is in favor of the defendants concerns the level of punishment that can be authorized by the panel. Successful complainants are limited to the withdrawal of equivalent concessions and cannot make punishments retroactive, both features that tend to reduce $W_c^1$ and raise $W_d^1$.

This model also suggests that the strengthening of the rules under the WTO process should lead to a higher expected return to filing a complaint. The elimination of the unanimity requirement should raise the probability of success for the complainant, and any shortening of the panel process should raise the expected present value of a successful outcome. Since complainants will only file if the expected value is positive, these stronger rules would be expected to increase the frequency of complaints.

The model of settlement bargaining with symmetric information predicts that all disputes will be settled prior to going to a panel decision. However, we do observe that a significant fraction of the disputes – around 40 percent – end up with a panel decision. This proportion is actually much higher than that observed in domestic U.S. civil disputes, where only about five percent of cases go to trial. Therefore, it is useful to examine models that have the potential to explain why settlements do not occur prior to a panel decision. Models of bargaining between a plaintiff and defendant when there is asymmetric information have been developed in the law and economics literature to explain why some cases go to trial and others are settled out of court. Although these
models are not explicitly designed to analyze the dispute settlement process, they can provide insights about why settlement bargaining may fail.

Early models took a non-Bayesian approach to the settlement breakdown, in which the disputants may have different and possibly inconsistent priors about the outcome at trial. If either of the disputants is too optimistic about her chance to win at the trial, out-of-court negotiations could break down. Specifically, there will be values of \((p_c, p_d)\) such that \(p_c\) is sufficiently high and \(p_d\) sufficiently low that the inequality in (2) fails and no settlement can occur.\(^{20}\)

4.2. Incomplete-information models

While the model with differing initial beliefs can generate a failure to settle, it is not very helpful in generating predictions about what types of cases are more likely to lead to breakdowns in settlement. Subsequent modeling has focused on the premise that information may be revealed as part of the bargaining process and this revelation could mitigate the inefficiencies caused by information asymmetry. According to the models of settlement bargaining under asymmetric information, the disputants may act strategically in offering and accepting settlement options in order to screen their partner’s type or signal their own type. Under both signaling and screening models, a fraction of disputes fail to reach out-of-court settlement. We can illustrate this point using two of the canonical models of settlement bargaining under asymmetric information in which the complainant makes a settlement offer to the defendant. In the first model, due to Bebchuk (1984), the defendant has better information about the probability of a court decision and the offer is made by the uninformed party. In the second model, due to Reinganum and Wilde (1986), the complainant has better information about the size of the losses and the offer is made by the more informed party.\(^{21}\)

Bebchuk (1984) assumes that the defendant knows the true probability of success of the complaint, \(p\). The complainant is assumed to know that the likelihood of success is distributed according to a probability distribution, \(F\), with support on the interval \([p_L, p_U]\). The size of the damages is common knowledge. It is assumed that the complainant makes a settlement demand, \(S\). The defendant decides whether to accept the demand and settle the case without litigation, or to reject it and proceed to the court. From Equation (2), the defendant would accept \(S\) if and only if \(S < L_d(p)\), which requires that \(p > q(S) \equiv (S - c_d)/D\). The complainant country will then make an offer that maximizes its expected gain, which is the solution to

\[
\max_S \left(1 - F(q(S))\right) S + \int_{p_L}^{q(S)} (Dp - C_c) f(p) \, dp. \tag{3}
\]

\(^{20}\) For more discussion of this approach and for bibliographical information see Spier (2005).

\(^{21}\) For a thorough review of the economic theories of the settlement bargaining see Daughety and Reinganum (2005) and Spier (2005).
An increase in the settlement demand will extract more from defendants when the true probability is low, but will reduce the probability that the defendant actually accepts the demand. Letting \( \tilde{s} \) denote an interior solution to (3), the equilibrium will be one in which all defendants with \( p \geq q(\tilde{s}) \) will settle, while those with \( p < q(\tilde{s}) \) will reject the settlement demand and wait for the panel decision. The equilibrium settlement is denoted \( \tilde{r} = 1 - F(q(\tilde{s})) \). Generally, an injured party would file a complaint if and only if \( c_c \leq D \int_{p_L}^{p_U} x f(x) dx \). It can be shown that for \( c_c < p_L D \), an increase in the cost of litigation for either party or a reduction in the level of damages imposed will make settlement more likely. This result is intuitive as it implies that a dispute is escalated to the court only if the potential court award is so large that it becomes worthwhile to incur the costs of litigation.

Figure 2 illustrates how the outcome of settlement bargaining is affected by the level of litigation costs \( \{c_c, c_d\} \) when \( p \) is uniformly distributed on \( [p_L, p_U] \). For any values above the horizontal line, \( a \), no complaint will be filed because the complainant’s threat of litigation would not be credible. Thus, no defendant would be willing to settle a complaint. For values below \( a \), a complaint will be filed. The line \( b \) is the locus of values of \( c_c \) and \( c_d \) at which the probability of settlement is unity. Settlement will always occur for values below \( a \) and to the right of \( b \). The probability of settlement in the region to the left of \( b \) and below the dashed line is given by

---

22 Bebchuk imposes a restriction on the range of complainant costs to ensure that the complainant’s threat of litigation is always credible. Specifically, he assumes \( c_c < p_L D \), which is more restrictive than necessary, but it simplifies the analysis. Since we are also interested in the behavior of parties before filing a complaint, our description of the model considers the entire range of complainant’s costs.
Fig. 3. Equilibrium probability of settlement in the signaling model.

\[(c_c + c_d)/(D(p_U - p_L))\]. This means that the probability of litigation occurring decreases in both litigation cost parameters. In the region bounded by \(b\) and the dashed line, the settlement rate is given by \((p_U + p_L - 2c_c/D)/(p_U - p_L)\). A mean-preserving spread of the distribution of \(p\), which reflects an increase in the degree of informational asymmetry, would result in a rightward shift in line \(b\). Therefore, the probability of settlement is weakly decreasing in the degree of uncertainty. Thus, more litigation occurs when the degree of uncertainty is greater.

In contrast to Bebchuk (1984), Reinganum and Wilde (1986) develop a model in which it is the informed party who proposes a settlement. Suppose the complainant country knows more about the level of damages to the domestic industry due to the alleged violation of the agreement than does the defendant. We model this by assuming that the level of damages \(D\) is known with certainty by the complainant, while the defendant only knows that \(D\) is drawn from a probability distribution \(G\) with support on the interval \([D_L, D_U]\). By proposing a settlement, the complainant country effectively signals the extent of damages it suffers. In the equilibrium the defendant adopts a mixed strategy as to whether to reject or accept the complainant’s demand, conditional on the transmitted signal. The equilibrium probabilities are illustrated in Figure 3.

The screening and signaling models differ in that the former has a pure-strategy equilibrium, whereas the latter has a mixed-strategy equilibrium. However, the comparative statics of both models are similar in regard to the effect of litigation costs on the settlement rate. Therefore, the predictions of the two models concerning the relationships between country characteristics are similar.

4.3. Compensation methods: cash transfer versus policy adjustment

The discussions above have treated the negotiations between countries as if the damages and settlements represent cash transfers between countries. These cash transfers have two features that simplify the analysis of settlement bargaining. First, the bargaining between countries with cash transfers will be a zero-sum game because the monetary loss to one country will be of equal magnitude to the gain of the other.
Second, the cash transfers are divisible and can be treated as a continuous variable. In practice, the penalties imposed by panels and the settlements negotiated between countries almost always involve policy changes rather than cash transfers. In this section we examine how the analysis would be modified in cases where the policy changes do not involve a zero-sum game between countries and where there may be indivisibilities in policy choice.

When damages and settlement negotiations are conducted in terms of policy changes, there is a presumption that the settlement and dispute process is a positive-sum game. If a complaint results from the deviation by one country from a negotiated value, then the initial bargaining position should be an efficient policy choice for the countries as noted above. In contrast, a deviation by one country will involve an inefficient policy choice that benefits the deviator at the expense of the complaining country.

If the penalty imposed by the panel is a return to the initial negotiated policy pair, then the efficiency of the initial policy choice will mean that the gains to the complainant, \( W_{c}^{1} - W_{c}^{0} \), exceed the losses to the defendant, \( W_{d}^{0} - W_{d}^{1} \). In order for a settlement to be satisfactory to both parties, it must provide payoffs \( (W_{Sd}, W_{Sc}) \) to the respective parties such that \( W_{Sd} \geq W_{d}^{0} - L(p_{d}) \) and \( W_{Sc} \geq W_{c}^{0} + G(p_{c}) \).

When settlements are made by negotiating policy changes between countries, we can summarize the possible payoffs for the countries by the settlement frontier \( W_{d} = \varphi(W_{c}) \) that indicates the maximum settlement payoff for the deviating country as a function of the settlement welfare obtained by the complaining country. Clearly \( W_{d}^{i} \leq \varphi(W_{c}^{i}) \) for \( i = 0, 1 \) since the initial policy and settlement choices are known to be feasible. Since the expected payoff pair \( \{W_{d}^{0} - L(p), W_{c}^{0} + G(p)\} \) from litigation is strictly less than \( \{pW_{c}^{1} + (1 - p)W_{c}^{0}, pW_{d}^{1} + (1 - p)W_{d}^{0}\} \) when \( c_{c}, c_{d} > 0 \), it follows that a sufficient condition for a settlement to exist with \( p_{c} = p_{d} \) is that the settlement frontier be a continuous and concave function. This is clearly satisfied in the case of monetary transfers considered above, where the settlement frontier is \( W_{d} = W_{c}^{0} + W_{d}^{0} - W_{c} \). If policy choices are divisible and can be treated as continuous variables, then the settlement process may yield efficiency gains for the two countries by allowing choice of more efficient policies. Thus, the assumption of a continuous and concave settlement function seems plausible in this case.

This result shows that with continuous policy choices, the fact that compensation occurs in the form of policy choices rather than cash transfers does not alter the conclusion that settlement should always be reached when the countries have the same beliefs about the probability of a successful complaint. However, as Guzman and Simmons (2002) point out, if the subject matter of the dispute has an all-or-nothing character (e.g., disputes over health and safety regulations) and leaves little room for compromise, the parties’ ability to reach an agreement through policy adjustments may be limited.
is restricted. When monetary transfers are impractical, lumpiness of the subject matter of the dispute can lead to litigation even if the disputants share all the available knowledge relevant to the case. Consider for example the extreme case in which the only two possible outcomes for the country are to remain with the current policy or to choose the policy that will ultimately be imposed by the panel if it supports the complaint. The only circumstance under which the defendant would be willing to settle in this case is that if \( c_c \) and \( p \) are sufficiently high that 
\[
(1 - p)(W_0d - W_1d) < c_d.
\]

4.4. Political and institutional considerations

Under the above models, an inefficient outcome of the settlement-bargaining process is due to imperfect information. We also discussed the inefficiency that may arise from the restricted means of compensation in Section 4.3. But there are still other sources of inefficiency that seem to be important in the context of the international trading system. We discuss some of these problems that have been analyzed in the literature.

4.4.1. Signaling to domestic interest groups

Domestic politics may provide an alternative explanation for prevalence of litigation in the WTO. In dealing with trade disputes, governments are under pressure from their domestic interest groups such that any concession during the consultations may be associated with high political costs. As Davis (2006, p. 226) has put it, “... governments need a justification to give their domestic regulatory agency and lobby groups before they can change policies that were adopted to protect sensitive sectors.” By bringing a case before the dispute panel and obtaining official rulings of the WTO, a government can diffuse the political pressure from the interest groups by appealing to its international obligations. This explanation is more in line with the view of trade agreements as commitment mechanisms which are used to offset the bargaining power of local interest groups. Interestingly, as reported by Busch (2000) and Guzman and Simmons (2002), democratic governments, which are naturally under more influence from their domestic interest groups, are less likely to settle a dispute in the consultation stage.

4.4.2. DSU rules

Prevalence of late-stage settlements might also be a consequence of the DSU rules. Whereas any unilateral retaliation by a complaining country is discouraged by the DSU, an offending country is not held liable for the past injury imposed on the complainant and will not suffer a sanction if it discontinues the disputed action. That is, the DSU effectively allows an offending party to enjoy a period of time in violation of its obligations without any punishment. This may encourage some member countries to violate their obligations and then drag their feet along different stages of the

\[24\] In Butler and Hauser (2000), which is a complete information model of dispute settlement process, the lumpiness of the transfers between the disputants drives the results regarding the timing of settlement.
DSP and finally correct their actions just before retaliation is authorized. However, the offended country may opt for unauthorized retaliation in order to deter the offending country from feet-dragging in the negotiation process. The fact that so many tit-for-tat disputes are filed in the WTO supports the view that, during the negotiation process, complainants retaliate against responding parties by suspending some of the previously-granted market access.

Moreover, the DSU rule that allows third parties to join a WTO dispute may have bearings on the prospect of early settlements. As was explained above, most of the models of settlement bargaining predict a lower settlement rate for cases involving larger stake. Therefore, the engagement of third parties in the dispute, by raising the responding country’s loss from a concession, can increase that country’s incentive to go to trial. Involvement of third parties can further undermine the prospect of out-of-court settlement, as it is more costly and time-consuming to negotiate a settlement with several complainants. In Busch and Reinhardt (2006), empirical analysis of this issue implies that a single-complainant dispute with a 50-percent chance of settlement would only have a 16-percent chance of settlement if at least one third party joins the dispute.

In deciding whether to pursue a contentious case, the largest countries in the WTO may also consider the effect of litigation on the evolution of the multilateral trading system. Garrett and Smith (2002) point out that the United States and the EU have on occasion decided not to push contentious disputes through the WTO system for fear of de-legitimizing it. For example, they discuss the EU’s decision to withdraw its complaint against the U.S. Helms–Burton Act, and conclude (p. 3) that “the EU suspended the proceedings because it did not want to put WTO arbitrators in an invidious situation that inevitably would have damaged the DSU. A decision for the U.S. would have been difficult to justify on legal grounds, whereas Washington openly threatened to defy any ruling against it... [In such situations] the complainant may prefer to settle if it believes the defendant is unlikely to comply and hence that the only consequence of keeping the case on the WTO agenda will be to undermine the authority of the DSU.”

Developing countries may pursue other goals by engaging in WTO disputes as third parties. Participating in dispute settlement cases as a third party or jointly filing a case along with other more experienced WTO members can represent an easy window into learning how the system works, permitting governments to improve their legal capacity for future disputes. In fact, China has adopted this strategy and frequently joins cases as a third party.

25 Klimenko et al. (2006) and Reinhardt (1999) document the widespread use of tit-for-tat disputes. Further, according to Prusa (1999), two-thirds of all antidumping complaints are tit-for-tat responses to antidumping actions of other countries.

26 Sometimes, tit-for-tat disputes resemble frivolous lawsuits, which may be a strategic move by the defendant to impose more costs on the complainant from pursuing the original dispute.

27 This idea has been expressed by Christina Davis in an interview with The Viet Nam News, http://vietnamnews.vnagency.com.vn/showarticle.php?num=02COM220706.
4.5. Developing countries and the dispute settlement process

One of the complaints that has often been raised about the WTO dispute settlement process is that developing countries are at a disadvantage because they have high costs of litigation. High litigation costs could disadvantage smaller and poorer countries in two ways. First, these countries tend to have smaller trade stakes, both overall and in individual commodities, which makes it less profitable to absorb the litigation costs of pressing charges against countries that maintain illegal trade measures (Nordström, 2005). Second, because of their poor legal capacity, developing countries face higher marginal costs in WTO lawsuits. Greater legal capacity lets developed countries maintain teams of legal and economic experts responsible for gathering and analyzing evidence to detect potential infringements and to defend against potential complaints. Bown (2005) presents evidence in support of both of these views. He has found that while the size of exports at stake is an important determinant of a country’s decision to challenge a deviating country, its retaliatory and legal capacity and its international political economy relationships also matter.28

According to settlement-bargaining models with asymmetric information (presented in the previous sections), higher litigation costs should result in higher rate of early settlement. Therefore, if poorer countries have higher relative litigation costs, those models predict higher settlement rates for dispute cases involving them. Thus, a dispute between the United States and the EU is less likely to settle than a dispute between the United States and Argentina, and both are less likely to settle than a dispute between Argentina and Chile.

Empirical observations provided by Beshkar (2006) support this hypothesis. He shows that the likelihood of early settlement under DSP is inversely correlated with the GDP and GDP per capita of both disputants. However, in contrast with the predictions of settlement-bargaining models, the likelihood of settlement in the WTO is significantly more sensitive to the defendant’s wealth than to the complainant’s wealth. Beshkar (2006) extends these models with asymmetric information to situations where the relationship between the two countries is characterized by a Prisoner’s Dilemma. This makes the line $b$ in Figure 2 steeper. Under the modified models the perceived litigation costs of the defendant are higher than they actually are. Therefore, small changes in the defendant’s actual costs would have a pronounced effect on the likelihood of early settlement. For example, in a dispute between the United States and Argentina, it matters who is the defendant and who is the complainant.

The prospect of costly negotiations and lack of human capital may deter developing countries from filing disputes that are otherwise worth pursuing. In practice, larger trading nations have been far more active in using the WTO dispute settlement process.

---

28 In order to help developing countries improve their participation in the multilateral trading system, the Advisory Centre on WTO Law (ACWL) was established in 2001 and provides developing countries access to legal aid for an hourly charge that ranges from $25 for the least developed countries to $200 for the highest income developing countries (see http://www.ACWL.ch). Regarding implementation of the TRIPS agreement, similar assistance is provided by the World Intellectual Property Organization.
than smaller and poorer nations. Horn et al. (1999) argue that this dominance might be a reflection of the fact that larger countries encounter more questionable trade-related measures than smaller ones. Large countries tend to have wider export varieties, a situation that translates into a larger number of disputable trade measures. Assigning a fixed probability of dispute to each exporting product, a country’s frequency of filing complaints will be given by a binomial distribution function with the distribution parameter being the number of exporting products. Using data on export varieties of the WTO members and the dispute settlement data from the first four years of the DSU, they tentatively conclude that the seeming under-representation of developing countries, reflects their low diversity and value of trade.

5. Conclusions

In this chapter we have presented the state of the literature on dispute settlement and its relation to the protection of intellectual property. We conclude with some remarks on what seem to be promising areas for future research. First, the application of models of incomplete contracts to the analysis of dispute settlement is in its early stages and shows promise. These models suggest a role for the dispute settlement procedure for introducing flexibility into trade agreements under situations where there are policy adjustments with the potential to benefit both parties. It would be useful in this modeling to draw a distinction between the DSP and the safeguards system, which allows for adjustments to agreements in response to events that unduly harm domestic industries. When should flexibility be introduced through these rules as opposed to through the DSP? In particular, it is unclear why the DSU seems to discourage countries from “buying” adjustments in their tariffs in return for the withdrawal of equivalent concessions. Also, it would be useful to consider models of incomplete contracts that incorporate the requirement that agreements be self-enforcing.

A second point concerns the role of multilateralism in the DSP. Most of the models of the DSP involve just two countries. However, the fact that unilateral retaliatory actions are discouraged suggests that the WTO is concerned about their potential spillover effects to the trading system. Since unilateral retaliation could be accomplished more quickly than waiting for cases to work through the DSP, the deterrence benefits it would imply must be offset by negative effects on the stability of the trading system. It would be useful to have models that incorporated this spillover.

Uncited references

(Horn and Mavroidis, 2001)

References


