

THE ROLE OF DISPUTE SETTLEMENT PROCEDURES IN INTERNATIONAL TRADE AGREEMENTS*

Giovanni Maggi
Yale University and NBER

Robert W. Staiger
Stanford University and NBER

October 2008

Abstract

While disputes are typically treated as synonymous with concerns about enforcement in formal economic analyses of trade agreements, in reality most WTO disputes involve disagreements of interpretation concerning the agreement, or instances where the agreement is simply silent. And some have suggested that the WTO's Dispute Settlement Body (DSB) might serve a useful purpose by granting "exceptions" to rigid contractual obligations in some circumstances. In each of these three cases, the role played by the DSB amounts to "completing" various dimensions of an incomplete contract. Moreover, there is a debate among legal scholars on whether or not precedent-setting in DSB rulings may enhance the performance of the institution. All of this points to the importance of understanding the implications of the different possible degrees of *activism* in the role played by the DSB. In this paper we bring formal analysis to bear on this broad question. We characterize the choice of contractual form and DSB role that is optimal for governments under various contracting conditions. A novel feature of our approach is that it highlights the interaction between the design of the contract and the design of the dispute settlement procedure, and it views these as two components of a single over-arching institutional design problem.

*We thank Pierpaolo Battigalli, Mostafa Beshkar, Rick Bond, Alvaro Bustos, Christina Davis, Avinash Dixit, Eric French, Nuno Limao, Kristopher Ramsay, Alan Sykes, Thierry Verdier, Ben Zissimos, seminar participants at the Chicago Federal Reserve, Georgetown, George Washington University, Georgia Tech, Paris School of Economics, Stanford, Vanderbilt, Yale and participants in the 2007 UCSC International Conference, the 2008 AEA Meeting in New Orleans, and the 2008 Princeton-Yale Conference on Reputations and International Agreements for helpful comments. Maggi and Staiger gratefully acknowledge financial support from the NSF (respectively SES-054552435 and SES-0518802).

1. Introduction

Formal economic analysis of trade agreements typically treats disputes as synonymous with concerns about enforcement.¹ But at least in the World Trade Organization (WTO), observed disputes seem to be more about interpreting “vague” aspects of the agreement and filling “gaps” in the agreement than about enforcing unambiguous obligations specified by the agreement.² And it has been suggested by some legal scholars that the WTO’s Dispute Settlement Procedure (DSB) might serve a useful purpose by granting “exceptions” to rigid contractual obligations in some circumstances (see for example the discussion in WTO, 2007, pp. 279-282). In each of these three cases, the *activist* role played by the DSB amounts to “completing” various dimensions of an incomplete contract.

These distinctions are of more than passing interest: the dispute settlement system in the WTO – as well as in many of the major regional trade agreements – is very much in flux, and there is intense debate among scholars and practitioners about the role that a DSB should serve in a trade agreement: How activist a role should the DSB play? Should it be allowed to engage in “judicial lawmaking,” or should it be allowed only to enforce the obligations that are clearly stated in the agreement? And if the DSB is allowed to play an activist role, should it have authority to set precedent for future rulings? These questions have attracted considerable attention from scholars in international law, political science and economics, but the literature in this area is still largely informal in nature.

In this paper, we bring formal economic analysis to bear on the questions posed above. We evaluate the possible role that a DSB might play in helping to complete an incomplete contract, and we characterize the choice of contractual incompleteness and DSB design that would be optimal for governments under various contracting conditions. More specifically, we ask under what conditions it may be desirable to allow the DSB to *interpret* vaguely-stated obligations, to *fill gaps* in the agreement, or to *modify* rigid contractual obligations; and we evaluate the implications of allowing the DSB to set *precedent* for future rulings. A distinctive and novel feature of our approach is that it highlights the interaction between the design of the contract and the design of the dispute settlement procedure, and it views these as two components of a

¹See for example Bagwell and Staiger (2002, Chapter 6), and the literature cited therein.

²A recent WTO dispute (WTO, 2001) involving the use of special “safeguard” protection illustrates the point. In this case, the following issues were central to the resolution of the dispute (see Horn and Mavroidis, 2003): the meaning of “unforeseen developments,” the definition of “serious injury,” the acceptable method for establishing “causality,” the meaning of “increased imports,” and the meaning of “threat.”

single over-arching institutional design problem. The WTO DSB is a natural reference point for our analysis, and it serves as the lead example throughout the paper, but we emphasize that our analysis applies not just to the WTO but to trade agreements more generally.³

We consider a two-country setting where, in a given industry, the importing country selects a trade policy and the exporting country can then file a complaint, in which case the DSB steps in and issues a ruling. The environment is uncertain and the joint-payoff-maximizing policy varies with the underlying state of the world. We assume that the relevant contingencies are costly to describe, as in Battigalli and Maggi (2002), and this gives rise to two possible forms of contractual incompleteness: *rigidity* and *discretion*. We then use a simple form of fuzzy logic to introduce a contracting language which is vague in nature, and which gives rise to a third form of contractual incompleteness: *vagueness*. Our analysis then naturally pairs the three possible activist roles of the DSB described above with our three forms of contract incompleteness: the DSB can interpret vague obligations; it can fill gaps where the contract leaves governments with discretion; and it can grant exceptions to rigid obligations. Or, the DSB can serve none of these functions and simply enforce contractual obligations that are unambiguous. Furthermore, for each of the three activist roles, the DSB may or may not be given precedent-setting authority.

An important feature of our model is that it is costly for both the complainant (exporter government) and the defendant (importer government) to engage in a dispute, and this gives rise to non-trivial strategic interaction between them. Indeed, the relative performance of the various institutional arrangements depends crucially on the impacts of these arrangements on the equilibrium behavior of the governments, which in turn determines the circumstances under which a trade dispute will arise in equilibrium: as we later demonstrate, some of the more subtle insights generated by our framework hinge on how the institutional arrangement affects the equilibrium occurrence of disputes.

As a first step of the analysis, we abstract from the issue of precedent and consider a static (one-period) setting. A critical feature of our contracting environment is the accuracy of DSB rulings in representing government preferences. We characterize the optimal combination of contract and DSB role, as a function of the degree of DSB accuracy. We show that when

³For example, most preferential trade agreements have their own dispute settlement procedures, and our analysis can help shed light on the role and design of these dispute procedures as well. We believe that our model can also be applied to dispute settlement mechanisms in international agreements and organizations more broadly, not just trade agreements. On the other hand, as we argue in the concluding section, our model is less immediately applicable to domestic court systems.

DSB rulings are sufficiently accurate, it is optimal for governments to build discretion into the contract and provide the DSB with a mandate to fill the gaps when disputes arise. On the other hand, if DSB rulings are sufficiently inaccurate, we show that it is optimal for governments to negotiate a contract which is either vague or rigid and to then bar the DSB from engaging in any attempts to “complete” the contract when disputes arise. Finally, if the accuracy of DSB rulings falls in an intermediate range, we show that it is optimal for governments to negotiate a vague contract and provide the DSB with a mandate to interpret the vague aspects of the contract when disputes arise.

Interestingly, our analysis does not offer support for the “modification” role of the DSB: it is never optimal to allow the court to void obligations that are clearly stated in the agreement. Also notable is our finding that, if DSB rulings are sufficiently accurate and the DSB has an activist role, the first best outcome can be achieved even though the contract is highly incomplete, the use of the DSB is costly and the DSB rulings are imperfect. The reason is that the threat of invoking the DSB and the expectation of a sufficiently precise DSB ruling is sufficient to induce governments to act efficiently. Therefore, our model suggests that the imperfection in the DSB information does not necessarily impair the performance of the institution; the presence of an activist DSB can potentially generate dramatic efficiency gains in spite of its (inevitable) information limitations.

At the same time, our analysis offers an important warning. We find that, if litigation costs are not too high, the equilibrium policy tends to be efficient when the DSB is *not* invoked in equilibrium. Moreover, we find that equilibrium disputes are more frequent when the DSB is *less* accurate. Thus, in effect the motives that trigger a DSB filing are inefficient, and the efficiency-enhancing effect of the DSB tends to be associated with its *off-equilibrium* impacts. What underlies these findings is the fact that, anticipating the possibility of error by the DSB, each government is tempted to game the system within the leeway offered by the incompleteness of the contract: the importer is tempted to erect trade barriers when it should not, hoping to get away with it; and the exporter is tempted to force free trade by filing a dispute when it should not. One implication of these findings is that the intensity of DSB use is not a reliable indicator of the performance of the institution.

Our model can also shed light on the issue of “bias” in observed DSB rulings: both under the GATT and the WTO, complainants have mostly won their cases. What accounts for this observed pro-trade bias? Our model suggests that a pro-trade bias in DSB rulings arises

when disputes are mostly triggered as a result of the importer – rather than the exporter – acting opportunistically and exploiting the incompleteness of the contract; and in our model, this selection occurs when the costs of dispute fall sufficiently on the exporter relative to the importer (reflecting, perhaps, the allocation of the burden of proof). We can also address a related question: When DSB rulings exhibit a pro-trade bias, does the equilibrium *policy outcome* exhibit a pro-trade bias as well, in the sense that trade protection is under-utilized relative to the first best outcome? A natural conjecture would be that the bias in DSB rulings and in policy outcomes would go hand-in-hand. But, perhaps surprisingly, we find that there is an inverse relationship between the two biases: if there is a pro-trade bias in DSB rulings there tends to be an anti-trade bias in policy outcomes, and vice-versa.

We next extend the static model to a dynamic (two-period) setting to examine whether DSB rulings should set legal precedent for future rulings. There is a growing discussion in the legal literature as to whether and to what extent the actual WTO DSB operates on a precedent system.⁴ Jackson (2006, page 177) argues that “there is quite a powerful precedent effect in the jurisprudence of the WTO, but ... it is not so powerful as to require panels or the Appellate Body considering new cases to follow prior cases,” and concludes that “the ‘flavor’ of the precedent effect in the WTO is still somewhat fluid.” Given this state of affairs, where the role of precedent in the WTO is very much an open issue of institutional design, it seems particularly important to consider the implications of precedent from a formal perspective. We believe our framework is well-suited for this kind of analysis.⁵

In our dynamic setting, we show that the equilibrium frequency of trade disputes rises with the introduction of legal precedent set by DSB rulings, and that this effect by itself *diminishes* the overall performance of the contract/DSB pair. More broadly, our findings suggest that the introduction of precedent can enhance the performance of a contract/DSB-design pairing only if the expected value of the elimination of duplicative filing costs afforded by precedent outweighs

⁴In fact, the appropriate role of precedent in the WTO has become an issue in the recent “zeroing” disputes (see WTO, 2007a). In its recent report on the matter (see WTO, 2008), the Appellate Body sought to clarify the role of precedent in the WTO with the following statement: “Ensuring ‘security and predictability’ in the dispute settlement system, as contemplated in Article 3.2 of the DSU, implies that, absent cogent reasons, an adjudicatory body will resolve the same legal question in the same way in a subsequent case.”

⁵We are not aware of studies that examine the implications of precedent in the context of international agreements. But there are several papers in the literature that examine the effects of precedent in the context of a domestic court system (see for example the classic analysis by Posner, 1973, and the recent paper by Gennaioli and Shleifer, 2007). These papers focus mostly on the question whether precedent setting makes it more or less likely that the body of law will become more efficient over time.

the expected costs associated with the additional filing behavior that precedent induces. When we examine how the resolution of these opposing effects vary with key parameters of the model, we find that the introduction of precedent is more likely to enhance the performance of the institution (i) when the accuracy of DSB rulings is sufficiently low, and (ii) when governments care sufficiently little about the future.

Our paper is broadly related to Horn, Maggi and Staiger (2006), in the sense that both papers view trade agreements as incomplete contracts, where the incompleteness arises endogenously from contracting costs. But there the similarity between the present paper and Horn, Maggi and Staiger largely ends. In addition to the notion of vagueness which is novel to the present paper, the focus of the present paper is on the potential role of the courts in helping to complete the incomplete contract, a feature that is absent from Horn, Maggi and Staiger. As a result of the different focus, the structure of the model across the two papers is also very different, because modeling courts and disputes requires modeling the strategic interaction between governments in the dispute process, as well as the behavior of the court – elements that are absent from Horn, Maggi and Staiger but which form the heart of the present paper.

In addition to the work of Battigalli and Maggi (2002) and Horn, Maggi and Staiger (2006) discussed above, we are aware of several other papers that relate to various themes in our paper. In particular, a number of authors have begun to explore the interplay between the form of written contracts and the nature and degree of ex-post activities performed by the court. Shavell (2006), for example, emphasizes how the possible methods of contract interpretation can affect the writing of contracts, and characterizes the optimal method of interpretation. Anderlini, Felli and Postlewaite (2006, 2007) similarly analyze the role of active courts in contracting environments where the ex-ante contracts are incomplete, and characterize optimal court behavior (see also Fon and Parisi, 2007). At a broad level, these papers and our paper have a similar focus, but the particular questions addressed and the formal treatments are quite distinct.⁶

A paper that is related to a different theme of our paper – the notion of vagueness in contracting language – is that of Bustos (2007a), who extends Battigalli and Maggi (2002)

⁶In the literature on trade agreements, another paper that considers the potential role of the DSB for completing an incomplete contract is Battigalli and Maggi (2003). They consider a setting in which two countries choose product standards, and can only write an agreement that covers existing products, not products that may emerge in the future. To remedy this contractual incompleteness problem, the countries can institute a DSB that fills the gaps of the contract ex post. However, the formal structure adopted and the questions addressed in their paper are very different from our paper.

to include a choice between precise and vague clauses in the contract. There are a number of substantial differences between our paper and Bustos', but among these differences are two: first, unlike our paper and those reviewed just above, Bustos is not concerned with the interplay between the form of written contracts and the nature and degree of ex-post activities performed by the court, and indeed the court in Bustos' model plays only a contract-enforcement role; and second, the way in which Bustos defines vagueness – a clause that demands “best effort” in performing the action – is very different from the concept of vagueness that we introduce below and which we believe permeates the contracting language used in trade agreements.⁷

The rest of the paper proceeds as follows. Section 2 presents our basic static model, characterizes the optimal choice of contractual incompleteness and DSB design in this static setting, and considers the issue of bias in DSB rulings. Section 3 extends the model and results to a dynamic setting, and considers the impact of precedent setting in contract/DSB performance. Section 4 briefly considers several extensions of the basic model. Section 5 concludes.

2. The Basic Static Model

We begin this section by describing the economic environment. We then describe the contracting language, the possible contracts, and our assumptions on contracting costs. With this description in hand, we present our model of the DSB and discuss the possible DSB-role/contract pairings, and then turn to an analysis of the equilibrium behavior of governments and optimal contract/DSB design.

2.1. The economic environment

We consider a very simple partial-equilibrium setup to develop our points. We focus on a single industry in which an importing government chooses a binary import policy $T \in \{FT, P\}$ (Free Trade or Protection) to maximize the payoff $\omega(T; s)$, where $s \equiv (s_1, s_2, \dots, s_N)$ is a state vector. We assume for simplicity that each state variable represents a binary event, such as “there is/is

⁷This is not to say that the notion of vagueness explored by Bustos (2007a) is irrelevant for trade agreements. On the contrary, it is possible to find specific examples of the kind of vagueness that Bustos seems to have in mind in the GATT/WTO, where for example GATT Article XVI on subsidies uses the language “seek to avoid” in describing the obligation of contracting parties, or where for example in GATT Article XXXVII it is stated that “the developed contracting parties shall to the fullest extent possible...” . These phrases seem to represent the kind of clause that fits well with Bustos' conception of vagueness. Our point here is simply that there is a more pervasive kind of vagueness in the GATT/WTO contracting language, and our modeling represents an attempt to formalize it.

not an import surge” or “the domestic industry does/does not shut down.” We will often refer to the random vector s simply as the “state.” We let $p(s)$ denote the probability that state s occurs. The exporting government is assumed to remain passive in this industry (i.e., there is no exporter policy), and its payoff is given by $\omega^*(T; s)$.

Let $\gamma(s) \equiv \omega(P; s) - \omega(FT; s)$ denote the importing government’s gain from protection. This gain may be thought of as arising from some combination of terms-of-trade and political considerations. We assume that $\gamma(s) > 0$ for all states s . Similarly, let $\gamma^*(s) \equiv \omega^*(P; s) - \omega^*(FT; s)$. We assume that $\gamma^*(s) < 0$ for all s : the exporting government always dislikes import barriers. Finally, we assume that there cannot be transfers between governments at the ex-post stage (after the state s is realized).⁸

For future reference, we define the “first best” policy for a given state s as the policy that maximizes the governments’ joint payoff $\Omega(T; s) \equiv \omega(T; s) + \omega^*(T; s)$. More substantively, we will assume that the institution is designed to maximize the ex-ante expected value of Ω .⁹

Let $\Gamma(s) \equiv \gamma(s) + \gamma^*(s) = \Omega(P; s) - \Omega(FT; s)$ denote the joint (positive or negative) gain from protection for the two governments. We let σ^{FT} and σ^P denote the sets of states for which the first best policy is respectively FT and P , or equivalently, $\Gamma(s) < 0$ for $s \in \sigma^{FT}$ and $\Gamma(s) > 0$ for $s \in \sigma^P$. To make the problem interesting, we assume that the first best policy is Protection in some states and Free Trade in others; this amounts to assuming that σ^{FT} and σ^P are non-empty. We also assume that each of the N state variables is relevant for determining the first-best policy (in the particular sense that, for each of the N state variables, there is at

⁸In practice, direct transfers are rarely used as part of international trade agreements, but indirect transfers may be feasible. For instance, in the process of negotiating a trade agreement, it is not uncommon for countries to grant market access through tariff concessions in return for concessions from their trading partners that are unrelated to market access (e.g., tighter intellectual property rights protection, or beefed-up drug control efforts), and these non-market-access concessions can be interpreted as indirect transfers. We could allow for such *ex-ante* transfers in our model (i.e., transfers that occur at the stage of writing the contract), and need only rule out *ex-post* transfers (i.e., transfers that occur at the time of a dispute). The resolution of WTO disputes almost never involve direct transfers (the one exception to date is the *US-Copyright* case; see WTO, 2007, pp. 283-286), and indirect transfers of the sort described above are typically not feasible in the time-frame of dispute resolution. Nevertheless, a more realistic assumption might be that transfers can be enacted ex-post at some cost, and this cost varies case by case. This could help explain why some of the disputes are settled before going to court, as we discuss in section 4.3.

⁹There are three ways to justify this emphasis on the maximization of the governments’ joint payoff: one possibility is to allow for ex-ante transfers, i.e., transfers at the time the institution is created (this is not in contradiction with the assumption of no ex-post transfers – see footnote 8); a second possibility would be to keep the single-sector model and introduce a veil of ignorance, so that ex-ante there is uncertainty over which of the two governments will be the importer and which the exporter; and a third possibility would be to introduce a second mirror-image sector.

least one state in which the first-best policy hinges on the realized value of that state variable).

The governments observe the realized state s and the value of $\Gamma(s)$. However, Γ is not verifiable, so contracts cannot be made directly contingent on Γ .¹⁰ The underlying state variables (s_1, s_2, \dots, s_N) on the other hand are verifiable, and they can be specified in a contract at some cost. Therefore, the first-best outcome can be implemented by a contract that makes the policy contingent on the state variables.

2.2. Contracting language and contracting costs

We next describe the language that is available to write a contract, the possible contracts that can be written, and the costs of writing them. We begin by describing the contracts that can be written with the “crisp” (unambiguous) language that is typically assumed to be available in contract theory.

Consider first the benchmark case of a *complete* contract, which specifies T as a function of the full vector of state variables. As discussed in the previous section, the first best can be achieved with such a contract that takes the form:

$$T^{FB}(s_1, s_2, \dots, s_N) = \begin{cases} P & \text{if } (s_1, s_2, \dots, s_N) \in \sigma^P \\ FT & \text{if } (s_1, s_2, \dots, s_N) \in \sigma^{FT}. \end{cases}$$

This contract achieves the first best, but it is likely to be very costly to write, especially if the number of state variables N is large. Following Battigalli and Maggi (2002) and Horn, Maggi and Staiger (2006), we assume the cost of including a state variable in the contract is c_s . For now, we assume that specifying the policy (FT or P) is costless. The complete contract $T^{FB}(s_1, s_2, \dots, s_N)$ therefore costs $N \cdot c_s$.

Given our assumption on contracting costs, it is possible to reduce these costs by writing a contract that makes T contingent only on a subset of the state variables. The cost of such a contract is given by $n \cdot c_s$, where n is the number of state variables that are specified in the contract. Note that, given our payoff structure, if $n < N$ the first best cannot be achieved. In the limit, all state variables may be excluded from the contract, in which case we have two possibilities: a completely *rigid* contract, which specifies $T = FT$ for all states or simply “Free trade always;”¹¹ and the *discretionary* contract “Trade protection always allowed” (or

¹⁰That payoff levels are not verifiable is a standard assumption in this kind of model (this assumption is made also in Battigalli and Maggi, 2002, and Horn, Maggi and Staiger, 2006).

¹¹There is no need to consider the rigid contract $T = P$ for all s , because this is outcome-equivalent to the discretionary (or empty) contract that we describe next.

equivalently, the empty contract). These two contracts represent in a very stylized way the two types of contractual incompleteness that have been highlighted by the contract-theoretic literature: rigidity and discretion. We label these contracts respectively R and D . According to our contracting cost assumptions, there is no cost to writing either of these two contracts.

In addition to the crisp-language contracting possibilities described above, we now introduce a language which is *vague* in nature, in the sense that its meaning is only partially defined. To formalize vague language we use a simple type of *fuzzy logic*. To illustrate, consider a sentence \mathcal{F} , which could sound for example like “there is substantial injury to the domestic industry,” and suppose that the *truth function* of sentence \mathcal{F} is the following:

$$\text{Sentence } \mathcal{F} \text{ is } \begin{cases} \textit{True} & \textit{if } s_1 = \dots = s_N = 1 \\ \textit{False} & \textit{if } s_1 = \dots = s_N = 0 \\ \textit{Undefined} & \textit{otherwise.} \end{cases}$$

If T is made contingent on the truth value of such a sentence, something novel can be accomplished relative to the crisp-language contracts described above: T is made responsive to *some* states of the world, but in other states the contract remains vague. If for example $N = 3$, with $s_1 = 1$ ($s_1 = 0$) when there is (is not) an import surge, $s_2 = 1$ ($s_2 = 0$) when the domestic industry does (does not) shut down, and $s_3 = 1$ ($s_3 = 0$) when the majority of workers in the domestic industry are (are not) unemployed, then sentence \mathcal{F} could mean the following: there is certainly substantial injury to the domestic industry if there is an import surge, the domestic industry shuts down and the majority of workers in the industry are unemployed; there is certainly not substantial injury to the domestic industry if none of these events has occurred; but in the remaining states of the world it is not defined whether or not there is substantial injury – in other words, in these states it is a matter of interpretation.¹²

We assume that the truth function of sentence \mathcal{F} is common knowledge to the governments and the court, so the governments anticipate perfectly what truth function the court will assign to sentence \mathcal{F} . Henceforth we assume that $s^1 \equiv (1, \dots, 1) \in \sigma^P$ and $s^0 \equiv (0, \dots, 0) \in \sigma^{FT}$. In light of the sentence \mathcal{F} , we will sometimes refer to s^1 and s^0 as “extreme” states, and to all others as the “intermediate” states.

¹²A richer model would allow for a multiplicity of vague sentences, some more vague than others; for example there may be sentences that are true in more than one state and false in more than one state. Also, among the sentences with a similar degree of vagueness, some might “match” the first-best policy mapping (in our case, FT for $s \in \sigma^{FT}$ and P for $s \in \sigma^P$) better than others. In our basic model we abstract from these issues by considering only one simple vague sentence. In the context of a richer model, this could be interpreted as the best-performing among the set of possible vague sentences.

The vague language can be used to write what we label the *vague* (V) contract: “Trade protection is allowed only if there is substantial injury to the domestic industry.” We assume that there is no cost to writing the vague (V) contract.¹³

Summarizing, we have described three forms of contract incompleteness, each of which can reduce or eliminate the costs of contracting: *rigidity*, *discretion* and *vagueness*. At this point, in principle there are many contracts that we should consider. But in order to bring out the main insights in the simplest possible way, we assume that c_s is sufficiently high that specifying any state variable in the contract is suboptimal.¹⁴ Under this assumption we can restrict attention to the three benchmark types of incomplete contract: the rigid (R) contract, the discretionary (D) contract, and the vague (V) contract.

Finally, we have assumed that the R , D and V contracts cost nothing to write. But notice that, unlike the D contract, both the R and V contracts *would* cost something to write if specifying the policy (FT or P) were costly. Later we will introduce a small cost of specifying the policy as a way to break ties between the discretionary contract (D) and the rigid/vague contracts (R/V).¹⁵

2.3. The Dispute Settlement Body

Here we describe the potential roles that the DSB might play; we identify which possible DSB roles could arise under each type of contractual incompleteness; and we describe our assumptions on DSB behavior and information.

A first, basic, role of the DSB is to oversee enforcement of the obligations that are specified unambiguously in the contract.¹⁶ To the extent that the DSB is able to ensure enforcement,

¹³Of course, the performance of the V contract will depend in part on the likelihood of the “extreme” states $s_1 = \dots = s_N = 1$ and $s_1 = \dots = s_N = 0$; if N is large and the elementary events are *iid* these states would have low likelihood, but recall that we allow for elementary events to have different probabilities and to be correlated, so these states need not have low likelihood. Alternatively, it could be assumed that \mathcal{F} is true in more than one state and is false in more than one state. From this perspective, one could then interpret each extreme state as capturing a *cluster* of states rather than a single state.

¹⁴This seems like a reasonable first step, since in this paper we want to focus on the potential role of the DSB for completing an incomplete contract. Reducing c_s would intuitively diminish the potential role for the DSB, since this would make it optimal to write a more complete contract. But aside from this, it is not clear that reducing c_s would change the main qualitative insights of the paper. In any event, we leave the exploration of the more general case for future research.

¹⁵The assumption that specifying a policy is costly would be even more compelling in a richer setting with multiple policy dimensions. We discuss this extension briefly in section 4.2.

¹⁶There is a deep question about why a DSB is needed for enforcement (see, for example, Maggi, 1999 and Klimenko, Ramey and Watson, 2008). We address this question briefly in section 4.4.

this role is clearly a desirable feature of institutional design, and therefore we feel justified in taking this role for granted and keeping it in the background of the model. Our analysis will thus focus on whether and to what extent it is desirable for the DSB to go beyond a pure enforcement role. Since the enforcement role of the DSB is not the focus of our model, we make the simplest possible assumption regarding the enforcement of the contract: any obligation that is stated unambiguously in the contract is automatically enforced, just as in a standard contracting model where the contract is automatically enforced by a court.^{17,18}

Beyond the enforcement role, the DSB can play three potential roles:

- (1) Interpret obligations that are left vague in the contract. We denote this DSB role by *i*.
- (2) Fill gaps in the contract. Under this role the DSB introduces new obligations that are unambiguously *not* specified in the contract. We denote this DSB role by *g*.¹⁹
- (3) Introduce exceptions to rigid aspects of the contract. Under this role the DSB modifies obligations that are unambiguously specified in the contract. We denote this DSB role by *m*.

These roles span the potential (non-enforcement) activities of the DSB that find significant representation in the writings of WTO legal scholars. In effect, the interpretive (*i*), gap-filling (*g*) and modification (*m*) roles entail DSB efforts to “complete” the ex-ante contract. We will sometimes refer to a DSB whose mandate extends beyond the mere enforcement of the contract to include either the interpretation, gap-filling or modification role as an *activist* DSB, and to a DSB whose mandate is confined to contract enforcement as a *non-activist* DSB. Henceforth, when we make reference to the DSB being “invoked” or a “dispute” arising, this always refers to one of the three activist DSB roles.

¹⁷To be more precise, we assume that unambiguous obligations are automatically enforced unless the DSB has a mandate to grant exceptions (which is the case under the DSB role *m*, to be introduced below) and is invoked under this mandate.

¹⁸A more realistic assumption would be that the DSB intervenes to enforce contractual obligations only if it is invoked by a government (enforcement “on demand”), but this would only complicate the analysis without changing the main insights of the model.

¹⁹The distinction between gap-filling and interpretation/clarification is described by Goldstein and Steinberg (2007, p. 20) in this way: “Gap-filling refers to judicial law-making on a question for which there is no legal text directly on point, whereas ambiguity clarification refers to judicial law-making on a question for which there is legal text but that text needs clarification.” A similar distinction is adopted by Jackson (2006, p. 184): “In some cases, there are actual gaps in the treaty, in other words, the treaty is totally silent with regard to how it should apply in some circumstances. In other cases, there is treaty language that is applicable, but the treaty language is sufficiently ambiguous that it could reasonably be interpreted in several different ways.”

To be clear, the language of the WTO Agreement provides explicitly for an interpretive (*i*) role of the DSB, but it would seem to rule out the possibility that the WTO DSB could serve the modification (*m*) or gap-filling (*g*) roles.²⁰ Nevertheless, gap-filling (adding obligations) and modification (diminishing obligations) activities of the WTO DSB are frequently discussed in both positive and normative contexts by legal scholars (see, for example, Goldstein and Steinberg, 2007, and WTO, 2007). As such, we include them as important and relevant potential DSB activities to evaluate.²¹

We may now describe which of the activist roles identified above are relevant for each of the three forms of contractual incompleteness. In the case of discretion (*D*), only the role of gap-filling (*g*) is relevant. In the case of rigidity (*R*), only the role of modification of obligations (*m*) is relevant. And in the case of vagueness (*V*), only the interpretation role (*i*) is relevant.²²

Whatever the role played by the DSB, we assume that disputes are costly. In particular, whenever the exporter (complainant) invokes the DSB, the exporter incurs cost c^* and the importer (defendant) incurs cost c . We have in mind the costs of litigation, which may reflect various administrative costs, the costs of lawyers and the cost of acquiring in-house expertise, the burden of proof, etc., but in the model we treat these costs as parameters.

We next describe our model of DSB behavior. We assume that the DSB always operates in accordance with its mandate. Thus, there are three possible circumstances under which a dispute might arise in our model: (1) if the contract is discretionary (*D*) and the DSB has a gap-filling (*g*) role, there may be a dispute in any state for which there is no contractually stipulated obligation (i.e., there may be a dispute in any state s); (2) if the contract is vague (*V*) and the DSB has an interpretation (*i*) role, there may be a dispute in any state for which the contractually stipulated obligation is left vague (i.e., there may be a dispute in any intermediate state s); and (3) if the contract is rigid (*R*) and the DSB has a modification of obligations (*m*)

²⁰As Article 3, paragraph 2 of the WTO Understanding on Rules and Procedures Governing the Settlement of Disputes indicates: “The dispute settlement system of the WTO is a central element in providing security and predictability to the multilateral trading system. The Members recognize that it serves to preserve the rights and obligations of Members under the covered agreements, and to clarify the existing provisions of those agreements in accordance with customary rules of interpretation of public international law. Recommendations and rulings of the DSB cannot add to or diminish the rights and obligations provided in the covered agreements.”

²¹It is also possible to find examples of courts playing a modification-type role in domestic settings, where the court may void an obligation in the contract (see, for example, Anderlini, Felli and Postlewaite, 2006, 2007).

²²We feel justified in referring to the utter emptiness of the *D* contract as a “gap,” because in a richer model the contract could be empty along some dimensions while including non-empty clauses along other dimensions. Also, if we allowed c_s to be low, the optimal contract could have rigidity, discretion and vagueness (or a subset of these) at the same time, and so the DSB may play more than one of the activist roles; our basic model represents an attempt to focus more sharply on the trade-offs between the various roles.

role, there may be a dispute in any state for which the contractually stipulated obligation is rigid (i.e., there may be a dispute in any state s).

As we noted in section 2.1, the DSB is assumed to observe the realized state s but not the value of Γ ; thus, the DSB does not know what the “best” (joint-payoff-maximizing) policy is for the realized state s . We assume that, if invoked, the DSB observes a noisy signal of Γ , denoted Γ^{DSB} , which can be interpreted as the outcome of an independent investigation. The DSB then issues a *ruling* – which is simply a policy determination T^{DSB} for the realized state – with the objective of maximizing the expected joint payoff of the governments given the signal.²³ Note that the DSB ruling T^{DSB} fills a gap in the D contract in case (1) above, it interprets a vague obligation in the V contract in case (2) above, and it simply ignores the obligation stipulated by the R contract in case (3) above. We assume that DSB rulings are automatically enforced.

We now describe in more detail our assumptions on the information available to the DSB. Consider the moment when the DSB is invoked, before it observes the signal of Γ . From this time on, the state s is known with certainty, so it will appear in the notation as a deterministic variable. The DSB has some beliefs on Γ , represented by the distribution $f_s(\Gamma)$. We assume that, for any state s , the DSB priors on Γ are uninformative, so that $\Pr_s(\Gamma > 0) = \Pr_s(\Gamma < 0) = 1/2$ for all s .²⁴ Also, we assume that the signal is unbiased, and more specifically that the distribution of Γ^{DSB} conditional on Γ , denoted $h_s(\Gamma^{DSB}|\Gamma)$, is symmetric and centered around Γ .

A simple measure of the noise in the DSB signal is $q(s) \equiv \Pr_s(\Gamma^{DSB} \neq \Gamma)$; this is the probability that, given state s , the DSB signal has the opposite sign from the true Γ . Since the DSB priors are uninformative and the signal Γ^{DSB} is unbiased, it is easy to show that the DSB will issue the ruling $T^{DSB} = P$ if and only if it observes a signal $\Gamma^{DSB} > 0$. Note that,

²³By our assumption that the DSB maximizes the joint payoff of the governments, we seek to capture the DSB’s attempt to complete the contract for the realized state in the same way that the governments would have done ex ante. In essence, this is what the WTO DSB is supposed to attempt. For example, in WTO (2005), the Appellate Body states: “We recall that Article 31(1) of the Vienna Convention stipulates that: ‘A treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose.’... Importantly, the ordinary meaning of a treaty term must be seen *in the light of the intention of the parties*...” (emphasis added). Even if this is what the WTO DSB is supposed to attempt, there is of course an implementation issue of ensuring that the DSB actually attempts this: in the WTO DSB, this implementation issue may in part be addressed through the rules on selection of panel members (for example, to enhance impartiality of the panel, no panel member can reside in a country that is party to the dispute) and through the appeals process. Finally, there is a remaining question whether alternative DSB objectives might better serve the governments. We consider this question briefly in section 4.

²⁴We could allow the DSB priors on Γ to be informative. This would imply a slightly more complicated updating rule for the DSB, but the results of the model would not be affected.

if the DSB observes no signal at all, it is indifferent between choosing FT and P , and hence it randomizes in its ruling; this is the benchmark case of maximum DSB noise. Thus, $q(s)$ is the probability that the DSB issues the “wrong” ruling when playing an activist role.

Note that, since $h_s(\Gamma^{DSB}|\Gamma)$ is symmetric, $q(s)$ is bounded above by $1/2$. For the purposes of comparative-static analysis, we consider equi-proportional changes in the precision of the DSB signal, letting $q(s) \equiv q \cdot k(s)$, where $k(s) \in [0, 1/2]$ for all s and $q \in [0, 1]$ is a parameter that captures (inversely) the overall quality of the DSB information. We will vary q while keeping $k(s)$ fixed. The case $q = 0$ corresponds to the case in which the DSB has perfect information.

2.4. The possible institutions and the timing of events

Collecting all the possibilities that we have discussed above for the contract and the potential roles of the DSB, we have the following candidate *institutions*:²⁵

| <u>Contract</u> <u>DSB Role</u> | Discretion | Rigidity | Vagueness |
|------------------------------------|---|---|--|
| Non-activist DSB | D_n : no obligations | R_n : rigid obligations | V_n : vague obligations |
| Activist DSB | D_g : no obligations, DSB fills gaps | R_m : rigid obligations, DSB allows exceptions | V_i : vague obligations, DSB interprets |

Observe that the institution D_n delivers the noncooperative equilibrium outcome, and so amounts to no institution; the drawback of D_n of course is that it does nothing to prevent opportunistic behavior by governments. The institution R_n prevents opportunistic behavior but its rigid insistence on free trade implies that it sometimes gets the policy wrong. The institution V_n prescribes the correct policy in the extreme states s_0 and s_1 , but in the intermediate states the importer effectively has discretion and so the wrong policy is sometimes chosen under V_n as well. Finally, the institutions D_g , R_m and V_i entail high DSB-use costs and subject governments to potential DSB error as the DSB works to “complete” the contract.

We describe now the timing of events. We start with a game that is essentially static. In the dynamic extension that we consider in the next section, this game is repeated and a number of additional issues can be explored. But many of the main points are best illustrated in the simple static setting of this section. We consider the following timing:

stage 0. The institution is created (the contract is written, specifying the substantive obligations and the mandate of the DSB).

²⁵In a richer model these forms of contract-incompleteness/DSB-roles could well co-exist in the same institution. But for simplicity we use the term “institution” to refer to the contract/DSB-role pairing.

stage 1. The state (s_1, s_2, \dots, s_N) is realized.

stage 2. The importer government chooses the policy $T \in \{FT, P\}$.

stage 3. The exporter government decides whether to file a complaint with the DSB.

stage 4. If invoked, the DSB issues a ruling within its mandate $T^{DSB} \in \{FT, P\}$.

stage 5. If the DSB rules against the defendant ($T \neq T^{DSB}$), the import policy is modified to comply with the DSB ruling.

stage 6. Trades occur and payoffs are realized.

We can allow governments to renegotiate, both after the state (s_1, s_2, \dots, s_N) is realized in *stage 1*, and after the DSB issues a ruling in *stage 4* (these are the two relevant renegotiation nodes, because at each of these nodes governments receive some new information). But note that, since transfers between governments are not available, the possibility of renegotiation is irrelevant in this setting. This is because the interests of the two governments are directly in conflict and no Pareto improvement is possible ex-post; hence there is no room for renegotiation.

2.5. Analysis

We look for the institution $\hat{I} \in \{D_n, D_g, R_n, R_m, V_n, V_i\}$ that maximizes the governments' ex-ante joint payoff $E\Omega(\hat{T}(I); s)$, where $\hat{T}(I)$ is the (subgame-perfect) equilibrium policy that emerges under institution I .

Let us first consider the institutions with a non-activist DSB (D_n, R_n, V_n) . Recall that in each of these cases, the DSB has no mandate to “complete” the contract, and no disputes can arise. The comparison between these three institutions is therefore straightforward.

Under the institution D_n , clearly the equilibrium outcome is the policy P for all realizations of s . Under the institution R_n , the equilibrium outcome is FT for all s . And under the institution V_n – recalling that the V contract calls for FT in state s^0 and for P in state s^1 , while for all other states the obligations are left vague and hence nothing is enforced – the equilibrium outcome is for the importing government to select FT in state s^0 and P in all other states.

We can now compare the institutions D_n , R_n and V_n in terms of the ex-ante joint payoff. First note that V_n performs strictly better than D_n : the former imposes FT in state s^0 and

allows P in the other states, whereas the latter allows P in all states; thus D_n implements the wrong policy in one more state than V_n . So we can ignore the institution D_n .

We turn next to a comparison of R_n and V_n . It is convenient to compare these contracts in terms of the expected efficiency losses relative to the first-best outcome. These are given respectively by:

$$L(R_n) = \sum_{s \in \sigma^P} p(s)\Gamma(s), \text{ and } L(V_n) = \sum_{s \in \sigma^{FT} \setminus s^0} p(s)|\Gamma(s)|. \quad (2.1)$$

In general the ranking between these two institutions is ambiguous, and it depends on the intensity of the preference for FT in states $s \in \sigma^{FT}$ versus the intensity of the preference for P in states $s \in \sigma^P$. Our main results will hold regardless of this ranking.

We next turn to consider the institutions with an activist DSB: D_g , R_m , and V_i . It is helpful to distinguish between two cases: (a) those states s where, for the given institution, the activist role of the DSB is not applicable; and (b) those states s where the activist role of the DSB is applicable (under its mandate). In case (a), the equilibrium outcome is exactly as characterized above for the case of a non-activist DSB, while in case (b) we have to consider the possibility of disputes and derive the equilibrium outcome by backward induction.

Note that, within the set of activist-DSB institutions – D_g , R_m , and V_i – case (a) applies only for V_i , and only then under states s^0 and s^1 ; for all other states and institutions, case (b) applies. Thus, the equilibrium outcome for the institution V_i in states s^0 and s^1 is the same as for the institution V_n examined above. It remains to determine equilibrium government behavior for all those states and institutions in which case (b) applies.

Consider first the exporter government’s filing behavior. For a state s where the activist role of the DSB is applicable under the given institution, the exporter government files a complaint if and only if $T = P$ and the expected benefit to the exporter government of filing exceeds the exporter government’s cost of filing, that is

$$\Pr(\text{DSB ruling is } FT \mid s) \cdot |\gamma^*(s)| > c^*. \quad (\text{F})$$

Condition (F) is the “filing” condition for the exporter government to invoke the DSB in response to a policy choice by the importer government of $T = P$.

Next consider the importer government’s policy choice in a state s where the activist role of the DSB is applicable under the given institution. The importer government chooses $T = P$ if either (F) fails – because then the importer government can set $T = P$ without triggering

a dispute – or if (F) holds and the expected benefit to the importer government from trade protection exceeds the cost to the importer government of a DSB dispute:

$$\Pr(\text{DSB ruling is } P \mid s) \cdot \gamma(s) > c. \quad (\text{P})$$

To reduce the number of cases and focus on the more interesting ones, we assume that for each disputant the cost of a dispute is relatively small. More specifically, we assume that even in the benchmark case of maximal DSB noise, i.e. even where the DSB randomizes when playing an active role, the (F) and (P) conditions are both satisfied for all s , or:

$$\frac{1}{2}|\gamma^*(s)| > c^* \text{ and } \frac{1}{2}\gamma(s) > c \text{ for all } s. \quad (2.2)$$

Condition (2.2) implies that c^* and c are sufficiently small so that, for any level of DSB noise, (i) if the first best policy is P , the importer government chooses P whether or not this triggers a complaint by the exporter government, and (ii) if the first best policy is FT but the importer government still chooses P , the exporter government files a complaint.

We can now examine more closely each of the activist-DSB institutions D_g , R_m , and V_i .

Consider first the institution D_g . It is direct to derive the equilibrium actions of the governments for each state:

1. In states $s \in \sigma^{FT}$: if $qk(s) < \frac{c}{\gamma(s)}$ then $T = FT$ and the DSB is not invoked; if instead $qk(s) > \frac{c}{\gamma(s)}$ then $T = P$ and the DSB is invoked.²⁶
2. In states $s \in \sigma^P$: if $qk(s) < \frac{c^*}{|\gamma^*(s)|}$ then $T = P$ and the DSB is not invoked; if instead $qk(s) > \frac{c^*}{|\gamma^*(s)|}$ then $T = P$ and the DSB is invoked.

Notice an interesting feature of the D_g institution: the expected quality of DSB rulings has a perverse effect on the equilibrium utilization of the DSB, in the sense that *the DSB is invoked more frequently in equilibrium when the accuracy of its rulings is low* (i.e., when q is high). Specifically, in states $s \in \sigma^{FT}$, the importing government chooses the inefficient policy P and induces a DSB filing by the exporting government if and only if $qk(s)$ – the probability that the DSB will rule in error – is sufficiently high. Similarly, in states $s \in \sigma^P$, the exporting

²⁶To see this, note that condition (F) in this state is $(1 - qk(s)) \cdot |\gamma^*(s)| > c^*$, or $qk(s) < 1 - \frac{c^*}{|\gamma^*(s)|}$; given (2.2) this is always satisfied, hence the exporter files iff $T = P$. Condition (P) is given by $qk(s) \cdot \gamma(s) > c$, or $qk(s) > \frac{c}{\gamma(s)}$; thus the importer chooses $T = P$ iff $qk(s) > \frac{c}{\gamma(s)}$, as claimed. The claim about states $s \in \sigma^P$ that we make below can be shown analogously.

government chooses to file with the DSB and challenge the importing government's efficient policy choice P if and only if $qk(s)$ is sufficiently high. Note also that, *for a given state s , the first best outcome is achieved if and only if the DSB is not invoked.*²⁷ These two observations play a central role in the analysis to follow, and indeed they are at the heart of the main results of this section. Together they imply that *the equilibrium motives that trigger a DSB filing are inefficient from an ex-ante perspective, and it is the off-equilibrium impacts of the activist role of the DSB that are efficiency-enhancing.*

We are now in a position to calculate the expected efficiency loss associated with the institution D_g relative to the first-best outcome. This loss is given by:

$$L(D_g) = \sum_{s \in \hat{\sigma}^{FT} \cup \hat{\sigma}^P} p(s)[qk(s)|\Gamma(s)| + c^{all}], \quad (2.3)$$

where $\hat{\sigma}^{FT}$ is the set of states under the D_g institution for which FT is efficient, the importing government chooses to protect, and the exporting government files a complaint (i.e., s such that $s \in \sigma^{FT}$ and $qk(s) > \frac{c}{\gamma(s)}$), $\hat{\sigma}^P$ is the set of states under the D_g institution for which P is efficient, the importing government chooses to protect, and the exporting government files a complaint (i.e., s such that $s \in \sigma^P$ and $qk(s) > \frac{c^*}{|\gamma^*(s)|}$), and $c^{all} \equiv c + c^*$. Notice that in the set of states $\hat{\sigma}^{FT}$, it is the importer government who acts opportunistically and exploits the incompleteness of the contract, thereby triggering a dispute; and in the set of states $\hat{\sigma}^P$, it is the exporter government who acts opportunistically and exploits the incompleteness of the contract, thereby triggering a dispute.

As (2.3) makes clear, the institution D_g entails two inefficiencies relative to the first-best: one arising from the probability of DSB error in its activist role, and the other arising from the cost of the dispute. The expected loss $L(D_g)$ is given by this double inefficiency summed over two sets of states: the set of states $\hat{\sigma}^{FT}$ for which FT is the first best policy, the importer government exploits a gap in the contract and there will be a dispute in equilibrium; and the set of states $\hat{\sigma}^P$ for which P is the first best policy, the exporter government exploits a gap in the contract and there will be a dispute in equilibrium.

²⁷This particular feature depends on the assumption that c and c^* are relatively small. If c and c^* were large, inefficient policies would arise also for states where the DSB is not invoked in equilibrium. This is because the threat of invoking the DSB may not be credible if the cost of doing so is high. In this case (which might potentially be relevant for small developing countries, but probably not for large or developed countries), intuitively the role of the DSB would be diminished, but our main qualitative results would still hold, at least to the extent that c and c^* are not prohibitive, i.e. so high that they shut down the impact of the DSB altogether.

Note that $L(D_g)$ is increasing in q according to (2.3), for two reasons. First, a higher q implies more frequent mistakes by the DSB when the DSB is invoked, and this increases the expected cost of DSB error given a dispute in any state s , $qk(s)|\Gamma(s)|$. And second, as we have observed above, a higher q also increases the number of states in which the DSB is invoked in equilibrium, and therefore increases the size of the sets $\hat{\sigma}^{FT}$ and $\hat{\sigma}^P$.

Also note from (2.3) that no loss arises ($L(D_g) = 0$) if q is lower than a critical level $\tilde{q} > 0$. This is the level of q below which the DSB is not invoked in any state (that is, the level of q below which the sets $\hat{\sigma}^{FT}$ and $\hat{\sigma}^P$ are empty). Intuitively, if the DSB noise is small ($q < \tilde{q}$) then the governments, expecting the DSB (in its active role) to make the right decision with high probability, will act efficiently and avoid the DSB intervention to save on the dispute cost (the importer will always choose the first best policy and the exporter will never file complaints). This reflects the off-equilibrium impacts of the DSB described above.

Consider next the institution R_m . Under this institution, the contract specifies a rigid free trade rule, but the DSB can allow exceptions if invoked. Notice, though, that the DSB is unconstrained by the (rigid) contract when its mandate includes the modification of obligations (m), and hence if invoked, the DSB makes the same ruling as it would under the discretionary contract. But it should now be clear that the equilibrium outcome under R_m must then be the same as under D_g . This is because the DSB will rule in the same way in both cases, hence the governments will make the same decisions in both cases – the fact that under R_m the contract specifies free trade is irrelevant at the end of the day. Based on this observation, and applying our tie-breaking assumption – a small cost of specifying the policy (FT or P) that raises the contracting cost of R_m slightly above D_g – we conclude that R_m is dominated by D_g .

Finally consider the institution V_i . Clearly, in all intermediate states (i.e., states other than the extreme states s^0 and s^1), the equilibrium government actions are the same as those we have described above under the institution D_g . But in states s^0 and s^1 things may be different, because in these states, as we have already observed, the activist role of the DSB is not applicable under the institution V_i , and hence there is no dispute cost in these states and the equilibrium outcome is first-best, that is, respectively $T = FT$ and $T = P$.

We are now in a position to calculate the expected efficiency loss associated with the institution V_i relative to the first-best outcome. This loss is given by:

$$L(V_i) = \sum_{s \in \hat{\sigma}^{FT} \cup \hat{\sigma}^P} p(s)[qk(s)|\Gamma(s)| + c^{all}], \quad (2.4)$$

where $\bar{\sigma}^{FT} = \hat{\sigma}^{FT} \setminus s^0$ and $\bar{\sigma}^P = \hat{\sigma}^P \setminus s^1$. The interpretation of $L(V_i)$ is the same as that given for $L(D_g)$, except that the sets $\bar{\sigma}^{FT}$ and $\bar{\sigma}^P$ now replace $\hat{\sigma}^{FT}$ and $\hat{\sigma}^P$: the set $\bar{\sigma}^{FT}$ is the set of states for which FT is the first best policy, the importer government exploits the vagueness of the contract and there will be a dispute in equilibrium; and the set $\bar{\sigma}^P$ is the set of states for which P is the first best policy, the exporter government exploits the vagueness of the contract and there will be a dispute in equilibrium.

To compare the performance of the D_g and V_i institutions, we define $\check{q} \equiv \min\left(\frac{c}{k(s^0)\gamma(s^0)}, \frac{c^*}{k(s^1)|\gamma^*(s^1)|}\right)$. It is direct to verify that, if $q < \check{q}$, the DSB will not be invoked in the extreme states s^0 and s^1 under the D_g institution, and the equilibrium outcome is respectively $T = FT$ and $T = P$. But this corresponds exactly to the outcome under the V_i institution in the extreme state s^0 and s^1 ; and as we have already observed, the equilibrium outcome under D_g and V_i is always the same for intermediate state realizations. Hence, if $q < \check{q}$, the equilibrium outcome under D_g is the same as under V_i in all (intermediate and extreme) states. By contrast, if $q > \check{q}$, the DSB *will* be invoked in at least one of the extreme states s^0 and/or s^1 under the D_g institution, implying that the double inefficiency described above will be incurred in at least one of these states under the D_g institution, and implying as a consequence that the expected efficiency loss under D_g is higher than under V_i . Thus for $q > \check{q}$ we find using (2.3) and (2.4) that V_i is strictly preferred to D_g , while for $q < \check{q}$ we have a tie. That is:

$$L(D_g) > L(V_i) \text{ if } q > \check{q}, \text{ and } L(D_g) = L(V_i) \text{ if } q < \check{q}.$$

Applying again our tie-breaking assumption – that specifying the policy (FT or P) entails a small cost, raising the contracting cost of V_i slightly above D_g – we conclude that V_i is better than D_g if $q > \check{q}$, and vice versa if $q < \check{q}$. Note that, if $k(s^0)$ and $k(s^1)$ are small, so that the DSB would have a very good idea about the “right” policy in the extreme states, then \check{q} is higher than one, and therefore $q < \check{q}$ for all $q \in [0, 1]$; in this case D_g is preferred to V_i for all $q \in [0, 1]$.

We are now ready to draw our conclusions about the optimal institution. We have narrowed the choice to four institutions: V_n , R_n , V_i and D_g . The best of these institutions is the one that minimizes the expected efficiency loss L . First note from (2.1) that $L(V_n)$ and $L(R_n)$ are both independent of q , and hence the ranking between V_n and R_n does not depend on q . Next recall from (2.3) and (2.4) that V_i is preferred to D_g iff $q > \check{q}$. Finally note from (2.4) and (2.1) that V_n is preferred to V_i iff q is above a threshold level,²⁸ and the same can be said for the ranking

²⁸To see this, note that (i) $L(V_n) > L(V_i)$ for $q = 0$; and (ii) $L(V_i)$ is increasing in q . It follows that V_n is

between R_n and V_i . Putting together these results, we have the following:

Proposition 1. *There exist critical levels q_1 and q_2 (with $0 < q_1 \leq q_2 \leq 1$) such that: for $q < q_1$ the optimal institution is D_g ; for $q_1 < q < q_2$ the optimal institution is V_i ; and for $q > q_2$ the optimal institution is either V_n or R_n .*

Proposition 1 carries with it a number of implications. We highlight five of these below.

First, Proposition 1 implies that leaving gaps in a contract can be optimal, but only when the DSB is given an activist mandate; by contrast, introducing rigidity into a contract can be optimal, but only when the DSB is given a non-activist mandate; and finally vague clauses can be optimal under either an activist or non-activist DSB mandate (although note that each of the intervals (q_1, q_2) and $(q_2, 1)$ may be empty). More broadly, these observations reflect a simple underlying point: it is optimal to leave governments with discretion in more states of the world and endow the DSB with a mandate to rein in the exercise of that discretion the better the quality of the DSB information.

Second, Proposition 1 implies that, if q is sufficiently small, the first best outcome can be achieved under D_g or possibly V_i even though (i) the contract is highly incomplete, (ii) the use of the DSB is costly, and (iii) the DSB rulings are imperfect. The reason is that the threat of invoking the DSB and the expectation of a sufficiently precise DSB ruling is sufficient to induce governments to act efficiently. But notice that the first-best cannot be achieved if the DSB is not given an activist mandate.

Third, Proposition 1 implies that there is no modification role for the DSB in the optimal institution, contrary to the suggestions of some legal scholars (see WTO, 2007). Intuitively, rather than placing a rigid obligation into the contract and then endowing the DSB with a mandate to modify the obligation ex post, it is always better to simply leave a gap in the contract to begin with and then endow the DSB with a mandate to fill the gap ex post.

Fourth, Proposition 1 suggests an interesting question: Can an activist DSB role be desirable even if the DSB has no information in any state and simply randomizes (i.e. even if $q = 1$ and $k(s) = 1/2 \equiv k$ for all s , so that $qk = 1/2$)? Under these conditions, if we let the costs of a dispute (c^{all}) go to zero, it is direct to show using (2.1) and (2.4) that $L(V_i) \rightarrow \frac{1}{2}L(R_n) + \frac{1}{2}L(V_n) + \frac{1}{2}p(s_0)|\Gamma(s_0)|$, which implies $L(V_i) > \min[L(R_n), L(V_n)]$. Our model therefore

preferred to V_i iff q is above a threshold level q^v . Note that q^v may be higher than one, in which case V_i is preferred to V_n for all $q \in [0, 1]$.

predicts that, at least if the costs of a dispute are sufficiently small, the DSB needs to have at least *some* information if an activist DSB role is to improve efficiency.

And finally, a corollary of Proposition 1 is that the probability of a DSB dispute – and hence the expected cost of disputes – is non-monotonic in q , and in particular it is increasing for low levels of q and decreasing for high levels of q (provided the interval $(q_2, 1)$ is nonempty). The reason is that when q is sufficiently low the DSB is not invoked at all in equilibrium, and when q is sufficiently close to 1 it is optimal to have a non-activist DSB. Notice as well that, because the efficiency achieved by the optimal institution is decreasing in q , there is a nonmonotonic relationship between the equilibrium frequency of DSB use and the performance of the optimal institution. Therefore, our model implies that one cannot rely on information about the equilibrium frequency of DSB use to gain information about the performance of the optimal institution in terms of how close it gets to the first best.

2.6. A pro-trade bias in the DSB?

It has been pointed out in the literature on WTO disputes that there is an apparent “pro-trade bias” in DSB rulings. For example, according to the WTO (WTO, 2007, p. 273), “...both under the GATT (82 per cent) and the WTO (88 per cent) complainants have mostly won their cases (counting the ones that went through to an adopted report and ‘decisive’ ruling respectively).” In an effort to offer an interpretation of this bias, Goldstein and Steinberg (2007) argue that, if it is costly for the *exporter* to file a dispute, there may be a pro-trade (selection) bias in DSB rulings, because the exporter will file only when there is a high probability of success.

In this section we show that our model points to a somewhat different interpretation of this pro-trade bias in DSB rulings. Rather than the absolute *level* of dispute costs faced by the exporter, our model suggests that this bias arises when the costs of dispute fall sufficiently on the exporter *relative* to the importer (reflecting, perhaps, the allocation of the burden of proof). Intuitively, as we demonstrate below, a pro-trade bias in DSB rulings arises in our model when disputes are mostly triggered as a result of the importer – rather than the exporter – acting opportunistically and exploiting the incompleteness of the contract, and this occurs when the exporter (importer) faces relatively high (low) dispute costs.

We also raise a related question: Under what conditions do the equilibrium *policy outcomes* skew in favor of free trade relative to the first best outcome, and in this sense exhibit a pro-trade bias? One might conjecture that the two dimensions of bias – the bias in rulings and the bias

in policy outcomes – would go in the same direction, but interestingly, we will show that the two biases are inversely related: if there is a pro-trade bias in DSB rulings, then there tends to be an anti-trade bias in policy outcomes, and vice versa.

We first consider the bias in DSB rulings. In order to focus on selection as the only source of bias, we assume that if disputes were initiated randomly, the outcome of disputes would be unbiased.²⁹ Formally, we assume

$$\sum_{s \in \sigma^{FT}} p(s) = \sum_{s \in \sigma^P} p(s) \quad (2.5)$$

and

$$\sum_{s \in \sigma^{FT}} q(s)p(s) = \sum_{s \in \sigma^P} q(s)p(s), \quad (2.6)$$

where to avoid unnecessary notation we now use $q(s)$ rather than $qk(s)$ as our (inverse) measure of the quality of DSB information. Condition (2.5) says that the first-best policy is FT or P with equal probability. Condition (2.6) says that the DSB error is not correlated with the first-best policy.³⁰

For simplicity we consider only the D_g institution, but similar results apply to the V_i institution. Recalling that $\hat{\sigma}^{FT}$ denotes the subset of σ^{FT} in which a dispute is filed under the D_g institution and similarly $\hat{\sigma}^P$ denotes the subset of σ^P in which a dispute is filed, we can write:

$$\begin{aligned} & \Pr(T^{DSB} = FT \mid \text{file}) \\ &= \frac{\Pr(T^{DSB} = FT \mid s \in \hat{\sigma}^{FT}) \cdot \Pr(s \in \hat{\sigma}^{FT}) + \Pr(T^{DSB} = FT \mid s \in \hat{\sigma}^P) \cdot \Pr(s \in \hat{\sigma}^P)}{\Pr(s \in \hat{\sigma}^{FT}) + \Pr(s \in \hat{\sigma}^P)} \\ &= \frac{\sum_{s \in \hat{\sigma}^{FT}} (1 - q(s))p(s) + \sum_{s \in \hat{\sigma}^P} q(s)p(s)}{\sum_{s \in \hat{\sigma}^P} p(s) + \sum_{s \in \hat{\sigma}^{FT}} p(s)}. \end{aligned} \quad (2.7)$$

It follows from (2.7) using condition (2.5) and condition (2.6) that $\Pr(T^{DSB} = FT \mid \text{file}) > 1/2$ if and only if

$$\sum_{s \in \hat{\sigma}^{FT}} (1 - 2q(s))p(s) > \sum_{s \in \hat{\sigma}^P} (1 - 2q(s))p(s). \quad (2.8)$$

²⁹By focusing on selection bias we abstract from other possible causes of the observed bias in DSB rulings. For example, the underlying cause might be simply that the probability of being in the set σ^{FT} is higher than $1/2$. This is a possible explanation, but not a very interesting one. The analysis of this section should be interpreted as shedding light on the determinants of selection bias, not as suggesting that selection bias is the only possible explanation for the observed bias in DSB rulings.

³⁰To see this, notice that $\sum_{s \in \sigma^{FT}} q(s)p(s) / \sum_{s \in \sigma^{FT}} p(s)$ gives the probability of DSB error conditional on $s \in \sigma^{FT}$, and $\sum_{s \in \sigma^P} q(s)p(s) / \sum_{s \in \sigma^P} p(s)$ gives the probability of DSB error conditional on $s \in \sigma^P$. Given condition (2.5), these two expressions are equal if and only if $\sum_{s \in \sigma^{FT}} q(s)p(s) = \sum_{s \in \sigma^P} q(s)p(s)$, which is condition (2.6).

Our goal now is to identify the region in (c, c^*) -space where (2.8) holds. First note that, if $c < \min_{s \in \sigma^{FT}} q(s)\gamma(s)$ and $c^* < \min_{s \in \sigma^P} q(s)|\gamma^*(s)|$, filing occurs in equilibrium in every state s . In this case, clearly there is no bias. Next focus on the case in which $c > \min_{s \in \sigma^{FT}} q(s)\gamma(s)$ or $c^* > \min_{s \in \sigma^P} q(s)|\gamma^*(s)|$, so that filing occurs in some but not all states. Recall that, by our small-cost assumption (2.2), the number of states in $\hat{\sigma}^{FT}$ depends only on c (and is weakly decreasing in c) and the number of states in $\hat{\sigma}^P$ depends only on c^* (and is weakly decreasing in c^*). This implies that the left hand side of (2.8) is a weakly decreasing function of c and the right hand side of (2.8) is a weakly decreasing function of c^* . It follows immediately that (2.8) is satisfied if and only if $c^* > y(c)$, where $y(\cdot)$ is a weakly increasing function. Thus, at a broad level we find that DSB rulings tend to have a pro-trade bias if the dispute costs incurred by the exporter are high relative to the dispute costs incurred by the importer.

We next characterize the bias in the policy outcomes under the D_g institution (again, similar results apply to the V_i institution). Since we are assuming here that the first-best policy is FT or P with equal probability, it is natural to say that the policy outcome exhibits a pro-trade bias if $\Pr(\hat{T}(D_g) = FT) > 1/2$, where recall that $\hat{T}(D_g)$ is the equilibrium policy that emerges under institution D_g .

A key observation is that, in the sets of states where no dispute is filed (i.e. $\sigma^P \setminus \hat{\sigma}^P$ and $\sigma^{FT} \setminus \hat{\sigma}^{FT}$), the equilibrium policy is the first-best policy. Using this observation, and with some straightforward algebra, we can write:

$$\begin{aligned} & \Pr(\hat{T}(D_g) = FT) \\ &= \Pr(s \in \sigma^{FT} \setminus \hat{\sigma}^{FT}) + \Pr(T^{DSB} = FT | s \in \hat{\sigma}^{FT}) \cdot \Pr(s \in \hat{\sigma}^{FT}) + \Pr(T^{DSB} = FT | s \in \hat{\sigma}^P) \cdot \Pr(s \in \hat{\sigma}^P) \\ &= \frac{1}{2} - \sum_{s \in \hat{\sigma}^{FT}} q(s)p(s) + \sum_{s \in \hat{\sigma}^P} q(s)p(s). \end{aligned}$$

This implies that $\Pr(\hat{T}(D_g) = FT) > 1/2$ if and only if

$$\sum_{s \in \hat{\sigma}^{FT}} q(s)p(s) < \sum_{s \in \hat{\sigma}^P} q(s)p(s). \quad (2.9)$$

Intuitively, there is no bias in policy outcomes under the D_g institution if no disputes ever arise, because as we have shown in that case the equilibrium policy is always the first-best policy; and so, as (2.9) indicates, the extent of anti- or pro- trade bias in policy outcomes depends on the relative size of the sets $\hat{\sigma}^{FT}$ and $\hat{\sigma}^P$.

Using arguments analogous to those above for the analysis of bias in DSB rulings, it is straightforward to show that (2.9) is satisfied if and only if $c^* < z(c)$, where $z(\cdot)$ is a weakly

increasing function. Thus, in sharp contrast with our finding concerning the bias in DSB rulings, we find that the policy outcomes tend to exhibit a pro-trade bias if the dispute costs incurred by the exporter are *low* relative to the dispute costs incurred by the importer. This result becomes even more crisp in the special case where $q(s)$ and $p(s)$ are the same for all states. In that case, it is direct to see that the two curves $y(c)$ and $z(c)$ coincide, and hence (2.9) is satisfied if and only if (2.8) is violated, and vice-versa.

The following proposition summarizes these results:

Proposition 2. *Consider the D_g and V_i institutions: (i) There is a pro-trade bias in DSB rulings if and only if $c^* > y(c)$, where $y(\cdot)$ is a (weakly) increasing function. (ii) There is a pro-trade bias in policy outcomes if and only if $c^* < z(c)$, where $z(\cdot)$ is a (weakly) increasing function. (iii) In the symmetric case where $q(s) = q$ for all s and $p(s) = p$ for all s , $y(c) = z(c)$, and hence there is a pro-trade bias in DSB rulings if and only if there is an anti-trade bias in policy outcomes.*

Figure 1 conveys the main message of Proposition 2 for the symmetric case where $q(s) = q$ for all s and $p(s) = p$ for all s :³¹ the conditions leading to a *pro-trade bias* in the DSB rulings are essentially the same as those leading to an *anti-trade bias* in the *policy outcomes*. Evidently, our model suggests that when c^* is high relative to c , the joint behavior that is induced by both the importer government and the exporter government under the D_g and V_i institutions means that the most common reason for a dispute to arise is because the importer government is trying to exploit the incompleteness of the contract and “get away with protection” (rather than the alternative that the exporter government is trying to exploit the incompleteness of the contract and “get away with forcing free trade”); and this ensures that the policy outcomes tend to be biased toward Protection relative to the first-best policies even though the DSB, when invoked, will usually find in favor of the complainant and rule for Free Trade.

3. Precedent Setting

We next develop a dynamic extension of the static model described in the previous section. With this dynamic extension, we consider a further important issue of DSB design: whether or

³¹In Figure 1, we define $c_1 \equiv \min_{s \in \sigma^{FT}} q(s)\gamma(s)$, $c_1^* \equiv \min_{s \in \sigma^P} q(s)|\gamma^*(s)|$, $c_2 \equiv \max_{s \in \sigma^{FT}} q(s)\gamma(s)$ and $c_2^* \equiv \max_{s \in \sigma^P} q(s)|\gamma^*(s)|$. Note that by our assumption (2.2) we can focus on the region where $c < c_2$ and $c^* < c_2^*$. Also note that the “No-Bias” locus is technically not a function but a correspondence, due to the discrete nature of the state space. If the state space were continuous this locus would be a curve.

not DSB rulings should set legal *precedent* for future rulings.

For simplicity, we consider a two-period version of the static model developed above, with a prior period (*Period 0*) in which the institution is created. *Period 1* and *Period 2* then proceed as in the static model of the previous section. We assume that the state (s_1, s_2, \dots, s_N) is *iid* across the two periods,³² and we let $\delta \geq 0$ denote the factor by which governments discount *Period-2* welfare: because “the future” is collapsed into *Period 2*, we allow that δ may be arbitrarily large. Finally, we denote by T_t and T_t^{DSB} ($t \in \{1, 2\}$) the *Period-t* importer-government policy choice and DSB ruling, respectively. Given that *Period 2* is the repetition of *Period 1*, and given the *iid* assumption, there is nothing truly dynamic in the contracting environment. The dynamic aspect of the analysis will arise from the presence of the DSB institution, *if* the DSB has precedent-setting authority.

As in the static model, we can allow the governments to renegotiate the contract at two stages within each period: after the state (s_1, s_2, \dots, s_N) is realized in *stage 1*, and after the DSB issues a ruling in *stage 4*. From the perspective of *Period 2*, the two possibilities for renegotiation look identical to those in the static model; that is, since transfers between governments are not available, the possibility of renegotiation – both at the end of *stage 1* and at the end of *stage 4* – is irrelevant in *Period 2*. This is because in *Period 2*, as in the static model, the interests of the two governments are directly in conflict and no Pareto improvement is possible ex-post; hence there is no room for renegotiation. But does this conclusion hold as well from the perspective of *Period 1*? After all, if DSB rulings set a precedent for the future, it seems plausible that governments might want to renegotiate after a “bad” ruling in *stage 4*. Nevertheless, upon further reflection it can be seen that the “winning” country under the DSB ruling in *Period 1* would never agree to a renegotiation without the possibility of transfers.³³ Hence, we may conclude that there is no room for renegotiation in our dynamic model.

We next describe the formal meaning of “precedent” within our model. When DSB rulings set *precedent*, we assume that a *Period-1* ruling $T_1^{DSB}(s')$ for the realized state s' implies henceforth that the contract specifies $T(s') = T_1^{DSB}(s')$. Therefore, if the DSB operates under precedent and the DSB is invoked in *Period 1* for the realized state s' , then in *Period 2* the

³²At the end of the section we consider an extension of the model that allows for persistence in the state of the world, in order to examine how the optimal institution depends on the degree of persistence.

³³More broadly, it is tempting to think of the periodic GATT/WTO multilateral negotiating “Rounds” as environments in which government-to-government transfers can more effectively be orchestrated (see note 8) and where the precedent set by “bad” DSB rulings might be addressed through renegotiation. This interpretation, however, is well beyond the scope of our simple model.

contract is “complete” for the state s' .³⁴

An interesting question is whether and to what extent it is costly for the DSB to describe the realized state s' , which it must do if its ruling is to set a precedent for future realizations of this state. In reality, this cost is probably non-negligible but significantly lower than the cost of describing a state *ex-ante*, before its realization. What we have in mind is similar to Battigalli and Maggi (2002)’s notion of “unforeseen events” (see their Section II.B6): *ex ante* there is a set of elementary events s_i that the contracting parties are aware of, but also a large number of “latent” elementary events that they do not have in mind because these latent events are normally turned “off.” If a latent elementary event is turned “on” *ex post*, at that point it becomes relatively easy to describe. For simplicity we assume that the cost of describing the realized state s' is zero. This makes the difference between *ex-ante* and *ex-post* description costs rather extreme in our model (recall the assumption that c_s , the cost of describing a state variable *ex-ante*, is sufficiently high that specifying a contingent contract is suboptimal), but it would be easy to allow for a positive cost of *ex-post* description; intuitively, this would have the straightforward effect of making precedent-setting less attractive, other things equal.

Finally, we assume for simplicity that the DSB has no recall in *Period 2* of its observed *Period-1* signal, so that if the same state occurs in both periods and if the DSB is invoked in both periods (which is possible in the absence of legal precedent), then the DSB uses only the *Period-2* signal when making its *Period-2* ruling.³⁵ It is now immediately clear that, if invoked in *Period 2*, the DSB behaves exactly as it did in the static model characterized in

³⁴This is not the only conceivable way to introduce precedent in the DSB. In principle, one could consider an institution where the DSB “completes” the contract not just for the realized state s' but also for other states. But given our assumptions, this is unlikely to be an attractive option, because the DSB observes a signal of $\Gamma(s)$ only for the realized state s' , while it has no information for any other state. Moreover, describing states that are different from the realized state s' is likely to be costly, just as it is costly to describe states *ex-ante*. But one can think of a slightly different model where it might be attractive to extend the application of precedent beyond the realized state s' . Suppose for example that $\Gamma(s)$ is increasing in each state variable s_i (e.g. one can think of each s_i as a political shock that increases the payoff from protection), and this is known to the DSB. Then the following precedent system might be desirable: if the DSB ruling is *FT* for state s' , the same ruling will apply for any $s \preceq s'$, and if the ruling is *P* for state s' , the same ruling will apply for any $s \succeq s'$ (e.g., if the DSB chooses *FT* when there is an import surge and the import-competing industry has shrunk, then *FT* will apply also if there is an import surge but the import-competing industry has not shrunk.) We suspect that the qualitative results of our model would extend to this setting.

³⁵We make this assumption to keep the model as simple as possible and to focus on the main interesting results. But it is clear that allowing the DSB to recall its past signals would introduce another cost of precedent, because under precedent the DSB cannot then reduce the noise of its signal with repeated draws over time for the same state. Notice, though, that our assumptions stack the deck in favor of precedent, by assuming no DSB recall and ignoring the DSB cost of describing the realized state s' , and yet we will show below that precedent still performs less well than one might expect.

the previous section, because at this point precedent for the future is irrelevant. Moreover, the DSB behaves exactly as it did in the static model even if it is invoked in *Period 1*: the prospect that its ruling will set a precedent for the future does not alter the conclusion that, to maximize the expected joint payoff of the governments (given its unbiased signal and its uninformative priors), the DSB will issue the ruling $T_1^{DSB}(s) = P$ if and only if under the state s it observes a signal $\Gamma_1^{DSB}(s) > 0$.

With these preliminaries established, we can now delve into the analysis. In the absence of legal precedent, it is straightforward to see that our dynamic model behaves exactly as a (twice repeated) version of the static model, and our analysis from the previous section carries over. Moreover, it is immediate that introducing precedent has no impact on the performance of any of the institutions with a non-activist DSB, that is, on the institutions (D_n, R_n, V_n) . Also, R_m is again outcome-equivalent to D_g , and hence dominated. Therefore, to evaluate the role of legal precedent in the optimal institution, we need only derive the impact of precedent on the performance of the institutions V_i and D_g .

Again to reduce the number of cases and focus on the more interesting ones, we impose for the dynamic model a slight strengthening of the static-model condition (2.2). We assume that c^* and c are sufficiently small so that, for any level of DSB noise, (i) if the first best policy is P , the importer government chooses P in *Period 1* whether or not this triggers a complaint by the exporter government, and (ii) if the first best policy is FT but the importer government still chooses P in *Period 1*, the exporter government files a complaint. This is ensured by the following condition (as we show in the proof of Remark 1):

$$\frac{1}{2}(1 - \delta p(s))|\gamma^*(s)| > c^* \quad \text{and} \quad \frac{1}{2}(1 - \delta p(s))\gamma(s) > c \quad \text{for all } s. \quad (3.1)$$

We now make an important observation:

Remark 1. *If the DSB plays an activist role (gap-filling or interpretation), the equilibrium frequency of (Period-1) disputes rises with the introduction of legal precedent.*

The proof of this and the remaining results of the paper is in the Appendix. The intuition for Remark 1 is that for $s \in \sigma^{FT}$, precedent magnifies the gain that the importer anticipates if it provokes a DSB filing by setting $T = P$ and the DSB rules in error (which the DSB does with probability $qk(s)$).³⁶ And similarly, for $s \in \sigma^P$, precedent magnifies the gain that the exporter

³⁶Of course, if the DSB ruling is FT , precedent also extends this (correct) ruling into the future, and so it

can anticipate if it files against $T = P$ and the DSB rules in error (which the DSB does with probability $qk(s)$).

As we noted in the previous section in the context of the static model, the equilibrium motives that trigger a DSB filing are always inefficient from an ex-ante perspective, and it is the off-equilibrium impacts of the activist role of the DSB that are efficiency-enhancing. By itself, therefore, the observation that precedent increases the equilibrium filing frequency suggests that introducing precedent into an activist-DSB institution could *diminish* the efficiency of the institution. Also note that, even though introducing precedent makes the DSB in some sense “more activist,” this does not increase the beneficial off-equilibrium effect of the DSB. However, weighing against these negative effects of precedent is the benefit of eliminating the duplicative costs of filing each period.

The next step is to assess the net impact of the opposing effects highlighted above. A complete analysis can be found in the Appendix; here we only outline the broad logic of the arguments.

We focus first on the D_g institution. Let D_g^P denote the D_g institution with precedent-setting, and $L(D_g^P)$ the associated loss relative to the first best. As we show in the Appendix, the differential loss implied by the introduction of precedent can be written as:

$$L(D_g^P) - L(D_g) = \sum_{s \in \hat{\Delta}} p(s)[qk(s)|\Gamma(s)| + c^{all}] + \delta \sum_{s \in \hat{\Delta}} p^2(s)qk(s)|\Gamma(s)| - \delta \sum_{s \in \hat{\sigma}} p^2(s)c^{all} \quad (3.2)$$

where $\hat{\sigma}$ is the set of states where *Period-1* filing occurs under the D_g institution and $\hat{\Delta}$ is the set of states for which *Period-1* filing occurs under D_g^P but not under D_g .

The first term on the right-hand-side of (3.2) represents the *Period-1* efficiency loss generated by precedent, coming from the additional equilibrium filing behavior and associated DSB error and filing cost. The second term on the right-hand-side of (3.2) represents the discounted *Period-2* efficiency loss generated by precedent: this can be understood by observing that $p^2(s)$ is the probability that state s will occur in both *Period 1* and *Period 2*, and therefore that the precedent set from a *Period-1* ruling in state s will be relevant in *Period 2*. Finally, the third term on the right-hand-side of (3.2) is the discounted *Period-2* savings in duplicative filing cost

might be thought that precedent magnifies the loss from a correct ruling as well. But this is the policy that the importer would have to choose for $s \in \sigma^{FT}$ in order *not* to provoke a DSB filing, and so relative to that alternative a correct DSB ruling imposes no loss from filing (in either period). A similar statement applies to the exporter for $s \in \sigma^P$.

over the states in which *Period-2* filing would have occurred in the absence of precedent (i.e., $s \in \hat{\sigma}$).

Evidently, $L(D_g^{\mathcal{P}}) < L(D_g)$ – and hence the introduction of precedent enhances the performance of the D_g institution – only if the savings in duplicative filing costs outweighs the inefficiencies associated with the additional filing behavior.

We turn next to the vague contract and the interpretive role of the DSB. Let $V_i^{\mathcal{P}}$ denote the V_i institution with precedent-setting, and $L(V_i^{\mathcal{P}})$ the associated loss relative to the first best.

If the *Period-1* realized state is neither s^0 nor s^1 , then the statements made above apply, with $V_i^{\mathcal{P}}$ and V_i taking the place respectively of $D_g^{\mathcal{P}}$ and D_g . If instead the *Period-1* realized state is s^0 or s^1 , then there is no activist role for the DSB in *Period 1* under the V_i or the $V_i^{\mathcal{P}}$ institution, the importer government makes the first-best *Period-1* policy choice, and *Period 2* behaves exactly like the static model under V_i .

Hence, we can write the differential loss implied by the introduction of precedent as:

$$L(V_i^{\mathcal{P}}) - L(V_i) = \sum_{s \in \bar{\Delta}} p(s)[qk(s)|\Gamma(s)| + c^{all}] + \delta \sum_{s \in \bar{\Delta}} p^2(s)qk(s)|\Gamma(s)| - \delta \sum_{s \in \bar{\sigma}} p^2(s)c^{all} \quad (3.3)$$

where $\bar{\sigma}$ is the set of states where *Period-1* filing occurs under institution V_i and $\bar{\Delta}$ is the set of states for which *Period-1* filing occurs under $V_i^{\mathcal{P}}$ but not under V_i .

Clearly, introducing precedent into the V_i institution involves identical trade-offs to those described above for the D_g institution, except that the two extreme states s^0 and s^1 are excluded from (included in) this tradeoff for the V_i institution (D_g institution).

Armed with expressions for the efficiency losses associated with the $D_g^{\mathcal{P}}$ and $V_i^{\mathcal{P}}$ institutions, the next step of the analysis is to examine the optimal choice of institution. In this setting, we need to consider six candidates for an optimum, namely the same four institutions considered in the static model (R_n, V_n, V_i and D_g) plus the two institutions with precedent ($D_g^{\mathcal{P}}$ and $V_i^{\mathcal{P}}$). In general, any of these six institutions can be optimal depending on parameter values, but we can say something about how changes in q and δ affect the optimal choice of institution.

Consider first the impact of changes in the DSB noise q . In analogy with the previous section, an increase in q clearly favors non-activist-DSB institutions over activist-DSB institutions, so that as q increases the optimal institution can only switch from one of the latter to one of the former. Also, it continues to be true that an increase in q favors institutions where the DSB interprets a vague contract (V_i and $V_i^{\mathcal{P}}$) over institutions where the DSB fills gaps (D_g and $D_g^{\mathcal{P}}$), so that as q increases the optimal institution can only switch from one of the latter to

one of the former. Thus, the result of Proposition 1 generalizes to this dynamic setting: as q increases, the optimal institution switches from a discretionary contract with a gap-filling DSB to a vague contract with a contract-interpreting DSB to a vague or rigid contract with a non-activist DSB (possibly “skipping over” one or more of these).

But how does an increase in q affect the desirability of introducing precedent-setting in the DSB? To answer this question we must understand how q affects the differences $L(D_g^{\mathcal{P}}) - L(D_g)$ and $L(V_i^{\mathcal{P}}) - L(V_i)$. Let us assume that $k(s)$ is sufficiently close (or equal) to $1/2$ for all s , so that as q goes from zero to one, the DSB signal goes from perfect to uninformative. This is a natural way to parameterize the degree to which the DSB is informed, with $q = 0$ then indicating that the DSB is perfectly informed and $q = 1$ indicating that the DSB is completely uninformed. We have the following result:

Proposition 3. *Conditional on the DSB playing an activist role (gap-filling or interpretation), it is optimal to give the DSB precedent-setting authority when the DSB is sufficiently uninformed, while it is preferable not to do so when the DSB is sufficiently informed.*

Proposition 3 reports a somewhat surprising result, but it can be understood to reflect a simple logic as follows. Consider the D_g and $D_g^{\mathcal{P}}$ institutions (analogous arguments hold for the V_i and $V_i^{\mathcal{P}}$ institutions). In the case where the DSB is sufficiently uninformed, the DSB will be invoked in every state realization under the D_g institution, and so when precedent-setting power is introduced into the DSB the equilibrium *Period-1* frequency of filings cannot rise (because it is already 1): the only effect of moving from D_g to $D_g^{\mathcal{P}}$ in this case is then to eliminate duplicative filing costs in *Period 2*, which is efficiency enhancing. For this reason, endowing the DSB with precedent-setting powers in the D_g institution is *desirable* when the DSB operates with sufficiently *little* information. Now consider the case where the DSB is sufficiently informed so that the DSB will not be invoked in any state realization under the D_g institution: introducing precedent and moving from D_g to $D_g^{\mathcal{P}}$ in this case will either (i) preserve the absence of DSB filings in all state realizations, in which case D_g and $D_g^{\mathcal{P}}$ are outcome-equivalent, or (ii) lead to a *Period-1* filing in at least some state realization under $D_g^{\mathcal{P}}$, in which case the equilibrium *Period-1* frequency of filings is increased while there is no elimination of duplicative filing costs in *Period 2*, and therefore efficiency must be reduced. For this reason, endowing the DSB with precedent-setting powers in the D_g institution is *undesirable* when the DSB operates with sufficiently *good* information.

Next we consider the impact of changes in the discount factor δ . We find that, for q in an intermediate range, precedent-setting is desirable if governments are sufficiently impatient, while it is harmful if governments are sufficiently patient:

Proposition 4. *There exists an intermediate range of q such that, conditional on the DSB playing an activist role (gap-filling or interpretation), it is optimal to give the DSB precedent-setting authority if δ is sufficiently low, while it is preferable not to do so if δ is sufficiently high.*

The intuition for the finding reported in Proposition 4 is simple. If q is either sufficiently low or sufficiently high, we know from Proposition 3 that introducing precedent is respectively undesirable or desirable regardless of the level of δ . What Proposition 4 indicates is that, when q lies in an intermediate range, the level of δ becomes decisive. In particular, when δ is small, introducing precedent adds little additional impetus to initiate a dispute with the DSB, and the implied savings of duplicative filing costs then dominate the efficiency effects, making precedent-setting attractive. On the other hand, a large δ magnifies the additional impetus to initiate a dispute with the DSB that comes from the introduction of precedent, and this accentuates the efficiency-reducing impacts of a precedent-setting DSB to a sufficient degree that the introduction of precedent-setting powers becomes unattractive.

We conclude this section by considering a simple extension of the model with persistent shocks. Suppose that the probability of state s occurring in *Period 2* conditional on state s having occurred in *Period 1* is $\tilde{p}(s; \rho)$, where ρ is a persistence parameter, so that $\partial\tilde{p}/\partial\rho > 0$. Also let $\hat{p}(s; \rho)$ be the probability that the realized state in both *Period 1* and *Period 2* is s , with $\partial\hat{p}/\partial\rho > 0$.

Introducing persistence in the model changes the analysis only at two junctures. First, in the conditions that determine whether a dispute occurs in a given state s , $p(s)$ is replaced by $\tilde{p}(s; \rho)$: this implies that an increase in ρ has the same qualitative impact as an increase in δ on the sets $\hat{\Delta}$ and $\bar{\Delta}$ (see Appendix Proof of Remark 1). Second, in the expressions for $L(D_g^P) - L(D_g)$ and $L(V_i^P) - L(V_i)$, the term $p^2(s)$ is replaced by $\hat{p}(s; \rho)$, so ρ and δ enter these expressions through $\delta\hat{p}(s; \rho)$. It is direct to conclude that an increase in ρ has the same qualitative impact on these loss differences as an increase in δ .

This last observation has an interesting, and perhaps surprising, implication: a high degree of persistence in political shocks – or more generally shocks to government preferences – tends

to *disfavor* the use of precedent, even though it magnifies the savings of duplicative filing costs afforded by precedent-setting. The intuition for this result is analogous to the intuition for the impact of δ : when persistence is high, precedent-setting magnifies the governments' incentives to initiate disputes (which reduce efficiency) more than it magnifies the savings in duplicative filing costs.

4. Discussion

In this section we explore briefly a number of dimensions from which we have thus far abstracted in our theoretical analysis.

4.1. Further institutional design possibilities

We have considered a fairly simple set of institutional design possibilities. In particular, we have assumed that the DSB, if invoked, maximizes the governments' joint payoff based on its own (imperfect) information. Moreover we have assumed that the DSB's information is given by a noisy signal that is used to update the DSB's (uninformative) priors. Finally, the DSB in our model is non-strategic, in the sense that it does not try to infer the governments' true preferences from the governments' actions. One could consider at least three potential ways to enhance the efficiency of the institution. First, the DSB could try to extract information from the governments through a revelation mechanism, for example some form of "hearing." This is what happens in actual WTO trade disputes, where governments have a chance to present their arguments before the DSB makes a decision. Second, the DSB could be given a mandate to maximize a different objective than the governments' joint payoff. In principle, this could have beneficial effects through its impact on the governments' filing behavior.³⁷ And third, the DSB could be strategic and try to make inferences from the governments' choices.

We discuss first the possibility of a DSB hearing. There exists a small literature on cheap-talk games where "experts" send messages to a less informed decision maker. Two prominent examples of this literature are Krishna and Morgan (2001) and Battaglini (2002). In Battaglini's model, two experts simultaneously send cheap-talk messages, and then the policy-maker makes a decision; Krishna and Morgan consider a similar game, except that the experts send messages in a sequential fashion. The results of these papers suggest that, in a setting where the experts

³⁷The idea that biasing a court's objective may have a beneficial effect through the equilibrium frequency of litigation is present also in Bustos (2007b), though in a very different setting.

have conflicting interests, as in our model, the decision maker (DSB) may be able to elicit some information from the experts (governments) but is unlikely to learn the state with certainty.³⁸ This suggests a reinterpretation of the signal Γ^{DSB} observed by the DSB in our model: Γ^{DSB} can be thought of as incorporating the information that the DSP is able to extract from the hearings, and q can be interpreted as capturing the DSB's residual uncertainty about Γ after the hearings. Conditional on this reinterpretation, our qualitative insights are likely to still be valid even in this setting.³⁹

Let us focus next on the possibility of biasing the DSB's objective away from the governments' joint payoff. Consider for example biasing the DSB's objective slightly in favor of free trade: this can be captured by assuming that the DSB rules in favor of the exporter if and only if $\Gamma^{DSB}(s) > -\Delta$, where Δ is some small positive number. Recall that in our basic model the DSB rules in favor of the exporter if and only if $\Gamma^{DSB}(s) > 0$; thus $\Delta > 0$ introduces a pro-trade bias in the DSB decision.⁴⁰ It is immediate to verify that introducing a pro-trade bias of this kind has four effects: (i) it decreases the accuracy of the DSB rulings for states such that $s \in \sigma^P$ and the DSB is invoked in equilibrium; (ii) it increases the accuracy of the DSB rulings for states such that $s \in \sigma^{FT}$ and the DSB is invoked in equilibrium; (iii) it increases the frequency of equilibrium disputes for states $s \in \sigma^P$, which weighs negatively on the perfor-

³⁸We can be a little more precise here. The above mentioned papers analyze situations where the private information and the policy vector have the same dimensionality. Negative results (non-existence of fully revealing equilibria in the hearing game) are obtained in the one-dimensional case with strong conflict of interests (Krishna-Morgan, 2001, Propositions 1 and 4; Battaglini, 2002, Proposition 2). In our model, both the policy choice (T) and the private-information parameter (Γ) are one-dimensional, and the conflict of interests is strong, because in each state of the world the two governments prefer the opposite policy. Although our model does not perfectly fit the frameworks of Battaglini or Krishna and Morgan, their results suggest that full revelation would not occur in our case.

³⁹In this discussion we have implicitly assumed – as in Battaglini's and Krishna and Morgan's papers – that the DSB cannot impose monetary penalties (or more generally, transfers) conditional on the governments' reports. If such penalties were available, there would exist mechanisms that implement the first best. For example, suppose that governments simultaneously announce the value of Γ and, if the reports are different, both governments are hit with steep penalties. Clearly this kind of mechanism can implement the first best, because it is an equilibrium for the governments to reveal the true value of Γ . But we believe it is reasonable to abstract from this kind of mechanism; in reality the DSB does not have the power to impose penalties on governments for the policies they choose, let alone for the announcements they make. See subsection 4.4 below for a discussion of self-enforcement issues.

⁴⁰As we noted in section 2.6, some commentators have expressed the view that the WTO DSB is biased in favor of free trade. In that section, we explored the possibility that such a bias might reflect a *selection bias* in the equilibrium filing of DSB cases rather than a bias in DSB rulings conditional on a case being filed. Here we focus on the latter possibility, and consider whether introducing a bias into the DSB's decision might be warranted from an institutional design perspective (perhaps to offset the effects of the selection bias identified in section 2.6).

mance of the institution; and (iv) it decreases the frequency of equilibrium disputes for states $s \in \sigma^{FT}$, which is beneficial for the performance of the institution. The net effect can go either way, depending on the details of the underlying structure, including how many states are in the sets σ^P and σ^{FT} , and how many states in each set are close to the margin where a dispute is triggered in equilibrium. A similar argument can be made for the effects of an anti-trade bias in the DSB objective.

Based on this discussion, we suggest that biasing the DSB’s objective might in principle enhance the performance of the institution, but even if this is the case, the optimal direction of the bias (pro-trade or anti-trade) is far from obvious, and it requires a great deal of information about the fundamental structure of the problem, which governments are unlikely to possess in reality.

Finally we discuss briefly the possibility that the DSB might try to extract information from the observation of the governments’ choices. A full analysis of this question would require solving a complicated three-player signalling model, in which the governments are the (sequential) “senders” and the DSB is the “receiver.” While this goes beyond the scope of the present paper, our intuition suggests that the DSB would be unlikely to extract much information from the governments’ actions. This intuition derives from the fact that in our basic model, where the DSB is not strategic, the governments’ equilibrium actions contain little information about the sign of Γ : recall that in general the exporter files complaints both when $s \in \sigma^{FT}$ and when $s \in \sigma^P$, and thus the fact that a complaint has been filed says little about the sign of Γ ; and when a complaint is filed, the policy chosen by the importer is always $T = P$, so the importer’s choice conditional on a complaint being filed conveys no information at all on the sign of Γ .⁴¹

4.2. Multidimensional policies

Many WTO disputes in practice concern the presence of gaps and vague provisions in the agreement regarding which *policies* are allowed and which are not. For example, there have been a number of disputes concerning whether certain policies should or should not be categorized as prohibited subsidy programs under the GATT/WTO agreement.⁴² Our framework can be

⁴¹In principle there is a third action that could convey some information: the governments’ ex-ante choice of contract. Intuitively, however, this choice can convey little information about the value of Γ for the specific state of the world *ex post* when the dispute arises.

⁴²An example is the U.S – FSC case (WTO, 1999), in which provisions of the U.S. tax code which exempted U.S. multinationals from paying taxes on income earned outside the United States were ruled as prohibited export subsidies and therefore in violation of the WTO Agreement on Subsidies and Countervailing Measures.

extended to examine the potential role of the DSB in completing an incomplete agreement on the policy dimension. In this section we discuss briefly what such an extension might look like.

One natural way to explain why the agreement might be incomplete in specifying the allowed policies is to suppose that the policy space is complex, unlike our basic model where the policy is one-dimensional (and zero-one) in nature. In this spirit, one could assume that the Home country chooses a multidimensional policy, say $t = (t_1, \dots, t_M)$, with each t_i representing a binary policy instrument. For example, the vector t could be a subsidy program, and each t_i a component of the subsidy program (e.g., $t_i = 1$ might mean “foreign earnings are exempted from income taxation,” whereas $t_i = 0$ means “no foreign-earning exemption”).

Suppose that, to achieve the first-best outcome, a certain subset of instruments needs to be constrained. For simplicity suppose the state is deterministic, so that we can abstract from issues of state-contingency. In analogy with our basic model, suppose that it is too costly to describe the specific instruments t_i in the contract, thus leaving only a few contract possibilities: the contract could leave complete discretion on subsidies, or it could impose a “rigid” ban on subsidies (i.e. specify $t_i = 0$ for all i); or it could specify some “vague” rule on subsidies, as for example in the sentence “trade-distorting subsidy programs are not allowed.” Let us also suppose that the sentence “the subsidy program is trade-distorting” is true for $t = (1, \dots, 1)$, false for $t = (0, \dots, 0)$ and undefined for all other values of t . Each of these incomplete contracts (discretionary, rigid, vague) could be associated with an activist DSB that has a mandate to complete the contract in the relevant dimension, or with a non-activist DSB. Finally, it would be natural to assume that, if the DSB has an activist mandate, the Foreign country can choose to dispute one or more of the policy instruments chosen by Home, with the litigation cost increasing in the number of disputed instruments.

It is easy to imagine how a model of this kind could give rise to an institutional choice whereby the contract specifies a vague rule on subsidies and the DSB has an interpretive role, which is broadly consistent with what we observe in the WTO DSB. Also, given that the notions of vagueness, rigidity and discretion (and potential roles of the DSB – interpretation, gap-filling and modification) extend in a natural way to a setting of multidimensional policies, we suspect that a number of the qualitative results derived in our basic model would arise in this setting as well.

One result that might change in a setting with multidimensional policies is that discretion without gap-filling (D_n) may no longer be dominated. If there is heterogeneity across the policy

instruments, so that discretion is more desirable for some instruments than for others, and if the role of the DSB must be the same across instruments, then it might be optimal to have some gaps in the contract (discretion on some instruments) but give the DSB only a mandate to interpret vague provisions, not fill gaps.⁴³

We conclude this discussion by noting that, in a setting with multidimensional policies, legal precedent would have a different role than in our basic model. If the DSB's role is only to interpret the contract or fill a gap for a specific state of the world, as in our basic model, then legal precedent would be applicable only if the same exact state occurs again in the future. But if the DSB's role is also to clarify obligations regarding a given policy instrument, and these obligations apply across states of the world, then legal precedent would apply more broadly than in our basic model. Interestingly, though, recall from section 3 that when precedent applies with higher probability in the future (for example because of a higher ρ), precedent tends to be *less* desirable; hence in a multidimensional policy setting, precedent may end up looking less appealing than in our basic model.

4.3. Consultations and early settlement of disputes

Here we focus on a prominent feature of the WTO dispute settlement process that does not appear in our basic model. The WTO emphasizes consultations between the disputing parties prior to the filing of a formal complaint with the DSB, and indeed Busch and Reinhardt (2006) report that roughly half of WTO disputes are resolved in the consultation stage and roughly two-thirds are resolved at some point prior to a DSB ruling. This raises the question of what distinguishes disputes that are resolved early from disputes that proceed to a formal complaint and DSB ruling, and whether our model could account for this distinction.

Our model suggests – or at least seems broadly compatible with – a few possible reasons for the fact that some disputes are settled early.

First, from a pure enforcement perspective, the consultation stage may facilitate the exchange of information between the two governments concerning the presence or absence of a

⁴³A related, interesting issue is whether the mandate of the DSB can – and if so, whether it should – be made contingent on the policy instruments. This will depend on the cost of specifying policy instruments in the agreement, and on the possible presence of non-separabilities between policy instruments. But it may well be desirable to give different mandates to the DSB according to broad issue areas. So, for example, the DSB might be given a mandate to interpret rules on subsidies, but keep a non-activist role for labor and environmental standards.

suspected deviation.⁴⁴ If consultations can be used to put to rest suspicion that an unambiguous obligation has been violated, then some cases could be settled early. We have abstracted from pure enforcement issues in our model, but such issues could be introduced to accommodate early settlement along the lines just described without altering our basic findings.

Second, even when (as in our formal model) pure enforcement of an unambiguous obligation is not at issue, the consultation stage may lead the two governments to exchange information and learn more about the likely direction of a DSB ruling, which (as our model indicates) may remove the governments' incentives to trigger DSB intervention.⁴⁵

A third possibility is that, in some cases, the governments may agree on some form of compensation at the consultation stage. In models of pre-trial negotiation it is a standard argument that, if side payments are available and the parties have symmetric information (which is the case in our model), they will settle before going to court. In our basic model we assumed away transfers at the dispute stage; indeed, as we remarked earlier (see footnote 8), monetary transfers are almost never used in trade disputes, and effecting compensation through changes in other policies is likely to be costly.⁴⁶ But it is conceivable that, in at least some of the cases, some form of compensation can be enacted without too much cost, and this may help explain why some of the disputes are settled early.

A fourth possibility would be suggested if one introduced a constraint on the governments' ability to monitor and litigate policies. To illustrate, let us suppose that there are many policy instruments (over many industries), and the exporting country has the capacity to scrutinize – and possibly file a complaint – for at most M policies in a given time period. Knowing this, the importing country may find it advantageous to act opportunistically on a large number of

⁴⁴Indeed, Article 4.3 of the WTO Subsidies and Countervailing Measures Agreement states explicitly that the purpose of consultations is to “clarify the facts of the situation and to arrive at a mutually agreed solution,” and WTO rulings have confirmed that this is a central purpose of consultations in WTO disputes generally (see, for example, WTO, 2005).

⁴⁵A related argument can explain why roughly one sixth of all disputes are settled *after* the consultation stage but before the final panel report. In reality the timing of events is as follows: if the consultation stage does not produce a settlement, a DSB panel is formed, there are two rounds of testimony, and then the panel issues an “interim” ruling, at which point the disputants have a further chance to settle; if they do not, there is a final round of testimony and then the DSB issues the final ruling. It is after the interim ruling that the remaining one sixth of early-settlement cases occur. With this timing in mind, our model suggests that a fraction of cases may get settled after the interim ruling but before the final ruling. Suppose that it is costly both to file the initial complaint and to go through the final round of the dispute. Then, if the interim ruling removes some of the uncertainty about the direction of the final ruling, the governments' incentives to trigger the final ruling may be removed too.

⁴⁶Moreover, the WTO explicitly discourages compensation as a means of settling disputes, encouraging instead the parties to simply adjust the disputed policies.

policies (higher than M). Once the exporter chooses the M policies to complain about and requests consultations, the importer might at that point withdraw the policies that have a small chance of being upheld by the DSB (for example those that are in clear violation of the agreement), and stand firm to a final DSB ruling on the policies that have a higher chance of being upheld (for example those for which the agreement is vague and the DSB is more likely to interpret in favor of the importer). This kind of story seems broadly consistent with the evidence that governments often maintain a stock of “documented” suspected WTO violations of their trading partners, and from this stock initiate a limited number of WTO cases each year (see, for example, Davis and Shirato, 2007).

We conclude by mentioning that, according to Busch and Reinhardt (2006), an important determinant of settlement is whether or not third-parties become involved in the dispute (possibly as a way to prevent discriminatory settlements), suggesting that a more complete understanding of this issue would require an extension of our model to a multi-country setting. We leave such an extension for future work.

4.4. Self-enforcement

We have assumed that the DSB is able to enforce the obligations stipulated in the agreement, as well as its own rulings in case of a dispute. In reality, the DSB does not have direct enforcement power. Both the contractually stipulated obligations and the DSB rulings must be self-enforcing, if they are to be relevant. This raises a natural question: Can there still be a contract-completing role for the DSB – as proposed in this paper – in a world of self-enforcing agreements? And can the results of the model be re-interpreted in a meaningful way within such a world?

The standard way to think about self-enforcing agreements is to consider a repeated game played by two or more governments. Consider a world with N countries, and suppose for the sake of simplicity that the one-shot game is separable in the $N(N - 1)/2$ bilateral relationships (see Maggi, 1999, for an example of such a game). Let us focus on a given bilateral relationship, say the one between countries A and B. If countries A and B were sufficiently patient (assuming they have symmetric information), they would be able to enforce the first best policies simply by the threat of bilateral retaliation, without the need of a formal (complete or incomplete) contract of any kind.

But now suppose that A and B are relatively impatient, so that they are not able to enforce

the first best policies. Then they may benefit from the involvement of third countries in the enforcement process: to the extent that third countries can credibly threaten to inflict some punishment in case of violation, countries A and B may be able to sustain more efficient policies. Is there a potential contract-completing role for the DSB in this setting? Not necessarily. If all N countries have symmetric information, then again there is no clear role for a formal contract, let alone a DSB that fills gaps or interprets this contract.

Finally, suppose that, while countries A and B know the relevant payoff functions for their bilateral relationship (let us call it $\Gamma^{AB}(s)$), third countries do not. In this case, a role for a formal (written) contract and for a contract-completing institution *can* arise. A formal contract that stipulates obligations between A and B is a way to inform third parties of what A and B would like to enforce – which third countries are not otherwise able to infer because they do not know A and B’s true preferences. And if it is too costly to write a complete formal contract, then a role for a DSB that helps complete the incomplete contract may arise: in this setting, the DSB’s role would be to transmit its own (imperfect) guess of the Γ^{AB} -maximizing policy to third countries, so that third countries can lend (off-equilibrium) enforcement power to countries A and B.

Hence, we believe that the main ideas of our paper are relevant even from the perspective of self-enforcing agreements, provided that they are re-interpreted in the broader context of a multi-country setting in which the DSB serves to muster enforcement power from third countries by transmitting information to them.

5. Conclusion

We conclude by discussing two issues. The first concerns an often-heard argument for why trade agreements contain vague provisions and contractual gaps. It is often argued that, when governments are at the bargaining table and encounter a “sensitive” issue, such as whether to allow for trade protection in contingencies where a government may be subject to domestic political pressures, they may find it hard to agree on firm policy obligations, and choose instead to leave the agreement vague – or even to leave a gap – on that point. In other words, leaving vagueness or discretion in the contract may be a way to “agree to disagree,” rather than an optimal response to contracting costs. The question is: are the points we made in this paper still valid if (part or all of) the contractual incompleteness in trade agreements is due to political sensitivities of the sort just described?

We believe our framework is able – in spite of its simplicity – to capture this “political-sensitivities” explanation for the incompleteness of trade agreements. Indeed, we can interpret the states $s \in \sigma^P$ as states where the political cost of imposing free trade in the importing country exceeds the benefit for the exporting country, making it politically efficient to leave discretion to the importing country in those states. The next observation is that, while it is easy to rationalize the presence of discretion in the agreement based on the political-sensitivities theory, it is less easy to rationalize the *vagueness* of an agreement based solely on that argument: suppose there are no contracting costs; then the optimal contract will be one that specifies unambiguously the states in which free trade must be applied ($s \in \sigma^{FT}$), and leaves discretion in all other states ($s \in \sigma^P$). Moreover, the political-sensitivities theory – at least in its simple form considered above – cannot explain an activist role for the DSB: if it is politically efficient to leave discretion in certain states of the world, then it is not clear why the governments would give a mandate for the DSB to remove that discretion ex-post.⁴⁷

The second issue we would like to address is whether our model of trade agreements can be interpreted to apply more broadly to other legal settings. Do the same insights extend to the optimal design of the role of courts more generally?

While a similar model structure could perhaps be applied to other types of international agreements, we do not believe that the insights of the paper can be applied in a straightforward manner to the optimal design of domestic legal institutions, for at least three reasons. First, our assumption that ex-post transfers are not available is not realistic in a domestic setting; the use of monetary compensation as both a settlement tool and a legal remedy is pervasive in domestic settings. Second, at a more methodological level, we have adopted an institutional design approach whereby the structure of the contract and the role of the court are jointly designed ex-ante by the contracting parties (the governments). In a domestic setting, the design of the judicial system and the design of private contracts are decoupled: when two (or more) parties design a contract, they do so taking the court system and its procedures as given; and the court system is designed by legislators (or constitution writers), presumably taking

⁴⁷A slightly different version of the politically-sensitivities theory is that, in some contingencies, free trade may be the first-best policy but is not self-enforcing, because under those contingencies the importing government has a strong incentive to defect. Then, the agreement may leave discretion under those contingencies in order to preserve self-enforceability. The argument made in the text can be applied also to this version of the politically-sensitivities theory, by re-interpreting the notion of (political) efficiency as (political) efficiency subject to the self-enforcement constraint: again, if it is constrained-efficient to leave discretion in certain states of the world, it is not easy to rationalize a vague contract, or even less an activist role for the DSB.

into account that there is a large and heterogenous population of potential users of the court system.

Finally, trade agreements involve a relatively small number of players (the governments), each of whom can expect to utilize the court system relatively frequently, while in most domestic settings there is a large population of “small” players (the individuals), each of whom can expect to utilize the court system only rarely. This difference can itself produce very different implications. Consider for example the role of legal precedent. A key insight that emerges from our model is that legal precedent has the harmful effect of increasing the equilibrium frequency of disputes, whereas its main benefit lies in eliminating the duplicative costs of disputes. But the increase in the equilibrium frequency of disputes is a consequence of the fact that, when for example the Foreign government contemplates filing, it takes into account that with a certain probability it will find itself in the same situation again in the future. In a domestic legal setting where an individual contemplates using the court system only rarely, this consideration is unlikely to weigh significantly in the filing decision. This suggests that the main shortcoming of legal precedent that is highlighted by our model would be much less important in a domestic setting, and hence legal precedent would be more likely to be desirable in a domestic setting than in an international setting. More broadly, the “small numbers” feature of the trade-agreements setting is likely to distinguish it in a variety of ways from the typical domestic legal setting.⁴⁸

For these reasons, we believe that the simple framework adopted in our paper would have to be substantially modified before it can be fruitfully applied to the analysis of domestic legal systems. Introducing the necessary modifications, and exploring their implications for the optimal design of legal systems at the domestic and international level, seems to us to be an ambitious and important area for future research.

⁴⁸There do exist some domestic settings in which a few “large” players establish a private arbitration system within their contractual arrangement, for example in the area of labor relations or business-to-business transactions. Our framework is potentially applicable to these settings (we thank Eric French for suggesting this possibility to us).

6. Appendix

Proof of Remark 1:

We will prove that the equilibrium frequency of disputes is higher under D_g^P than under D_g . The proof of the statement concerning V_i^P and V_i is analogous.

We work backwards through time. Denote the *Period-1* realized state by s' . Observe first that, if there is no *Period-1* filing, then *Period 2* behaves like the static model under D_g for all s . If instead there is a *Period-1* filing with associated *Period-1* DSB ruling $T_1^{DSB}(s')$, then in *Period 2* the contract specifies $T_2(s') = T_1^{DSB}(s')$ and the DSB has no active role for $s = s'$, while for $s \neq s'$, *Period 2* behaves exactly like the static model under D_g .

Now consider *Period 1*. The exporter government files a complaint if and only if $T = P$ and the expected benefit of filing exceeds the cost of filing. Denote by $B_P^E(s')$ the expected *Period-2* value to the exporter of the precedent that would be set by a DSB ruling, conditional on the realized *Period-1* state s' (but before the filing decision is made and the *Period-1* ruling is known). Then the exporter files a complaint in *Period 1* if and only if $T = P$ and

$$\Pr(\text{DSB ruling is } FT \mid s') \cdot |\gamma^*(s')| + \delta B_P^E(s') > c^*. \quad (\text{F1})$$

It can be shown that the term $B_P^E(s')$ can be written as

$$B_P^E(s') = \begin{cases} p(s') \cdot c^* & \text{if } s' \in \hat{\sigma}^P \cup \hat{\sigma}^{FT} \\ p(s') \cdot [qk(s')|\gamma^*(s')|] & \text{if } s' \in \sigma^P \setminus \hat{\sigma}^P \\ -p(s') \cdot [qk(s')|\gamma^*(s')|] & \text{if } s' \in \sigma^{FT} \setminus \hat{\sigma}^{FT}. \end{cases} \quad (6.1)$$

where the sets $\hat{\sigma}^{FT}$ and $\hat{\sigma}^P$ are as defined in the static section. To understand this expression, observe first that precedent is only relevant if the *Period-2* state realization is also s' , which occurs with probability $p(s')$. This explains why each term on the right-hand-side is multiplied by $p(s')$. For $s' \in \hat{\sigma}^P \cup \hat{\sigma}^{FT}$, the exporter would have filed in *Period 2* if not for the precedent created by the *Period-1* DSB ruling, and so the exporter saves the *Period-2* filing cost c^* . For $s' \in \sigma^P \setminus \hat{\sigma}^P$, the exporter would not have filed in *Period 2* and the importer would have chosen $T = P$, so the exporter enjoys a *Period-2* benefit of $|\gamma^*(s')|$ from the precedent created by the *Period-1* DSB ruling provided the ruling is in error, which happens with probability $qk(s')$. Finally, for $s' \in \sigma^{FT} \setminus \hat{\sigma}^{FT}$, the exporter would not have filed in *Period 2* and the importer would have chosen $T = FT$, so the exporter suffers a *Period-2* loss in the amount $\gamma^*(s')$ from the precedent created by the *Period-1* DSB ruling if the ruling is in error, which happens with probability $qk(s')$.

Next consider the importer government's *Period-1* policy choice. Denote by $B_{\mathcal{P}}^M(s')$ the expected *Period-2* value to the importer, given the realized *Period-1* state s' , of the precedent set by a DSB ruling. The importer chooses $T = P$ if either (F1) fails or if (F1) holds and the expected benefit from choosing P exceeds the cost of a DSB dispute:

$$\Pr(\text{DSB ruling is } P \mid s') \cdot \gamma(s') + \delta B_{\mathcal{P}}^M(s') > c. \quad (\text{P1})$$

It can be shown that the term $B_{\mathcal{P}}^M(s')$ can be written as

$$B_{\mathcal{P}}^M(s') = \begin{cases} p(s') \cdot c & \text{if } s' \in \hat{\sigma}^{FT} \cup \hat{\sigma}^P \\ p(s') \cdot [qk(s')\gamma(s')] & \text{if } s' \in \sigma^{FT} \setminus \hat{\sigma}^{FT} \\ -p(s') \cdot [qk(s')\gamma(s')] & \text{if } s' \in \sigma^P \setminus \hat{\sigma}^P. \end{cases} \quad (6.2)$$

This expression has an analogous interpretation to that described above for $B_{\mathcal{P}}^E(s')$.

We can now derive the equilibrium *Period-1* actions of the governments under D_g when DSB rulings set precedent:

1. For $s \in \sigma^{FT}$: if $qk(s) < \frac{c}{(1+\delta p(s))\gamma(s)}$ then $T_1 = FT$ and the DSB is not invoked in *Period 1*; if instead $qk(s) > \frac{c}{(1+\delta p(s))\gamma(s)}$ then $T_1 = P$ and the DSB is invoked in *Period 1*.
2. For $s \in \sigma^P$: if $qk(s) < \frac{c^*}{(1+\delta p(s))|\gamma^*(s)|}$ then $T_1 = P$ and the DSB is not invoked in *Period 1*; if instead $qk(s) > \frac{c^*}{(1+\delta p(s))|\gamma^*(s)|}$ then $T_1 = P$ and the DSB is invoked in *Period 1*;

It is now direct to compare the *Period-1* equilibrium actions under $D_g^{\mathcal{P}}$ and under D_g , and conclude that the number of states in which a dispute is triggered is larger under $D_g^{\mathcal{P}}$. The claim follows. **QED**

Proof of Proposition 3:

Let us first focus on the D_g and $D_g^{\mathcal{P}}$ institutions. Define $\hat{\sigma}_{\mathcal{P}}^{FT}$ as the set of states s such that $s \in \sigma^{FT}$ and $qk(s) > \frac{c}{(1+\delta p(s))\gamma(s)}$, and $\hat{\sigma}_{\mathcal{P}}^P$ as the set of states s such that $s \in \sigma^P$ and $qk(s) > \frac{c^*}{(1+\delta p(s))|\gamma^*(s)|}$. Thus, $\hat{\sigma}_{\mathcal{P}}^{FT}$ is the set of states for which (i) FT is efficient and (ii) in *Period 1* there is a dispute in equilibrium under $D_g^{\mathcal{P}}$. And $\hat{\sigma}_{\mathcal{P}}^P$ is the set of states for which (i) P is efficient and (ii) in *Period 1* there is a dispute in equilibrium under $D_g^{\mathcal{P}}$. Thus, the set of states where *Period-1* filing occurs under D_g is $\hat{\sigma} = \hat{\sigma}_{\mathcal{P}}^P \cup \hat{\sigma}_{\mathcal{P}}^{FT}$, and the set of states for which *Period-1* filing occurs under $D_g^{\mathcal{P}}$ but not under D_g is $\hat{\Delta} = (\hat{\sigma}_{\mathcal{P}}^P \cup \hat{\sigma}_{\mathcal{P}}^{FT}) \setminus (\hat{\sigma}_{\mathcal{P}}^P \cup \hat{\sigma}_{\mathcal{P}}^{FT})$.

Now refer to the expression (3.2) for $L(D_g^{\mathcal{P}}) - L(D_g)$, and note first that, if $q = 0$, the sets $\hat{\Delta}$ and $\hat{\sigma}$ are both empty, and hence $L(D_g^{\mathcal{P}}) - L(D_g) = 0$. As q increases from zero, the first of these two sets to become non-empty is $\hat{\Delta}$, and hence $L(D_g^{\mathcal{P}}) - L(D_g) > 0$. To see this, note

that $\hat{\Delta}$ is the set of states s such that either (i) $s \in \sigma^{FT}$ and $\frac{c}{q(1+\delta p(s))\gamma(s)} < k(s) < \frac{c}{q\gamma(s)}$, or (ii) $s \in \sigma^P$ and $\frac{c^*}{q(1+\delta p(s))|\gamma^*(s)|} < k(s) < \frac{c^*}{q|\gamma^*(s)|}$, while the set $\hat{\sigma}$ is the set of states s such that either (i) $s \in \sigma^{FT}$ and $k(s) > \frac{c}{q\gamma(s)}$, or (ii) $s \in \sigma^P$ and $k(s) > \frac{c^*}{q|\gamma^*(s)|}$. As q increases from zero, the thresholds $\frac{c}{q(1+\delta p(s))\gamma(s)}$ and $\frac{c}{q\gamma(s)}$ drop down from infinity. It is clear from inspection of the inequalities above that the first set to become nonempty is $\hat{\Delta}$.

Next note that, if $q = 1$, filing occurs for all states both in the presence and in the absence of precedent, and therefore $\hat{\sigma}$ includes all states and $\hat{\Delta}$ is empty, and hence $L(D_g^P) - L(D_g) < 0$. It is then a small step to prove the result.

Next focus on V_i and V_i^P . Let $\bar{\sigma}_P^{FT}$ be the set of states s such that $s \in \sigma^{FT} \setminus s^0$ and $qk(s) > \frac{c}{(1+\delta p(s))\gamma(s)}$, and $\bar{\sigma}_P^P$ the set of states s such that $s \in \sigma^P \setminus s^1$ and $qk(s) > \frac{c^*}{(1+\delta p(s))|\gamma^*(s)|}$. Clearly, we have $\bar{\Delta} \equiv (\bar{\sigma}_P^P \cup \bar{\sigma}_P^{FT}) \setminus (\bar{\sigma}^P \cup \bar{\sigma}^{FT})$ and $\bar{\sigma} \equiv (\bar{\sigma}^P \cup \bar{\sigma}^{FT})$. The claim is then easily shown by following similar steps as the ones for D_g and D_g^P . **QED**

Proof of Proposition 4:

We prove the claim for $L(D_g^P) - L(D_g)$; the argument for $L(V_i^P) - L(V_i)$ runs analogously. The key observations are: (i) when $\delta = 0$, the set $\hat{\Delta}$ is empty; and (ii) the set $\hat{\sigma}$ is independent of δ , while the set $\hat{\Delta}$ is weakly increasing in δ and as $\delta \rightarrow \infty$ the set $\hat{\Delta}$ includes all states s that are not in $\hat{\sigma}$. Consider the case where q lies in an intermediate range, so that the set $\hat{\sigma}$ is non-empty but not “too large” (i.e., contains a small subset of all states); when $\delta = 0$, we have by (i) that $\hat{\Delta}$ is empty, and so (3.2) implies $L(D_g^P) = L(D_g)$; for $\delta > 0$ but sufficiently small, (3.2) implies $L(D_g^P) < L(D_g)$; and for $\delta \rightarrow \infty$, we have by (ii) that $\hat{\Delta}$ includes all states s that are not in $\hat{\sigma}$, and so (3.2) implies $L(D_g^P) > L(D_g)$; hence in this case introducing precedent is (strictly) desirable if and only if δ is not too high. The claim follows immediately. **QED**

References

- Anderlini, L., Felli, L. and A. Postlewaite (2006), "Should Courts Always Enforce What Contracting Parties Write?" mimeo.
- Anderlini, L., Felli, L. and A. Postlewaite (2007), "Courts of Law and Unforeseen Contingencies," *Journal of Law, Economics and Organization*, 23, 662-684.
- Bagwell, Kyle and Robert W. Staiger (1990), "A Theory of Managed Trade," *American Economic Review*, September.
- Bagwell, Kyle and Robert W. Staiger (2002), *The Economics of the World Trading System*, MIT Press.
- Battaglini, Marco (2002), "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70(4), pp.1379-1402, July.
- Battigalli, Pierpaolo and Giovanni Maggi (2002), "Rigidity, discretion and the costs of writing contracts," *The American Economic Review*, vol. 92(4), pp. 798-817.
- Battigalli, Pierpaolo and Giovanni Maggi (2003), "International Agreements on Product Standards: An Incomplete-Contracting Theory," NBER Working Paper No. 9533.
- Busch, Mark L. and Eric Reinhardt (2006), "Three's a Crowd: Third Parties and WTO Dispute Settlement," *World Politics* 58, April, pp. 446-77.
- Bustos, Alvaro (2007a), "Litigation and the Optimal Combination of Vague and Precise Clauses in Contracts," mimeo, Northwestern University.
- Bustos, Alvaro (2007b), "A dynamic theory of Common Law: the Rule-Making Role of Courts," mimeo, Northwestern University.
- Davis, Christina and Yuki Shirato (2007), "Firms, Governments and WTO Adjudication: Japan's Selection of WTO Disputes," *World Politics*, 59, no. 2, January: pp. 274-313.
- Fon, Vincy and Francesco Parisi (2007), "On the optimal specificity of legal rules," *Journal of Institutional Economics*, vol. 3(2), 147-164.

- Gennaioli, Nicola and Andrei Shleifer (2007), "The Evolution of Common Law," *Journal of Political Economy*, vol. 115(1), February, 43-68.
- Goldstein, Judith L. and Richard H. Steinberg (2007), "The Rise of Judicial Liberalization at the WTO," mimeo, May 1, Stanford University.
- Horn, Henrik, Maggi, Giovanni and Robert W. Staiger (2006), "Trade Agreements as Endogenously Incomplete Contracts," NBER Working Paper No. 12745, December.
- Horn, Henrik and Petros C. Mavroidis (2003), "*United States - Safeguard Measures on Imports of Fresh, Chilled or Frozen Lamb Meat from New Zealand and Australia: What Should be Required of a Safeguard Investigation?*," in H. Horn and P.C. Mavroidis (eds.) *The WTO Case Law of 2001: The American Law Institute Reporters' Studies*, Cambridge University Press (Cambridge).
- Howse, Robert and Robert W. Staiger (2005), "United States–Anti-Dumping Act of 1916 (Original Complaint by the European Communities) – Recourse to arbitration by the United States under 22.6 of the DSU, WT/DS136/ARB, 24 February 2004: A legal and Economic Analysis," *World Trade Review* 4:2, 295-316.
- Jackson, John H. (2006), *Sovereignty, the WTO, and Changing Fundamentals of International Law*, Cambridge University Press.
- Klimenko, Mikhail, Ramey, Garey and Joel Watson (2008), "Recurrent Trade Agreements and the Value of External Enforcement," *Journal of International Economics*, 74(2), pp. 475-499, March.
- Krishna, Vijay and John Morgan (2001), "A Model of Expertise," *Quarterly Journal of Economics*, 116(2), pp 747-775, May.
- Lawrence, Robert Z. (2003), *Crimes & Punishments? - Retaliation Under the WTO*, Institute for International Economics (Washington, D.C.).
- Maggi, Giovanni (1999), "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review*, 89, 190-214.
- Posner, Richard A. (1973), *The Economic Analysis of Law*, Boston: Little Brown.

- Sykes, Alan O. (2000), "The Remedy for Breach of Obligations Under the WTO Dispute Settlement Understanding: Damages or Specific Performance?" in Marco Bronckers and Reinhard Quick (eds.) *New Directions in International Economic Law*, Kluwer Law International (The Hague).
- Schwartz, Warren F. and Alan O. Sykes (2002), "The Economic Structure of Renegotiation and Dispute Resolution in the World Trade Organization," *Journal of Legal Studies* 31.
- Shavell, Steven (2006), "On the Writing and the Interpretation of Contracts," *Journal of Law, Economics & Organization*, vol. 22(2), October, 289-314.
- WTO (1999), *United States - Tax Treatment for "Foreign Sales Corporations" - Report of the Panel*, WT/DS108/R, August 10.
- WTO (2001), *United States - Safeguard Measures on Imports of Fresh, Chilled or Frozen Lamb Meat from New Zealand and Australia*, WT/DS177/AB/R, WT/DS178/AB/R, May 1.
- WTO (2005), *European Communities - Customs Classification of Frozen Boneless Chicken Cuts - Report of the Appellate Body*, WT/DS269/AB/R and WT/DS286/AB/R.
- WTO (2007), *World Trade Report*, Working Draft, Geneva.
- WTO (2007a), *United States - Final Anti-Dumping Measures on Stainless Steel from Mexico - Report of the Panel*, WT/DS344/R, December 20.
- WTO (2008), *United States - Final Anti-Dumping Measures on Stainless Steel from Mexico - Report of the Appellate Body*, WT/DS344/AB/R, April 30.

Pro-trade bias .