

Does WTO Dispute Settlement Enforce or Inform?

THOMAS SATTLER, GABRIELE SPILKER AND
THOMAS BERNAUER*

Whereas some researchers emphasize how World Trade Organization (WTO) dispute settlement reduces complexity and clarifies legislation, others argue that dispute rulings promote co-operation by providing an enforcement mechanism. This article identifies empirical implications from these distinct arguments and tests them on WTO disputes from 1995 to 2006. The study's analytical approach combines a three-step coding of dispute escalation with a strategic bargaining model and statistical backwards induction to account for governments' forward-looking behavior. It finds strong support for the argument that WTO dispute settlement primarily serves as an enforcement device. It finds much less support for the argument that dispute settlement reduces complexity and clarifies trade law. These results suggest that the role of WTO dispute settlement in generating information on acceptable trade policy standards is less relevant than proponents of the complexity argument tend to assume.

The general trend in recent decades towards stronger legalization of international relations has been associated with an increasing prevalence of judicial procedures in international institutions.¹ Important recent examples include the International Criminal Court, dispute settlement procedures in global environmental agreements and bilateral agreements on trade and investment, and the World Trade Organization's (WTO) dispute settlement system.²

The political science and international relations literatures regard dispute settlement as an important instrument for promoting international co-operation, but researchers emphasize different views of how dispute settlement affects co-operation, highlighting either enforcement or management. This article assesses the relative importance of these two mechanisms for WTO dispute resolution and their role in promoting international co-operation.

The enforcement perspective assumes that many international co-operative efforts have the structure of mixed-motive games.³ While countries are interested in the collective goods that result from co-operation, they maintain incentives to shirk or cheat, particularly when implementing an international commitment that is costly, and when free riding is associated with large benefits. In this context, dispute settlement is considered part of the monitoring and enforcement process. It is a mechanism that protects co-operating countries from non-compliance by others and deters countries

* Department of International Relations, LSE (email: t.sattler@lse.ac.uk); Center for Comparative and International Studies, ETH Zurich (emails: gabriele.spilker@ir.gess.ethz.ch and thbe0520@ethz.ch, respectively). A version of this manuscript has been presented at the 4th Annual Conference on the Political Economy of International Organizations, 27–29 January 2011, Zurich. Jeffrey Kucik and Laura Zoratto gave valuable comments on an earlier draft. Marianne Furrer, Quynh Nguyen and Gwen Tiernan provided excellent research assistance. Finally, we thank four anonymous reviewers and the journal's editor, Kristian Skrede Gleditsch, for extremely helpful advice that improved this research.

¹ Goldstein et al. 2000; Koremenos, Lipson, and Snidal 2001.

² Guzman 2002; Posner and Yoo 2005; Rosendorff 2005; Rosendorff and Milner 2001.

³ Barrett 2006; Downs, Rocke, and Barsboom 1996.

from violating their commitments in the first place. Dispute settlement thus promotes co-operation by deterring, detecting and punishing violators.⁴

The second – managerial – perspective views dispute settlement procedures primarily as information and rule clarification devices. International institutions are highly complex systems of behavioral standards and regulations.⁵ International rules are sometimes ambiguous or so complicated that governments do not fully understand whether their current policies and practices are really in compliance with those rules. In other words, this perspective views dispute settlement as part of a compliance management process under conditions of incomplete contracts.⁶ The basic modus operandi of dispute settlement mechanisms (DSMs) involves helping states implement their commitments by reducing complexity and providing information, rather than detecting and punishing violations of international rules.

This article examines the theoretical and empirical implications of these perspectives with regard to the WTO DSM. It emphasizes the importance of strategic and forward-looking government behavior in an effort to understand the dynamics of dispute settlement across three key phases of the process: the consultation phase, in which the parties have an opportunity to achieve an ‘out of court’ settlement; the formal WTO panel adjudication phase; and the ‘post-verdict’ phase, which handles non-compliance with WTO DSM verdicts.

The first (enforcement) perspective implies that disputes that are associated with greater domestic political stakes – in which the disputing countries’ positions are backed by powerful domestic interests – are more prone to escalation and harder to solve. Taking forward-looking behavior into account, however, we expect that governments will anticipate the high escalation risk, especially the fact that a dispute might receive increased public attention after a WTO panel ruling. Hence they carefully decide which of the potentially difficult cases to further pursue after the consultation stage.

The second (managerial) perspective implies that disputes over more complex issues should end after a WTO panel ruling, because a formal ruling presumably clarifies the legal situation and thus helps countries solve the problem and restore compliance. Since forward-looking governments anticipate this low escalation risk, they should not worry more about escalation risks involved in requesting panel rulings in complex cases than in non-complex cases.

Our results offer strong support for the political stakes hypothesis and less support for the complexity hypothesis. Disputes associated with greater domestic political stakes are particularly prone to escalate further than a panel ruling. Similarly, complex disputes escalate more frequently to the panel stage and are also more likely to escalate into non-compliance disputes. The latter result contradicts the complexity hypothesis in part: complex disputes, according to this hypothesis, should escalate to the panel ruling stage, but not beyond. Our results also show that taking governments’ strategic behavior into account is crucial for understanding dispute resolution dynamics in the WTO. Most notably, we find that forward-looking governments anticipate that settling disputes associated with greater political stakes becomes more difficult the further they have advanced in the dispute settlement process. Hence they often pre-empt such loss of control by settling disputes associated with greater domestic political stakes early on

⁴ Mitchell 2006; Downs, Rocke, and Barsoom 1996.

⁵ Goldstein et al. 2000; Young 1994.

⁶ Chayes and Chayes 1993.

through consultations. Overall, these results offer important insights into the role that dispute settlement plays in international trade policy.

Our analysis of WTO dispute settlement makes three innovations to the existing literature.⁷ First, most of the literature on WTO dispute settlement uses a binary definition of dispute escalation – whether a dispute reaches the stage of a panel ruling or not.⁸ The analysis of our arguments requires that we take into account what happens after a panel ruling. Hence we conceptualize dispute resolution as a three-step process. Our dependent variable measures whether a dispute is resolved during the consultation phase, whether it ends in compliance after a panel or appellate body ruling, or whether it further escalates into a dispute about compliance with the WTO verdict.

Secondly, measuring domestic political stakes and the complexity of disputes requires a systematic classification of disputes according to these theoretical constructs. We characterize complexity in terms of the properties of disputed trade barriers. Greater domestic political stakes are characterized by the type of economic sector that is subject to a dispute. Existing studies on dispute escalation in the WTO are chiefly interested in the implications of country characteristics, such as power, trade dependence, legal capacity or democracy.⁹ Some studies also include a partial classification of disputes (for example, agricultural disputes), but do not establish a consistent classification across all relevant types of disputes.¹⁰

The third innovation in our article is that we take into account governments' forward-looking behavior and design our empirical testing strategy accordingly. Forward-looking behavior means that government choices at a given point in the dispute settlement process depend, among other things, on their expectations about the future choices of their opponent(s). For instance, if faced with the choice of whether to escalate a dispute with greater domestic political stakes to the panel stage, forward looking governments should take a different decision than 'naive' – not forward-looking – governments. Knowing that a dispute associated with greater domestic political stakes is difficult to resolve once it has escalated to the panel stage, forward-looking governments should seek to avoid escalation in the first place. To model such forward-looking behavior, we set up a strategic empirical model that reflects the essential steps in the WTO dispute settlement process.¹¹ The empirical model is estimated using statistical backwards induction.¹²

BARGAINING DYNAMICS ACROSS DISPUTE TYPES

Since 1995, the WTO has provided its member states with a highly structured process for settling international trade disputes. The WTO's dispute settlement mechanism is used

⁷ For a comprehensive account of the literature on WTO dispute settlement, see Bernauer, Elsig, and Pauwelyn 2012.

⁸ Guzman and Simmons 2002.

⁹ Allee 2010a; Allee 2010b; Davis and Bermeo 2009; Reinhardt 2001; Sattler and Bernauer 2011.

¹⁰ Exceptions are studies by Francois, Horn, and Kaunitz 2008 and Guzman and Simmons 2002. Francois, Horn, and Kaunitz (2008) include a full classification of sectors, but only look at dispute initiation (rather than escalation), and they focus on developing countries only. Guzman and Simmons (2002) code disputes according to their 'all or nothing' character. They argue that disputes of this type are more likely to escalate to the panel stage (escalation is coded as a two-step process) because it is more difficult to arrange gradual concessions and transfer payments in such cases. This argument implies that, for example, disputes about tariff levels are easier to settle than disputes about environment, health and safety regulations.

¹¹ Signorino 1999.

¹² Bas, Signorino, and Walker 2008.

when one or more member countries believe that trade practices by other WTO members (for example, tariffs, import restrictions, export subsidies) violate existing trade rules. Whereas the dispute settlement procedure of the General Agreement on Tariffs and Trade (GATT) system – the predecessor of the WTO until 1995 – left accused countries with many ways to delay or stall dispute settlement proceedings, the WTO's DSM provides for a multistage process with clearly defined time lines and virtually no opportunities for accused countries to block the process.¹³

The WTO's DSM is crucial for the functioning of the global trading system. In the WTO's own words:

Dispute settlement is the central pillar of the multilateral trading system, and the WTO's unique contribution to the stability of the global economy. Without a means of settling disputes, the rules-based system would be less effective because the rules could not be enforced. The WTO's procedure underscores the rule of law, and it makes the trading system more secure and predictable.¹⁴

Empirical research confirms that the WTO DSM plays a major role in promoting co-operation and trade.¹⁵ However, different mechanisms can lead to this outcome. The statement above appears to allude to both the enforcement and managerial perspectives. But it does not tell us exactly what we should look for in real WTO dispute settlements in order to better understand the mechanisms through which dispute settlement relates to international co-operation.

In the next two sections, we elaborate on both the enforcement (political stakes) and managerial (complexity) perspectives to derive distinct empirical implications to be tested in the subsequent sections. In particular, we examine theoretically how dispute resolution dynamics should look if WTO dispute settlement primarily addresses enforcement or managerial concerns. We argue that enforcement problems mostly arise when high domestic political stakes are involved. Managerial difficulties in international trading relations generally originate from a high degree of complexity and uncertainty regarding the exact rights and obligations associated with international trade rules. As we will see below, the two perspectives have fundamentally different implications for expected dispute settlement dynamics and outcomes. Furthermore, the two distinct theoretical arguments, which we derive from the enforcement and managerial perspectives, distinguish between two scenarios. First, we discuss how both domestic political stakes and complexity affect dispute escalation when governments behave non-strategically. Secondly, we relax this assumption and argue how the dynamics of dispute escalation should look when governments behave strategically; that is, when they are forward looking in the sense that they take their opponents' possible future actions into account.

¹³ The parties to a dispute are allotted a period of sixty days for consultations and mediation. If they fail to reach an agreement, the plaintiff can ask for a dispute settlement panel to be established. This panel, which consists of several international trade experts, must be set up within forty-five days and has six months to issue its report. The Dispute Settlement Body (DSB) has to adopt the report within sixty days and can reject it only by consensus. The entire process takes up to one year, and approximately another three to four months if an appeal is brought before the WTO's appellate body. After the adoption of the report, the defendant country is given a 'reasonable' period of time to implement the panel's recommendations. If the defendant does not comply with the ruling, the complainant can ask the DSB to permit retaliatory measures, for example tariffs on imports from the defendant country. As of September 2011, more than 420 dispute settlement cases had entered the WTO system, many of which have been resolved.

¹⁴ http://www.wto.org/english/thewto_e/whatis_e/tif_e/disp1_e.htm.

¹⁵ Bechtel and Sattler 2012.

WTO DISPUTE SETTLEMENT AS INTEREST GROUP POLITICS

According to the enforcement perspective, many international co-operative efforts have the structure of mixed-motive games.¹⁶ Countries are interested in the benefits of co-operation, but they also have incentives to shirk or cheat. This is mainly the case when implementing an international commitment is costly or if free riding is associated with large benefits. In the logic of the enforcement perspective, dispute settlement is considered part of the monitoring and enforcement process that protects co-operating countries from non-compliance by others and deters countries from violating their commitments in the first place. Dispute settlement thus promotes co-operation by deterring non-compliance, since it increases the costs associated with non-cooperation.¹⁷ When do countries engage in non-compliance, even though it increases the risk of a costly dispute settlement process?

Even a casual look at news media reporting on WTO trade disputes quickly reveals that some disputes are associated with great political stakes, in the sense that they attract a great deal of attention and are associated with extensive and controversial public debates in the countries involved. Examples include EU-US disputes over the use of growth hormones in beef production and genetically modified (GM) crops. Strong political attention, in turn, limits the freedom of policy makers to strike bargains – in particular, bargains that are perceived as unfavorable by domestic producers, interest groups or the general domestic audience.¹⁸ Other disputes are hardly noticed by the news media and the domestic public. Examples include a dispute between the United States and Indonesia on ‘Measures Affecting the Production and Sale of Clove Cigarettes’ and a dispute between Chile and Mexico over ‘Measures Affecting the Import of Matches’. In such cases, policy makers have more room to maneuver and negotiate with the opponent(s) in a dispute and make concessions.

The domestic political stakes associated with a disputed trade policy affect governments’ motivations and thus their willingness and ability to co-operate with trading partners. Defendant governments are more likely to resist demands from complainants in disputes with greater domestic political stakes, because domestic constituencies will punish them for backing down and exposing domestic industries to more economic competition if the disputed trade-restricting policy is weakened or removed. Similarly, to the extent that a dispute is associated with greater political stakes, the complainant government faces higher domestic political costs if it gives in and accepts that the defendant makes only some concessions or does not lift the disputed trade-inhibiting policy at all.

To identify the political stakes of a dispute *ex ante* – that is, independently of the dynamics of a trade dispute once it has been launched – the existing literature has focused on economic sector characteristics because disputes concerning some sectors attract stronger political attention than others. Davis and Shirato, for instance, show that sector-specific factors influence government decisions during trade conflicts.¹⁹ The magnitude of domestic political stakes in any given sector should then be a function of the importance of that sector for the respective country.²⁰

¹⁶ Barrett 2006; Downs, Rocke, and Barsoom 1996.

¹⁷ Downs, Rocke, and Barsoom 1996; Mitchell 2006.

¹⁸ Domestic political stakes also influence the choice where to settle a specific dispute (Allee and Huth 2006; Davis 2008). Allee and Huth (2006), for example, show that governments opt for international instead of bilateral dispute settlement if they are in need of ‘political cover’ to settle a controversial dispute.

¹⁹ Davis and Shirato 2007.

²⁰ Francois, Horn, and Kaunitz 2008.

However the political stakes may also depend on other factors, notably each sector's ability to organize politically, form alliances with other stakeholders, publicize the dispute, and lobby for or against trade-restricting policies or practices.²¹ In the case of the EU-US dispute over GM crops, a heterogeneous but very large and well-organized political coalition of European farmers, consumer organizations, environmental groups and left parties has emerged. Although the importance of the issue in purely economic terms is quite modest, the high political stakes associated with the dispute have made it virtually impossible for the EU and its member states to accommodate US requests to open Europe's market to GM crops.²²

Under conditions of higher domestic political stakes, it should thus be less likely that a defendant government will give in and accept an adverse ruling. In general, a formal WTO ruling increases political pressure on the defendant country, if it is found guilty, to remove the disputed trade restrictions because failure to comply with WTO rules would adversely affect the defendant country's international reputation.²³ Such negative reputation effects could, in turn, negatively affect the defendant country's ability to reach advantageous trade or other agreements in the future. These negative reputation effects, however, could be offset by domestic political pressure when a very politically sensitive sector is affected. Hence we should expect trade disputes to be more prone to escalation if domestic political stakes are high.

So far, we have not considered that governments behave strategically when settling disputes. Strategic behavior is important because the exact implications of the enforcement perspective differ when we assume that governments are forward looking (that is, that they behave strategically). When governments decide whether or not to request the establishment of a dispute settlement panel, which constitutes a major step towards dispute escalation in the WTO, they are likely to pay close attention to the other country's expected behavior in subsequent phases of the process.²⁴ We therefore need to discuss how the effect of domestic political stakes on the likelihood of escalation may vary at different stages of the dispute settlement process. This is necessary because governments not only consider immediate gains and losses when deciding what to do at a particular point in the dispute process, but also base their decision on what they expect to happen in the future.

We submit that the escalation dynamics – that is, the likelihood of dispute escalation at a given stage of the dispute settlement process – of strategic governments are different than if governments act 'naively' and fail to take into account their opponent's future responses. When forward-looking governments consider whether or not to request a panel ruling in a given dispute, they know that the domestic political stakes associated with the dispute will become more relevant if the conflict escalates beyond the consultation phase of the WTO dispute process. This is because with every additional stage of the DSM process, the dispute is likely to receive more attention from the domestic public. Once a WTO panel has been established and a ruling is issued, the case becomes more visible to a larger public and audience costs become more important as well.²⁵ Forward-looking governments know that a dispute with high political stakes will become more difficult to control and settle when it escalates to the panel stage.

²¹ Olson 1971.

²² Bernauer 2003; Bernauer and Meins 2003.

²³ Abbott and Snidal 1998; Moravcsik 2000; Simmons 2000.

²⁴ Busch and Reinhardt 2000; Reinhardt 2001.

²⁵ Busch and Reinhardt 2000.

Anticipating such loss of control, governments will think twice before asking for a panel in the first place, for it will become more and more difficult to settle a dispute with greater political stakes the further it has moved through the WTO dispute process. A forward-looking government will therefore try to solve potentially high-stakes disputes at the consultation phase, given the reduced likelihood of solving such disputes once they have passed beyond the first phase of the dispute settlement process. In other words, strategic, forward-looking behavior mediates the effect of domestic political stakes at the first stage by reducing the likelihood of escalation to the panel stage (but not at subsequent stages).²⁶ Two implications follow from these considerations. First, if observable indicators suggest that a dispute is associated with higher domestic political stakes, we should not see a substantially higher escalation risk from the consultation to the panel phase, compared to disputes in which observable indicators suggest lower domestic political stakes. Secondly, the escalation risk of such disputes should be substantially higher once they have reached the panel stage. This leads to the following hypothesis:

HYPOTHESIS 1: Assuming that governments are forward looking, domestic political stakes do not affect the probability that disputes will escalate to the panel stage; however, disputes with higher domestic political stakes are more likely to escalate further once they reach the panel stage than disputes with lower domestic political stakes.

THE WTO DSM AS A COMPLEXITY-REDUCING DEVICE

In contrast to the enforcement perspective, the managerial perspective views dispute settlement procedures primarily as information and rule clarification devices. International rules are sometimes ambiguous or so complicated that governments might violate them inadvertently because they do not fully understand whether their current policies and practices are in line with those rules. Hence according to the managerial perspective, dispute settlement is part of a compliance management process under conditions of incomplete contracts.²⁷ The basic modus operandi of dispute settlement mechanisms, according to this perspective, pertains to helping states implement their commitments by reducing complexity and providing information, rather than detecting and punishing violations of international rules. Correspondingly, the decisive characteristic of trade conflicts is dispute complexity, rather than domestic political stakes.

Complexity pertains to the content of a dispute and reflects the varying degrees of uncertainty about the legitimacy of disputed trade restrictions. Some discriminating trade instruments and policies may be obviously inconsistent with WTO rules, but the WTO conformity of other trade-inhibiting measures may be difficult to assess. Disputes over tariffs or quotas, for example, mostly deal with non-complex issues because countries usually understand whether or not a given tariff violates WTO rules. In contrast, disputes

²⁶ This should be true for the defendant and the complainant country alike. The complainant country should opt for early settlement during the consultation stage in order to obtain any concessions by the defendant country. This is because the complainant anticipates that the defendant country is unlikely to make concessions once a panel ruling has been issued due to the high domestic political stakes of the dispute. Similarly, the defendant country should opt for early settlement to avoid economic and/or reputational costs associated with an ongoing WTO dispute. This is because the defendant anticipates that it will become hard to appease the respective sector once a panel ruling has been issued.

²⁷ Chayes and Chayes 1993; Underdal 1998; Young 1994.

over environment, health and safety issues can be very complex, since it is often unclear whether environmental, health or safety concerns are sufficient to justify a trade barrier according to WTO rules.

Similar to high domestic political stakes, a high level of complexity is likely to hamper the prospects for co-operation in an international system without central authority, albeit in different ways. In the context of WTO dispute settlement, more complex disputes can be viewed as situations with less complete information. All parties incur transaction costs in such situations, either because WTO rules are vague or ambiguous or because the disputing parties hold different interpretations of which rules are applicable and how those rules should be interpreted.²⁸ Under such conditions, both the complainant and the defendant will find it difficult to establish whose interpretation of the relevant WTO rules is correct.

Dispute complexity increases the likelihood that countries will disagree about their specific rights and obligations under WTO law. If there is a high degree of uncertainty, governments will often have inconsistent expectations about winning a conflict.²⁹ When one or both sides in a dispute are overly optimistic about being right, early dispute settlement is difficult to achieve: the bargaining range (that is, the range of potential agreements that both parties should prefer to an escalation) decreases or even disappears when parties have conflicting and high expectations that they will win the case.³⁰

The dispute between the United States and Japan over measures restricting the import of apples (DS No. 245, a similar dispute has also occurred between Australia and New Zealand) is such an example of variation in dispute complexity. This dispute, which may appear rather trivial at first sight, is relatively complex. It concerns the rather complicated question of whether Japan's health and safety regulations for the importation of apples are consistent with the WTO Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement) and the Agreement on Agriculture. In this case, dispute settlement needs to clarify whether (from scientific, economic and other viewpoints) Japan's practices meet particular environmental and health objectives – in particular to avoid transmission of the fire blight bacterium. A settlement also needs to clarify how those objectives should be weighed against losses in the exporting country.

In comparison, the WTO dispute on cigarette taxes, which was initiated by Chile against Peru in 2001 (DS No. 227), is arguably less complex. Peru's obligation to end its discriminatory policies was quite obvious to most trade specialists. In contrast to the apples dispute, there was virtually no legal uncertainty about whether the defendant country's trade-restricting policies were compatible with the relevant WTO rules. The principal purpose of legal action by Chile against Peru was not to obtain clarification on how WTO rules should be interpreted in this case, but to increase pressure on Peru in order to obtain the desired concessions.

When WTO member countries face a dispute over a complex issue, the role of dispute settlement differs fundamentally from what the enforcement perspective assumes. As illustrated by the US-Japan apples case, WTO dispute settlement serves primarily to

²⁸ Coase 1960; Milgrom and Roberts 1992; Williamson 1985.

²⁹ Jervis 1976.

³⁰ Fearon 1995, 391–2. In particular, Fearon (1995) refers to uncertainty of winning a conflict arising from incomplete information about the opponent's strength. A similar logic applies to situations with other sources of uncertainty about the outcome of a conflict, such as uncertainty about the strength of a party's legal case.

reduce complexity and clarify rules.³¹ This second (managerial) perspective subscribes to key assumptions of the legalization literature in international relations, which views international courts as institutions that help countries reduce uncertainty and transaction costs and clarify legal obligations.³² This view, in turn, is based on the more general argument that international institutions help countries solve international co-operation problems by providing information, making rules explicit and pointing out deviations from officially recognized rules. WTO disputes, viewed from this perspective, escalate because international contracts are incomplete, there is considerable uncertainty over whether the disputed trade measure is legal and both parties believe that they are correctly interpreting the law.

The main instruments for establishing an authoritative interpretation of applicable trade law under conditions of uncertainty are the WTO's panel and appellate body.³³ According to the managerial perspective, settling more complex disputes during the consultation phase should be less likely. Such disputes are more likely to escalate to the panel or appellate body stage because uncertainty about whether the disputed trade measures are legal under WTO law motivates both parties to hold out, hoping to obtain a verdict that supports their position. Trade disputes over complex issues should, in most cases, be resolved once the respective panel and (if an appeal is made) the appellate body have issued their verdicts. The reason is that the DSM's verdict has, at that point, reduced the legal uncertainty. This argument reflects the common assumption in legal research that court trials would not be needed in a world with zero transaction costs and full information. According to this argument, absent transaction costs and uncertainty, the parties in a dispute should always come to a mutually agreed solution without going to court, since adjudication always entails direct and indirect costs.³⁴

Again, anticipation is likely to play a significant role in governments' decisions. Both sides in a conflict should anticipate that the conflict will end once the court has issued a verdict if the procedure is mostly about rule clarification. They should therefore be willing to escalate a dispute to the court level, because both can benefit from reduced transaction costs and the lower level of uncertainty that results from the adjudication process. Compared to the enforcement argument, anticipation in the managerial logic works in the opposite direction: it increases, rather than decreases, the escalation of a dispute to the panel stage. Once the panel resolves the issue, however, governments should try to restore compliance with WTO law. Since governments know that compliance will most likely arise after the panel ruling, they are (according to the managerial argument) more inclined to ask for a ruling. Hence from the managerial perspective, forward-looking behavior increases the already existing likelihood of dispute escalation to the panel stage under conditions of complexity, whereas the opposite is true according to the enforcement perspective.

The principal empirical implication of the complexity (managerial) perspective on dispute settlement is that disputes over more complex issues are more likely to escalate, but only up to the level of the panel or appellate body ruling. Escalation to that level (but not beyond) presumably clarifies the legal situation and helps countries restore compliance and solve the problem. Hypothesis 2 reflects these arguments.

HYPOTHESIS 2: Assuming that governments are forward looking, disputes over more complex trade issues are more likely to escalate to the panel stage than

³¹ Stone, Slantchev, and London 2008.

³² Keohane 1984; Smith 2000; Stone, Sweet, and Brunell 1998.

³³ Reinhardt 2001, 180.

³⁴ Coase 1960; Williamson 1985.

disputes over less complex issues. However, once disputes over more complex issues have reached the panel stage, they are unlikely to escalate further.

To summarize, we expect fundamentally different dispute resolution dynamics depending on whether the causal mechanism proposed by the enforcement or the managerial perspective dominates. According to the enforcement perspective, we should expect that a panel verdict is unlikely to resolve the conflict if a dispute's domestic political stakes are high. The key driver of dispute settlement dynamics from this perspective – domestic political stakes – initially reduces the probability of reaching the panel stage (assuming that governments are forward looking), but increases the probability of escalation after a panel ruling. In contrast, the managerial perspective implicates that dispute escalation to the panel stage can be an important step towards dispute resolution. The key driver of dispute settlement dynamics from this perspective – issue complexity – has an effect that clearly differs from what one would expect from an enforcement perspective: it increases the probability of dispute escalation to the panel stage, but reduces the probability of further escalation thereafter.

EMPIRICAL DESIGN

The theoretical arguments presented above take into account anticipatory (forward-looking) behavior of states during the bargaining process between the complainant and defendant in WTO dispute settlement. The assumption of anticipatory behavior has important implications for the empirical analysis: when actors are forward looking, their choices at the early stages of the bargaining process are influenced by their expectations about their opponent's future choices. Failure to account for anticipation may lead to biased estimates of the probability that an actor will make a particular choice at the outset and during the bargaining process.³⁵ We develop a strategic model that reflects the essential steps in the WTO DSM. Based on this model, we can deduce the likelihood that the complainant and the defendant will take specific actions during the WTO dispute settlement process. The empirical model is then estimated using statistical backwards induction.³⁶

The strategic empirical approach used for this analysis differs in important ways from the traditional statistical view on selection dynamics. Specifically, our strategic empirical approach assumes that selection occurs because governments are forward looking, which motivates them to escalate some disputes but not others. Taking into account such forward-looking behavior requires the analyst to start at the end of the decision tree, which is the last stage of the dispute settlement process, and work backwards towards the beginning of the game. The decision to study strategic versus non-strategic behavior then implies that we cannot use traditional statistical selection models, which start at the beginning of the decision tree and hence do not capture strategic and forward-looking behavior. The choice between the traditional and the strategic approaches ultimately depends on theoretical considerations that specify which underlying processes lead to selection. Given the importance of forward-looking behavior in our theoretical discussion, we opted for the strategic approach. The results discussed below will show how important it is to properly model governments' forward-looking behavior in our analysis.

³⁵ Signorino 1999.

³⁶ Bas, Signorino, and Walker 2008.

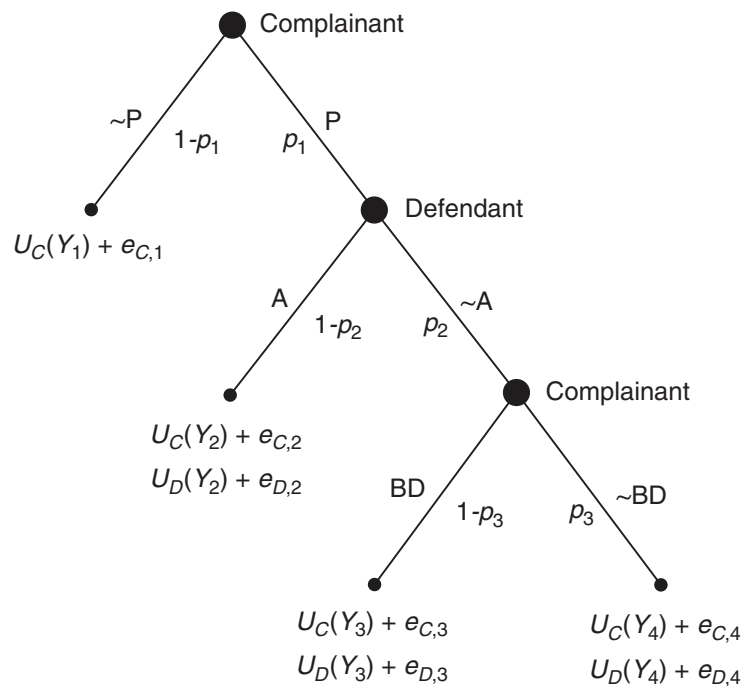


Fig. 1. The dispute bargaining game

STRATEGIC MODEL

The game shown in Figure 1 represents the most important stages of the WTO DSM, starting with the consultation phase. If the complainant is not satisfied with the consultation outcome, he can request a panel ruling (P). We define the WTO panel and appellate body process as a single stage in order to simplify the analysis.³⁷ The dispute ends at the consultation stage if no panel ruling is requested ($\sim P$). After the panel ruling, the defendant can accept the verdict (A) or not ($\sim A$). We assume that the game ends if the defendant implements the panel verdict, because there is no reason for the complainant to escalate the dispute further once it has reached its goal and the disputed trade barrier is removed. If the defendant refuses to implement the panel verdict, the complainant can decide to back down (BD) or escalate the conflict further; escalation leads to an implementation dispute ($\sim BD$).³⁸

³⁷ The WTO's DSM includes more steps, but these are mostly procedural and of minor importance for our analysis. For instance, the parties can appeal against the panel report, which leads to a review of the case by the appellate body, which then submits a report. To keep the model tractable, we do not include these additional steps, but consider the final decisions on each stage as the relevant outcomes. At the panel stage, this is the final report, whether the report is issued by a panel or the appellate body.

³⁸ In our representation of the WTO's DSM, we do not consider the panel as a third player, but treat its rulings as exogenous. Accounting explicitly for the panel decision (or even endogenizing its ruling) would make the model highly complex, with serious complications for the model estimation. Making the panel ruling endogenous would, in fact, imply that the panel behaves strategically and is therefore, in some way, politically motivated. For example, one could envisage the panel, taking into account power differences by ruling in favor of more powerful countries in disputes. At the same time, however, it is also possible to think of the panel trying to help countries with a lower legal capacity to counterbalance their potential disadvantage in the dispute process. Due to this theoretical ambiguity, and to keep the model as simple as possible, we decided to treat the panel as a technical and neutral body in the dispute settlement process.

The game has four outcomes: (1) Y_4 is observed if the dispute escalates to the implementation dispute stage; (2) Y_3 occurs if the defendant refuses to change its trade policy after a panel verdict, but the complainant decides not to pursue the issue further; (3) Y_2 is observed if a panel issues a verdict and the defendant complies with the ruling and implements the requested policy changes; and Y_1 occurs if, after consultations with the defendant, the complainant does not request a panel. $U_i(Y_j)$ is actor i 's utility of outcome j , and $e_{i,j}$ is the error assigned to this utility. The parameters p_1 through p_3 denote the probabilities that the respective move is realized.

We use quantal response equilibrium (QRE) to derive the outcome probabilities.³⁹ It follows from Figure 1 which outcomes will occur in equilibrium. The complainant will choose \sim BD if its utility of \sim BD is greater than the utility of BD, $U_C^*(Y_4) > U_C^*(Y_3)$ and vice versa, where $U_i^*(Y_j) = U_i(Y_j) + e_{i,j}$. The defendant will opt for A if its utility of A is greater than the expected utility of \sim A, $U_D^*(Y_2) > p_3 U_D^*(Y_4) + (1 - p_3) U_D^*(Y_3)$. In the same vein, the complainant will not demand a panel if its utility of \sim P is greater than the expected utility of P, $U_C^*(Y_1) > (1 - p_2) U_C^*(Y_2) + p_2 p_3 U_C^*(Y_4) + p_2 (1 - p_3) U_C^*(Y_3)$.

UTILITIES, VARIABLES AND DATA

We use several explanatory and control variables to measure governments' utilities. The political stakes reflect the audience costs that governments face at the domestic level. These audience costs should differ across economic sectors and in the extent to which various sectors can effectively organize to voice their concerns. Following the existing literature, we expect disputes involving the manufacturing sector to involve lower domestic political stakes than those related to agriculture, resource extraction and mining, and services.⁴⁰ The reasoning is that especially the agricultural sector has a particularly strong ability to organize politically, form alliances with other stakeholders, publicize disputes, and lobby for trade-restricting policies or practices. According to the seminal work by Olson, we should thus observe that due to this high level of organization, the agricultural sector should be able to politicize disputes and thus to pressure governments to take its side.⁴¹ While the importance of the issue at stake might often be modest in purely economic terms, the high political stakes associated with those disputes often makes it very difficult for the defendant country to accommodate the complainant's requests, as the EU-US dispute over GM crops illustrates.⁴² Hence we include a dummy variable that measures whether a dispute involves the manufacturing sector, assuming that manufacturing is associated with lower domestic political stakes.

To measure complexity, we use a dummy variable that captures whether a dispute refers to a complex trade-protective instrument – such as technical barriers to trade, anti-dumping, subsidies or various other instruments – in contrast to tariffs and quotas, which

³⁹ Signorino 1999. In particular, QRE is based on the assumption that a player may not entirely know his utility at a later decision node. This is plausible when the second move by the same player takes place in the more distant future (Bas, Signorino, and Walker 2008). This is generally the case in WTO dispute settlement, because it often takes quite some time to move from the consultation stage all the way to an implementation dispute. Moreover, the government can change within this period, which means that different actors take the decisions at different nodes. We leave the comparison of the QRE results with estimations based on Perfect Bayesian Equilibrium for future research.

⁴⁰ Bernauer 2003; Bernauer and Meins 2003; Davis and Shirato 2007; Elsig and Stucki 2012.

⁴¹ Olson 1971.

⁴² Bernauer 2003, Bernauer and Meins 2003.

TABLE 1 *Theoretical Concepts and Explanatory Variables*

Concept	Variable	Measurement
Political stakes	Stakes	Manufacturing sector involved (yes/no)
Complexity	Complex	Complex trade instrument used (yes/no)
Power	DifGDP _{ij}	Log[GDP(<i>i</i>)] – Log[GDP(<i>j</i>)]
International pressure	NumComp	Number of complainants
Legal strength	Strength	No of accepted claims/ no of claims made

we assume to be less complex. Coding tariff and quota cases as less complex follows a similar logic as the coding by Guzman and Simmons.⁴³ They argue that tariff and quota cases are easier to resolve since they refer to a continuous issue, in contrast to issues that have an ‘all or nothing’ character.

In additional analyses, we use more disaggregated measures that distinguish between more sectors and trade instruments. The results are largely similar. Since the number of disputes decreases with each additional stage in the settlement process, the results for the more disaggregated sector and instrument variables are based on few observations for each sub-sector or sub-instrument. This poses problems for statistical inference, and the use of more aggregated measures ensures that our conclusions are more reliable.

The empirical model also includes several control variables. The utilities of escalating the dispute to the implementation stage should depend on the retaliatory power of the two countries. Countries with less economic power will gain less from imposing countermeasures against a guilty country, and vice versa. Our measure of retaliatory power is the difference in economic size of the two countries.⁴⁴ We also add the number of complainants involved in a dispute, because a higher number of complainants may increase the pressure to comply with a WTO verdict due to higher international reputation costs. Finally, the defendant’s utility of compliance depends on the content of the ruling. To capture the extent to which the panel ruling strengthens the complainant’s or the defendant’s position, we include a variable that measures the ratio between the number of legal claims accepted by a panel and the total number of legal claims made by the complainant, as well as the square term of this ratio. The logic behind these two variables is that we expect compliance to be more easily achieved if only a few of the legal claims put forward by the complainant are upheld by the panel. Similarly, if the panel supports a higher proportion of the complainant’s legal claims, the potential reputation costs of non-compliance should be higher. When, however, only a portion of the legal claims made is accepted by the panel, the defendant country should have the strongest incentive not to comply. Table 1 summarizes how the explanatory variables relate to the theoretical concepts and our own theoretical discussion, and how they are measured.

To complete the model-building process, we need to specify which variable affects an actor’s utility at what stage in the WTO DSM. In other words, we need to assign the empirically observable measures of the utilities, $U_i(Y_1)$ through $U_i(Y_4)$, to estimate the probability of a particular outcome. Each of the utilities is a linear function of explanatory variables. Equations 1 through 7, define which variables are assigned to

⁴³ Guzman and Simmons 2002.

⁴⁴ Guzman and Simmons 2002; Sattler and Bernauer 2011.

which particular utility. This structure follows from our theoretical considerations and is largely consistent with previous formal analyses of conflict bargaining.⁴⁵

$$U_C(Y_1) = 0 \quad (1)$$

$$U_C(Y_2) = \beta_{C,4}\text{Complex} \quad (2)$$

$$U_D(Y_2) = \beta_{D,3}\text{Stakes} + \beta_{D,4}\text{Complex} + \beta_{D,4}\text{Strength} + \beta_{D,4}\text{Strength}^2 \quad (3)$$

$$U_C(Y_3) = \beta_{C,0} \quad (4)$$

$$U_D(Y_3) = \beta_{D,0} \quad (5)$$

$$U_C(Y_4) = \beta_{C,1}\text{DifGDP}_{CD} + \beta_{C,2}\text{NumComp} + \beta_{C,3}\text{Stakes} \quad (6)$$

$$U_D(Y_4) = \beta_{D,1}\text{DifGDP}_{DC} + \beta_{D,2}\text{NumComp} \quad (7)$$

We normalize the complainant's utility of outcome Y_1 to 0, as shown in Equation 1. Following the complexity argument, we assign the variable measuring dispute complexity to outcome Y_2 , as shown in Equations 2 and 3, because settling the dispute after a clarifying panel ruling should be valuable for both disputants. The variable measuring domestic political stakes (Manu) is added to the complainant's utility of Y_4 , as shown in Equation 6. If the complainant does not follow through with the complaint even though the defendant refuses to implement the panel verdict, the complainant incurs an audience cost for backing down. The audience cost should be smaller for disputes involving the manufacturing sector. The defendant incurs the same audience cost if it accepts an unfavorable panel verdict, as specified in Equation 3. Similarly, the content of the panel verdict influences the defendant's utility of compliance, and therefore the strength variable is added to the same equation. Finally, the number of complainants and amount of power affect the actors' utilities in an implementation dispute, as shown in Equations 6 and 7.

It is important to keep in mind that, due to the assumed forward-looking behavior of the two actors, explanatory variables that appear in the utility function at a later stage of the WTO DSM will, indirectly, also appear in the utility functions at an earlier stage. Using a strategic model therefore has important implications for understanding the effect of the different variables. When relying on regular regression models, the effect of each variable is denoted by its regression coefficient. In a strategic model, however, the effect of one variable has both direct and indirect components: direct as displayed by its regression coefficient and indirect through entering the probabilities that affect the earlier stages in the backwards induction model.⁴⁶ For example, the dummy variable for manufacturing that measures domestic political stakes should influence the utility of the complainant at the last stage $U_C(Y_4)$ and thus his decision whether to launch an implementation dispute. Due to the assumption of forward-looking behavior, the manufacturing variable will also

⁴⁵ For example, Fearon 1994.

⁴⁶ Signorino and Tarar 2006.

influence its decision at the first stage, when the actor decides whether to ask for a panel ruling.⁴⁷

The expected signs of the coefficients on the variables follow from the discussion above. A variable that has a negative effect on a country's utility from a specific outcome has a negative sign, and vice versa. These expectations generate the following implications for the probabilities of a particular move. As suggested by our theoretical arguments, and in line with the enforcement perspective, the probability of a panel ruling should increase if the dispute has greater domestic political stakes if we do not expect the government to be forward looking, but should decrease with greater domestic political stakes if we expect the government to be forward looking.

In both scenarios, however, disputes associated with greater domestic political stakes are more likely to continue and escalate further after the panel ruling. The probabilities that the defendant does not accept the ruling, p_2 , and that the complainant will not back down, p_3 , increase because the audience costs of these two choices are high for both parties. In line with the managerial perspective, complex cases should generally display a higher probability of escalating to a panel ruling. In contrast to disputes with greater domestic political stakes, the probability that a dispute is settled after a panel or appellate body ruling – that is, the probability of observing outcome Y_2 (compliance after a ruling) – should increase as the complexity of the disputed issue increases: both the probability that the defendant will accept the ruling, $1 - p_2$, and the probability that the complainant will request the establishment of a panel, p_1 , increase. The reason is that complexity positively affects both the complainant's and the defendant's utility of outcome Y_2 .

The data on trade disputes was collected from WTO documents. For each dispute, we obtained information on the defendant and complainants, and at which stage in the WTO DSM the dispute ended. Furthermore, we assigned a corresponding economic sector to each dispute and identified which trade policy instrument was used to protect this industry (for example, tariffs or subsidies). The first disputes entered the WTO system in 1995. We included all disputes for which a request for consultations was filed until 2006. Table 2 shows the frequency distribution of the dependent variable at the three dispute stages (Y_1 , Y_2 and Y_4). The table also lists, for each stage, how many disputes are, in our empirical definition, more complex cases and how many disputes involve higher domestic political stakes (the non-manufacturing sector). For example, the table shows that more complex cases end less often with a panel ruling (eighty cases) than with consultations (106 cases). Furthermore, more complex cases end more often with non-compliance (thirty-eight cases) and implementation disputes (twenty-four cases) than with compliance (thirty-seven cases) or backing-down by the complainant (fourteen cases).⁴⁸ Disputes with lower domestic

⁴⁷ Technically, the manufacturing sector variable does not explicitly appear in the complainant's utility for consultation $U_C(Y_1)$, but indirectly through the backwards induction process. Through backwards induction, all variables associated with the first stage are multiplied by the likelihood of facing an implementation dispute ($p_2 * p_3$), as specified at the end of the last section.

⁴⁸ To measure compliance, most existing studies on the implementation of GATT/WTO verdicts rely on the coding of Hudec (1993), which consists of three categories: whether the defendant liberalized the contested policy fully, partly or not at all. This measure does not capture whether the contested policy was changed as requested in a particular panel ruling, which is what we are interested in. To measure compliance, we first identified the cases in which a panel or the appellate body issued a ruling against the defendant. Following Hofmann and Kim (2009), we rely on the WTO's own assessment of whether or not a country complied with the ruling. Compliance, in our definition, occurs when the complainant country officially acknowledges that the defendant country's trade policies have been brought back into

TABLE 2 *Descriptive Statistics: Dispute Characteristics*

	Last stage		Second stage		First stage	
	Dispute	No dispute	Compliance	No compliance	Panel	No panel
Low-stakes disputes	10	8	34	18	53	47
Complex disputes	24	14	37	38	80	106
<i>All disputes</i>	32	32	77	64	152	179

TABLE 3 *Descriptive Statistics: Other Independent Variables*

	Mean	Std. Dev.	Min	Max
DifGDP _{CD}	-1.61	2.41	-9.22	4.61
Strength	0.69	0.22	0	1
NumComp	2.27	2.4	1	11

political stakes also end more often with panel rulings (fifty-three) than with consultations (forty-seven). Compliance occurs more often in these cases (thirty-four) than non-compliance (eighteen), and the number of implementation disputes is ten. Monadic and dyadic country-level trade data were taken from Barbieri, Keshk and Pollins and are log-transformed.⁴⁹ Table 3 presents the summary statistics.

RESULTS

The interpretation of the results from the strategic model is more complicated than in standard models, because explanatory variables can have a *direct* and an *indirect* effect on the complainant's and the defendant's utility. The direct effect refers to the immediate government utility at a particular stage of the dispute settlement. The indirect effect works through the impact of the same variable on the opponent's utility at a later stage of the dispute settlement process, and therefore through its impact on the probability of future escalation.⁵⁰ The indirect effect arises from the forward-looking behavior of governments, and the direct and indirect effects can reinforce or offset each other. For example, high political stakes (non-manufacturing) can have a positive direct effect on the complainant's utility of a panel ruling, because the complainant would like to increase pressure on the

(*F*'note continued)

compliance with WTO law during the specified time period. In eleven of the cases, we were not able to obtain information on compliance. We exclude these cases from our analysis at the second and last stages.

⁴⁹ Barbieri, Keshk, and Pollins 2008. In particular, we use data from the International Monetary Fund (IMF) Direction of Trade Statistics (DOTS) for Hong Kong, which is not included in Barbieri, Keshk, and Pollins 2008. Values for the EU are the sum of the values for the individual EU member states. For total EU exports and imports, we subtract intra-EU trade because, considering the EU as a single market and single actor in WTO disputes, we are interested in its dependence on trade with non-EU countries. Data on intra-EU trade is from the IMF DOTS.

⁵⁰ In Signorino and Tarar's (2006) words, 'in the deterrence model, relevant variables can affect the attacker's behavior in two ways: *directly* through its utilities for the various outcomes and *indirectly* through its belief p_d about whether the defender will defend' (595). In our context, the attacker is equivalent to the complainant, the defense is equivalent to non-compliance by the defendant and p_d is the probability that is called p_2 in our model.

TABLE 4 *Strategic Model of Dispute Resolution*

	Complainant: implementation	Defendant: non-compliance	Complainant: panel
Stakes	-0.631 (0.790)	-1.290** (0.658)	-6.305*** (2.245)
Complex		1.581** (0.680)	5.967*** (1.084)
DifGDP _{CD}	-0.361** (0.160)		-0.861*** (0.217)
DifGDP _{DC}		0.218 (0.239)	
NumComp	-0.347** (0.148)	0.783 (0.591)	1.125** (0.524)
Strength		6.125** (2.886)	
Strength ²		-6.409** (2.766)	
Constant	0.967 (0.651)	-3.895** (1.529)	-8.546*** (1.429)
Wald χ^2	7.75	10.80	42.41
Prob > χ^2	0.101	0.147	0.000
N	64	141	320

Standard errors in parentheses. Errors cluster on related disputes

p < 0.05, *p < 0.01

defendant when a high-stakes sector is concerned. But at the same time, it can also have a negative indirect effect if the complainant dislikes further escalation later on, and high political stakes increase the probability of escalation at the second and third stages of the DSM.

In view of these possibilities, we proceed in two steps. First, we discuss the estimated direct and indirect effects of the explanatory variables separately. These results provide an initial indication of whether state behavior in the WTO DSM is consistent with the mechanisms proposed by the two arguments. However, these results do not tell us exactly how these mechanisms jointly affect the escalation dynamics of WTO disputes, because direct and indirect effects can offset each other (see above). In a second step, we therefore add an analysis of the predicted probabilities of each of the possible outcomes of the dispute settlement procedure. These probabilities reflect the (*ex ante*) likelihood that no panel ruling (outcome Y_1 in Figure 1), compliance after a ruling (outcome Y_2), no implementation dispute after non-compliance (outcome Y_3) or an implementation dispute (outcome Y_4) will occur, and capture how these probabilities vary for different values of the political stakes and complexity variables. These probabilities are useful because they reflect both the direct and indirect effects of domestic political stakes and complexity.

Table 4 presents the estimation results for our strategic WTO dispute settlement model. The first column displays the results for the last stage, which represents the complainant's decision whether to launch an implementation dispute. The results for the defendant's decision to comply with a panel ruling are shown in the second column. The third column shows the estimates for the complainant's decision whether to request a panel.⁵¹

⁵¹ The statistical significance of the joint hypothesis tests is lower for the stage regressions in the first and second columns because the number of observations decreases with every additional stage in the DSM. To ensure consistency between the structural empirical model and theory, we prefer not to exclude

The coefficients on the explanatory variables indicate their *direct* effect on actors' utilities. As we can infer from the last column, the variable measuring lower domestic political stakes has a negative direct effect on the complainant's utility from requesting a panel ruling. Similarly, the second column shows that lower political stakes decrease the defendant's utility of non-compliance. The lower political stakes variable does not have a statistically significant effect on the complainant's utility of an implementation dispute, as the results in the first column show. The estimated *direct* effect therefore shows that disputes with lower political stakes are less likely to escalate, since they are more likely to end after consultation in the first stage and more likely to end in compliance after a panel ruling in the second stage. This is partially consistent with the implications of the enforcement perspective, as summarized in Hypothesis 1. The results also show that dispute complexity increases the complainant's direct utility of a panel ruling (last column) as well as the defendant's direct utility of non-compliance. The first result is consistent with the complexity argument and the associated Hypothesis 2. However the second is not consistent, because the complexity argument, in line with the managerial perspective, suggests that the defendant should accept the clarifying ruling, implying that its utility of non-compliance should decrease.

The *indirect* effect of the variables is reflected by the coefficients on the constant in each equation. These coefficients indicate how an actor's utility changes when the probability of dispute escalation at the subsequent stages of the WTO DSM increases. Specifically, a negative coefficient means that an actor's utility of reaching the next escalation stage decreases when the probability of further escalation beyond that stage increases.⁵² The results in the last column therefore suggest that the complainant's utility of a panel ruling decreases when the probability of non-compliance, and of an implementation dispute, increases. Since disputes concerning sectors with higher political stakes are more likely to end in non-compliance after a panel has ruled in the case (as indicated by the negative coefficients on manufacturing in the second column), the involvement of a sector with higher political stakes indirectly decreases the utility of a panel ruling at the first stage of the DSM. This result is consistent with Hypothesis 1, suggesting that complainants anticipate the greater escalation risk for higher stakes disputes and strategically select only those that are very important to them.

While the indirect effects are consistent with forward-looking behavior, as proposed by the domestic political stakes argument, they are not consistent with the complexity argument. The estimation results indicate that greater complexity makes non-compliance by the defendant more likely. The negative coefficient on the constant shows that greater complexity indirectly reduces the utility of a panel ruling for the complainant. The complexity argument

(*F*'note continued)

variables from a particular stage regression simply because they are not statistically significant and therefore reduce the significance of the joint hypothesis tests. What matters for our analysis is the implications of all three stage regressions for dispute resolution dynamics as represented by the predicted probabilities in Figures 2 and 3.

⁵² Following the procedure by Bas, Signorino, and Walker (2008, 28), the constant in the equation representing the complainant's utility of a panel ruling (last column) is weighted by the product of the probability of non-compliance and the probability of an implementation dispute, $p_2 * p_3$. A negative coefficient on the constant in the last column therefore indicates that the utility of a panel ruling decreases when the probability of non-compliance (p_2) and/or the probability of an implementation dispute (p_3) increases. The constant in the equation representing the defendant's utility of non-compliance is weighted by the probability of an implementation dispute, p_3 . A negative coefficient on this variable therefore indicates that the utility of non-compliance decreases when the probability of an implementation dispute increases.

suggests the opposite (that is, that greater complexity should increase the complainant's utility of a ruling). This should be the case because, in line with the managerial perspective, governments should comply with their international obligations once legal ambiguities have been clarified. In summary, the estimation results are consistent with the forward-looking domestic political stakes argument and the enforcement perspective, but they offer only limited support for the complexity argument and thus the managerial perspective.

As discussed above, the direct and indirect effects of the domestic political stakes variable work in opposite directions at the first stage. The question, then, is which effect dominates – and what this means for the particular outcomes in the DSM process. To assess this, we compute the predicted (*ex ante*) probabilities for each possible outcome, $\Pr(Y_1)$ through $\Pr(Y_4)$. These probabilities are calculated as follows:

$$\Pr(Y_1) = 1 - p_1 \quad (8)$$

$$\Pr(Y_2) = p_1 \times (1 - p_2) \quad (9)$$

$$\Pr(Y_3) = p_1 \times p_2 \times (1 - p_3) \quad (10)$$

$$\Pr(Y_4) = p_1 \times p_2 \times p_3 \quad (11)$$

where p_1 through p_3 are the probabilities in Figure 1.⁵³ We compute these probabilities for disputes that are characterized by lower and higher domestic political stakes (that is, those concerning the non-manufacturing and the manufacturing sectors), $\Pr^{ls}(i)$ and $\Pr^{hs}(i)$, and take the difference of these two probabilities, $\Pr^{ls}(i) - \Pr^{hs}(i)$. This difference shows how the probability of outcome i differs across the manufacturing and non-manufacturing sectors. Hence we can examine how the escalation risk differs for sectors with lower and higher political stakes by analyzing the differences in probabilities for all outcomes Y_1 through Y_4 .

The differences in probabilities for strategic (forward-looking) governments are plotted in the top left quadrant of Figure 2(a). The figure shows that the probability that no ruling is requested – or that the dispute ends at the consultation stage (and Y_1 occurs) – is smaller for manufacturing than for non-manufacturing. This implies that the indirect effect of domestic political stakes dominates: forward-looking governments are hesitant to request a ruling when they expect that the domestic political stakes associated with a conflict increase the escalation risk after the ruling. The difference in escalation probabilities at the consultation stage is substantially and statistically significant. The estimated difference for Y_2 suggests that compliance is significantly more likely for low- than for high-stakes cases. This confirms that cases characterized by lower domestic political stakes are more likely to end after a ruling and are thus less escalation prone once they have reached the panel stage. The *ex ante* probabilities for Y_3 and Y_4 do not differ much.

For further analysis, we compare these results with a simulation of how a 'naive' government (one that is not forward looking) would behave. This simulation has two purposes. First, it allows us to assess the importance of forward-looking behavior. If the forward-looking component is small, we should see only small differences between the

⁵³ These probabilities can be estimated using the estimation results in Table 4 for different values of the independent variables, for example for low political stakes (manufacturing) = 1 and low political stakes = 0 and for complexity = 0 and complexity = 1. All other variables are set to their median values.

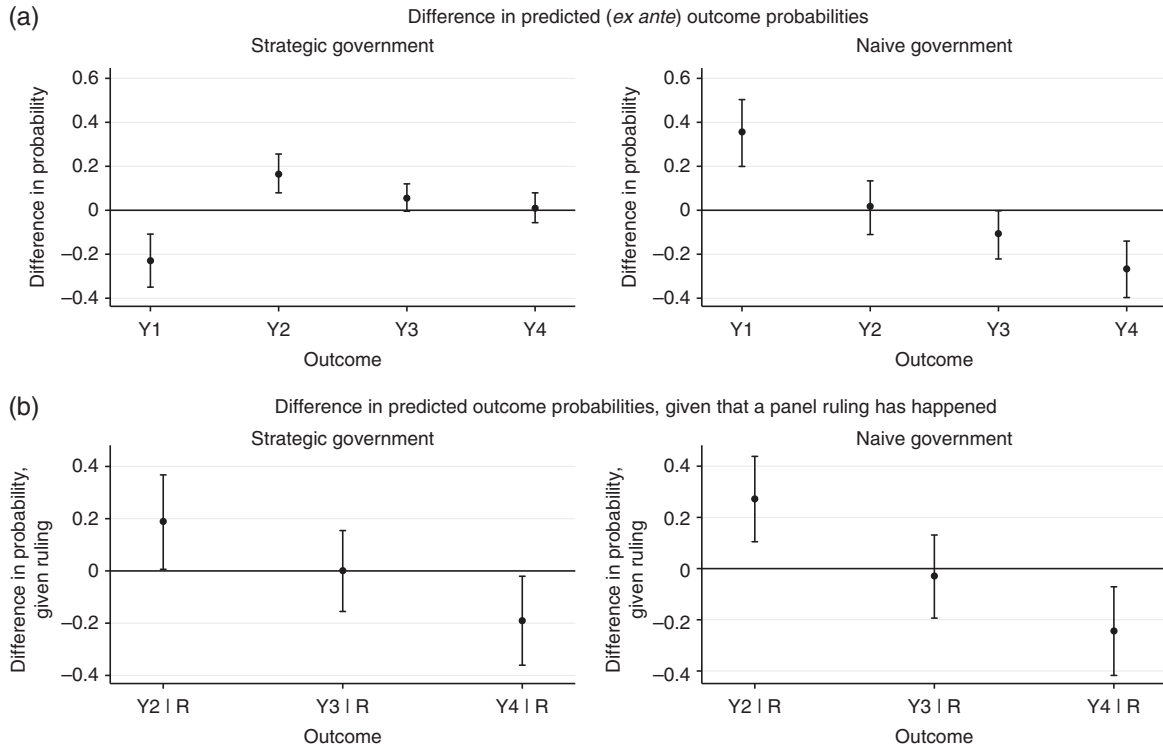


Fig. 2. Difference in predicted probabilities for **lower political stakes** and **higher political stakes** disputes that a particular outcome occurs

Note: the 'strategic' government is forward looking. The 'naive' government only considers the current stage of the WTO DSM. Bootstrapped confidence intervals are based on 10,000 draws.

probabilities for strategic and naive governments. Secondly, the simulation allows us to isolate and display the direct effect of the variables on the immediate utilities at a particular stage in the dispute settlement process. To simulate the behavior of a naive government, we treat the escalation probabilities at subsequent stages as exogenous and fix them at 0.5. This reflects the idea that a naive government is unable to anticipate the behavior of its opponent and simply assumes that there is a 50 per cent chance that the conflict will end at the next stage.⁵⁴

The results for the naive government are shown in the top right quadrant of Figure 2(a). For the domestic political stakes variable, they are substantially different from those for the forward-looking government, shown on the top left. Ignoring the indirect effect of forward-looking behavior, the predicted probability that a dispute will end with consultation – and thus no panel ruling (Y_1) – is significantly larger for cases with lower political stakes. This is consistent with the discussion of the estimation results above, and confirms that the direct effect of lower political stakes on the complainant's utility of a ruling is negative. The effect changes direction for implementation disputes (Y_4): the predicted probability of an implementation dispute is smaller for sectors with lower political stakes than for sectors with higher political stakes.

Finally, Hypothesis 1 implies that disputes with higher political stakes should escalate particularly often, given that they have experienced a panel ruling. To examine this claim,

⁵⁴ Note that p_2 depends on p_3 , and p_1 depends on both p_2 and p_3 . We therefore set $p_3 = 0.5$ when computing p_2 and $p_2 = p_3 = 0.5$ when computing p_1 . Setting the future escalation probabilities to other exogenous values does not change the main result here.

we compute and plot the conditional probability for compliance (Y_2), no implementation dispute after non-compliance (Y_3) and an implementation dispute (Y_4), given that a panel has ruled in the case. The results for both the forward-looking and the naive government are shown in the second row of Figure 2. These results confirm that lower political stakes cases are considerably less likely than higher political stakes cases to escalate: compliance after a ruling is much more likely, and an implementation dispute is much less likely, in cases involving sectors with lower political stakes. Moreover, there is little difference between the forward-looking and naive governments. Forward-looking behavior is therefore particularly important for the complainant's decision to request a panel ruling or not, but not afterwards. Governments that take a conflict with greater political stakes to the panel stage tend to be willing to continue the dispute if the defendant does not comply with the ruling.

Overall, these results lend strong support to the theoretical arguments underpinning Hypothesis 1. If we conceive of governments as strategic actors that do not simply think step by step but anticipate the potential reactions of their opponents, we obtain a more nuanced understanding of how dispute settlement works in the WTO. Anticipating that it is difficult to control a dispute that involves a sector with greater political stakes, governments tend to avoid letting such disputes escalate in the first place.

Turning to Hypothesis 2, which reflects the complexity or rule clarification argument of the managerial perspective, we observe that anticipating behavior does not play as important a role as in the case of greater domestic political stakes. As displayed in Figure 3(a), both strategic and naive governments are more likely to ask for a panel ruling in the case of more complex disputes, as predicted by Hypothesis 2. But contrary to Hypothesis 2, disputes over more complex trade-inhibiting instruments are not more likely to end after a panel ruling, as both strategic and naive governments are more likely to accept the escalation of complex disputes to the implementation dispute stage. If we look at the two pictures in Figure 3(b), we see that the conditional probability of observing compliance after a panel ruling has occurred ($Y_2|R$) is lower for more complex than for less complex disputes. Hence both strategic and naive governments are significantly less likely to comply with a ruling that was issued in a more complex dispute. In contrast, governments are more likely to go all the way through the DSM – and are even willing to incur the costs of an implementation dispute – if they clash over more complex trade barriers.

In summary, these findings suggest that more complex disputes are often not settled with a panel ruling, but disagreement between the defendant and the complainant continues after the panel has supposedly 'clarified' the case. These findings are not in line with the managerial perspective. Instead, they support an argument by Busch and Reinhardt⁵⁵ and Reinhardt,⁵⁶ who posit that a GATT/WTO panel ruling does not reveal much information or reduce complexity. Panels, according to this argument, mostly have an indirect effect because they make early settlement of trade disputes more likely.

The results concerning the control variables also warrant some discussion. We find that economic power reduces the probability of escalation within the WTO, which suggests that countries are more likely to settle their conflicts during consultations and after panel or appellate body rulings when power asymmetries are large. This means that more powerful complainants need less legal support from the WTO to extract concessions from

⁵⁵ Busch and Reinhardt 2000.

⁵⁶ Reinhardt 2001.

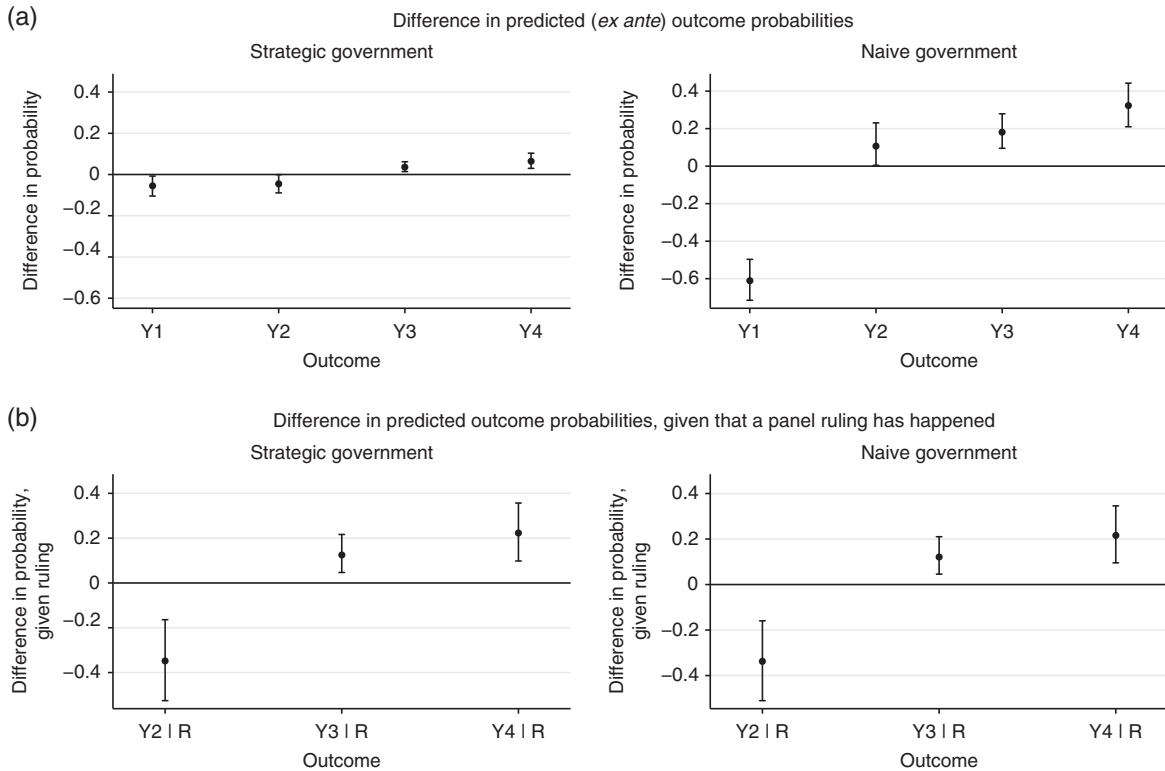


Fig. 3. Difference in predicted probabilities for **complex** and **non-complex** disputes that a particular outcome occurs

Note: the 'strategic' government is forward looking. The 'naive' government only considers the current stage of the WTO DSM. Bootstrapped confidence intervals are based on 10,000 draws.

the defendant or enforce a ruling than less powerful complainants.⁵⁷ The results also show that a larger number of complainants makes a panel or appellate body ruling more likely, but that once a ruling has been issued, more complainants reduce the risk of further escalation. This means that more complainants increase pressure on the defendant to comply with the ruling, which decreases the likelihood that the case will escalate into a compliance dispute. Finally, the legal strength of the case, measured as the proportion of claims accepted by the panel, has the hypothesized inverted U-shaped effect: if the panel accepts either a low or a high number of claims made by the complainant country, the likelihood of compliance by the defendant increases. This confirms our expectation that defendants have the greatest incentive to shirk when the panel accepts some, but not all, of the claims.

In summary, our findings are mostly in line with the predictions of the domestic political stakes argument and thus support the enforcement perspective on compliance. A dispute's domestic political stakes become increasingly important as countries escalate a trade conflict to higher levels of the WTO dispute settlement process. As the enforcement perspective implies, a panel verdict does not resolve the conflict if the domestic political stakes of a dispute are high. Domestic political stakes, therefore, reduce the probability of reaching the panel stage when governments are forward looking, but increase the probability of escalation after a panel ruling. In contrast to the managerial perspective, we

⁵⁷ Bown 2004a; Bown 2004b; Busch and Reinhardt 2000; Reinhardt 2001.

do not find evidence for the claim that a panel ruling reduces complexity; in fact, more complex disputes tend to escalate further even after a panel ruling.

CONCLUSION

Two perspectives on the functions of international judicial mechanisms have left a strong imprint on theoretical and empirical work on international dispute settlement and international co-operation more generally. The enforcement perspective regards dispute settlement as an enforcement instrument in settings characterized by incentives to violate international law, domestic interest group pressure and international power politics. The less antagonistic managerial perspective views dispute settlement as a managerial enterprise that reduces complexity and clarifies rules in settings in which international contracts are incomplete and states require help in co-ordinating and implementing their international obligations.

We derived two distinct empirical implications from these perspectives for dispute settlement in the WTO and used a novel empirical approach to test them on WTO disputes from 1995 to 2006. This approach combines a three-step coding of dispute escalation with a strategic bargaining model and statistical backwards induction to account for governments' forward-looking behavior. Our results offer more support for the enforcement than the managerial perspective. Disputes in which more powerful domestic interests are at stake and audience costs are higher are more prone to escalation throughout the WTO dispute settlement process, and the risk of escalation increases during the key stages of this process. However, governments that anticipate the increased risk of escalation try to pre-empt this loss of control by settling disputes associated with higher domestic political stakes early in the consultation stage. These results strongly support the conjecture that governments are forward looking, for they base their decisions at an early stage of the dispute settlement process on their expectations about how their opponents will behave at a later stage.

Interestingly, this forward-looking behavior plays a less significant role in disputes over more complex trade-restricting policy instruments. More complex cases are more likely to escalate through the different stages of the dispute settlement process. This finding only partially supports the managerial perspective on dispute settlement. The empirical implication of this perspective is that we should primarily see escalation up to the panel ruling stage, but not beyond. We observe the former, but not the latter. One reason for this finding could be that governments' political motivations dominate the complexity logic. Specifically, governments might choose increasingly complex trade barriers to protect politically sensitive economic sectors, hoping that these complex barriers make it difficult for the opposing government to demonstrate that these measures are illegitimate.

Our findings add to the existing literature on WTO dispute settlement in several ways. Whereas most of this literature uses a binary definition of dispute escalation – whether a dispute reaches the stage of a WTO panel ruling or not – we conceptualize and empirically analyze dispute settlement using a three-step process. The novel combination of conceptualizing and coding dispute escalation as a three-step process, and using a strategic model and statistical backwards induction to analyze dispute escalation dynamics, allows us to take into account that governments act strategically when using the WTO DSM. The results concerning how governments deal with disputes associated with higher domestic political stakes show that considering governments' forward-looking strategic behavior is important. Another contribution to the existing literature is

that the mixed evidence for the complexity hypothesis addresses an argument by Busch and Reinhardt⁵⁸ and Reinhardt,⁵⁹ who claim that a GATT/WTO panel ruling does not reveal much information or reduce complexity. Our results offer support for this claim.

Future research should also examine the role of learning and signaling in WTO dispute settlement. In our empirical model, governments do not learn from actions that were previously taken during the conflict. Nor do they learn about their opponents' capabilities or resolve to escalate the conflict further. Empirical analysis based on Perfect Bayesian Equilibrium and its empirical extensions could yield insights into interactions of this kind.⁶⁰ Finally, it could be useful to consider the panel as a separate actor that has its own interests and behaves strategically, rather than treating it as a technocratic institution that objectively interprets WTO law.

REFERENCES

- Abbott, Kenneth W., and Duncan Snidal. 1998. Why States Act Through Formal International Organizations. *The Journal of Conflict Resolution* 42:3–32.
- Allee, Todd. 2010a. The Distinctiveness of Developing Country Trade Dispute Initiation. Manuscript, University of Illinois.
- . 2010b. Legal Incentives and Domestic Rewards: A Litigation Model of GATT/WTO Dispute Resolution. Manuscript, University of Illinois.
- Allee, Todd, and Paul Huth. 2006. Legitimizing Dispute Settlement: International Legal Rulings as Domestic Political Cover. *American Political Science Review* 100:219–34.
- Barbieri, Katherine, Omar Keshk, and Brian Pollins. 2008. Correlates of War Project Trade Dataset Codebook, Version 2.01. <http://correlatesofwar.org>, accessed 19 July 2010.
- Barrett, Scott. 2006. *Environment and Statecraft: The Strategy of Environmental Treaty-Making*. Oxford: Oxford University Press.
- Bas, Muhammet A., Curis S. Signorino, and Robert W. Walker. 2008. Statistical Backwards Induction: A Simple Method for Estimating Recursive Strategic Models. *Political Analysis* 16:21–40.
- Bechtel, Michael, and Thomas Sattler. Forthcoming. What is Litigation in the WTO Worth? *International Organization*.
- Bernauer, Thomas. 2003. *Genes, Trade and Regulation: The Seeds of Conflict in Food Biotechnology*. Princeton, NJ: Princeton University Press.
- Bernauer, Thomas, Manfred Elsig, and Joost Pauwelyn. 2012. The Dispute Settlement Mechanism: Analysis and Problems. In *The Oxford Handbook on the World Trade Organization*, edited by Amrita Narlikar, Martin Daunton, and Robert M. Stern, 485–506. Oxford: Oxford University Press.
- Bernauer, Thomas, and Erika Meins. 2003. Technological Revolution Meets Policy and the Market: Explaining Cross-National Differences in Agricultural Biotechnology Regulation. *European Journal of Political Research* 42:643–83.
- Bown, Chad P. 2004a. Developing Countries as Plaintiffs and Defendants in GATT/WTO Trade Disputes. *The World Economy* 27:59–80.
- . 2004b. On the Economic Success of GATT/WTO Dispute Settlement. *Review of Economics and Statistics* 86:811–23.
- Busch, Marc L., and Eric Reinhardt. 2000. Bargaining in the Shadow of the Law: Early Settlement in GATT/WTO Disputes. *Fordham International Law Journal* 24:158–72.

⁵⁸ Busch and Reinhardt 2000.

⁵⁹ Reinhardt 2001.

⁶⁰ Bas, Signorino, and Walker 2008; Lewis and Schultz 2003; Whang 2010.

- Chayes, Abram, and Antonia H. Chayes. 1993. On Compliance. *International Organization* 47:175–205.
- Coase, Ronald. 1960. The Problem of Social Cost. *Journal of Law and Economics* 3:1–44.
- Davis, Christina L. 2008. The Effectiveness of WTO Dispute Settlement: An Evaluation of Negotiation Versus Adjudication Strategies. Paper Presented at the Annual Meeting of the American Political Science Association, Boston, MA 28–31 August.
- Davis, Christina L., and Sarah B. Bermeo. 2009. Who Files? Developing Country Participation in GATT/WTO Adjudication. *Journal of Politics* 71:1033–49.
- Davis, Christina L., and Yuki Shirato. 2007. Firms, Governments, and WTO Adjudication: Japan's Selection of WTO Disputes. *World Politics* 59:274–313.
- Downs, George, David M. Rocke, and Peter Barsboom. 1996. Is the Good News About Compliance Good News for Cooperation? *International Organization* 53:379–406.
- Elsig, Manfred, and Philipp Stucki. 2012. Low-Income Developing Countries and WTO Litigation: Why Wake Up the Sleeping Dog? *Review of International Political Economy* 19:292–316.
- Fearon, James D. 1994. Signaling Versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model. *Journal of Conflict Resolution* 38:236–69.
- . 1995. Rationalist Explanations for War. *International Organization* 49:379–414.
- Francois, Joseph, Henrik Horn, and Niklas Kaunitz. 2008. Trading Profiles and Developing Country Participation in the WTO Dispute Settlement System. International Centre for Trade and Sustainable Development Issue Paper No. 6.
- Goldstein, Judith O., Miles Kahler, Robert O. Keohane, and Anne-Marie Slaughter. 2000. Introduction: Legalization and World Politics. *International Organization* 54:385–99.
- Guzman, Andrew. 2002. The Cost of Credibility: Explaining Resistance to Interstate Dispute Resolution Mechanisms. *Journal of Legal Studies* 31:303–26.
- Guzman, Andrew, and Beth A. Simmons. 2002. To Settle or Empanel? An Empirical Analysis of Litigation and Settlement at the World Trade Organisation. *Journal of Legal Studies* 31:205–35.
- Hofmann, Tobias, and Soo Yeon Kim. 2009. The Political Economy of Compliance in WTO Disputes. Paper Presented at the International Political Economy Society Conference, College Station, TX, 13–14 November.
- Hudec, Robert E. 1993. *Enforcing International Trade Law: The Evolution of the Modern GATT Legal System*. Salem, NH: Butterworth Legal Publishers.
- Jervis, Robert. 1976. *Perception and Misperception in International Relations*. Princeton: Princeton University Press.
- Keohane, Robert. 1984. *After Hegemony: Cooperation, and Discord in the World Political Economy*. Princeton: Princeton University Press.
- Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. The Rational Design of International Institutions. *International Organization* 55:761–99.
- Lewis, Jeffrey B., and Kenneth A. Schultz. 2003. Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information. *Political Analysis* 11:345–67.
- Milgrom, Paul, and John Roberts. 1992. *Economics, Organization, and Management*. Englewood Cliffs: Prentice Hall.
- Mitchell, Ronald B. 2006. Problem Structure, Institutional Design, and the Relative Effectiveness of International Environmental Agreements. *Global Environmental Politics* 6:72–89.
- Moravcsik, Andrew. 2000. The Origins of Human Rights Regimes: Democratic Delegation in Postwar Europe. *International Organization* 54:217–52.
- Olson, Mancur. 1971. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.
- Posner, Eric A., and John C. Yoo. 2005. Judicial Independence in International Tribunals. *California Law Review* 93:3–74.
- Reinhardt, Eric. 2001. Adjudication Without Enforcement in GATT Disputes. *Journal of Conflict Resolution* 45:174–95.

- Rosendorff, B. Peter. 2005. Stability and Rigidity: Politics and Design of WTO's Dispute Settlement Procedure. *American Political Science Review* 99:389–400.
- Rosendorff, B. Peter, and Helen Milner. 2001. The Optimal Design of International Trade Institutions: Uncertainty and Escape. *International Organization* 55:829–57.
- Sattler, Thomas, and Thomas Bernauer. 2011. Gravitation or Discrimination? Determinants of Litigation in the World Trade Organization. *European Journal of Political Research* 50 (2):143–67.
- Signorino, Curtis S. 1999. Strategic Interaction and the Statistical Analysis of International Conflict. *American Political Science Review* 93:279–98.
- Signorino, Curtis S., and Ahmer Tarar. 2006. A Unified Theory and Test of Extended Immediate Deterrence. *American Journal of Political Science* 50:586–605.
- Simmons, Beth A. 2000. International Law and State Behavior: Commitment and Compliance in International Monetary Affairs. *American Political Science Review* 94:819–35.
- Smith, James M. 2000. The Politics of Dispute Settlement Design: Explaining Legalism in Regional Trade Pacts. *International Organization* 54 (1):137–80.
- Stone, Randall W., Branislav L. Slantchev, and Tamar R. London. 2008. Choosing How to Cooperate: A Repeated Public-goods Model of International Relations. *International Studies Quarterly* 52 (2):335–62.
- Stone Sweet, Alec, and Thomas Brunell. 1998. Constructing a Supranational Constitution: Dispute Resolution and Governance in the European Community. *American Political Science Review* 92:63–81.
- Underdal, Arild. 1998. Explaining Compliance and Defection: Three Models. *European Journal of International Relations* 4:5–30.
- Whang, Taehee. 2010. Empirical Implications of Signaling Models: Estimation of Belief Updating in International Crisis Bargaining. *Political Analysis* 18:381–402.
- Williamson, Oliver. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.
- Young, Oran R. 1994. *International Governance: Protecting the Environment in a Stateless Society*. Ithaca, NY: Cornell University Press.