

Between the Letter of the Law and the Demands of Politics: Judicial Balancing of Trade Authority within the WTO

Cosette D. Creamer*

Abstract

The judicial bodies of the World Trade Organization (WTO) operate virtually free from direct government influence. Yet they are also politically savvy and at times conform their rulings to the preferences of member states. This article employs original measures of judgment outcomes and aggregate levels of government support for the WTO's dispute settlement bodies to identify the conditions under which they do so. It finds that dispute panels tend to signal less deference to governments' regulatory choices—asserting international authority over a vast range of policy areas—only when they enjoy relatively greater support among the membership as a whole. Yet panels are not purely political actors, as they are accountable to a higher judicial authority. This article finds further that the legal constraints imposed by precedent moderate the influence of political pressures on panel validation of trade measures. In this way, panels simultaneously seek to maximize their support among their legal and their political audiences, thereby balancing greater legalism and predictability within the trade regime against continued flexibility for governments.

*Visiting Assistant Professor, Boston University School of Law; Ph.D. candidate, Department of Government, Harvard University; J.D., Harvard Law School. Email: creamer@fas.harvard.edu. For helpful comments and useful discussions, I would like to thank Beth Simmons, Dustin Tingley, Mark Wu, Gerald Neuman, Rachel Brewster, Sara B. Mitchell, Karen Alter, Mikael Rask Madsen, Anton Strezhnev, Joan Cho, Emily Clough, George Yin, Zuzanna Godzimirska, Todd Tucker, and participants in the University of Copenhagen's iCourts Centre research workshop and the American Society of International Law's Mid-Year Meeting Research Forum. Elise Barankouski provided invaluable research assistance.

1. INTRODUCTION

The turn toward law within world politics is undisputed. Over the past century, governments have increasingly relied on legal agreements to facilitate cooperation, with over five hundred major multilateral treaties and hundreds of thousands of international agreements currently registered with the United Nations.¹ In step with an increased reliance on legal rules, the spectacular growth in the number of transnational judicial and quasi-judicial bodies has sparked considerable debate over why states would create courts that could interfere with critical elements of national governance.² Policymakers and commentators frequently warn that these tribunals have the potential to expand their authority over time, engage in judicial lawmaking, or impose greater restrictions or sovereignty costs on states (Bradley and Kelley 2008; Hathaway 2008; Keohane, Moravcsik and Slaughter 2000). In addition to posing such potential risks, however, international courts do clarify obligations, help states overcome cooperation problems, and contribute more generally to the development of an international legal system (Alter 2014).

The transition from the General Agreement on Tariffs and Trade (GATT) dispute panels to the World Trade Organization's (WTO) Dispute Settlement Mechanism (DSM) represented a notable instance of the turn toward courts within world politics in that it considerably strengthened the independence of the bodies delegated authority to resolve trade disputes (Jackson 2001; Raab and Bevers 2006). The WTO Appellate Body and dispute panels (which together comprise the DSM) possess compulsory jurisdiction and act as the authoritative interpreters of WTO law, with their decisions effectively impossible for members to reverse. This has led many to assert that they operate virtually free from direct state control in enforcing WTO rules against governments (Alter 2008; Cass 2001; Stone Sweet 1997; Stone Sweet and Brunell 2013: 62). Yet as a number of studies on international courts have stressed, these institutions often purposefully conform their rulings to the expressed preferences of member states, particularly politically or economically influential ones, and do not always engage in expansive judicial lawmaking or assertions of authority (Busch and Pelc 2010; Garrett, Kelemen and Schulz 1998; Helfer and Alter

¹See the *United Nations Treaty Collection*, available at: <https://treaties.un.org/>.

²By last count, there are more than twenty permanent international courts and hundreds of quasi-judicial and non-permanent judicial bodies (Alter 2011; Romano 2011). See *The Project on International Courts and Tribunals*, available at: <http://www.pict-pcti.org/>.

2013; Kelemen 2001; Smith 2003; Steinberg 2004).

This article evaluates these seemingly contrary expectations in the context of international economic law by identifying the conditions under which WTO dispute settlement panels seek to expand or restrict their delegated authority. Similar to recent empirical efforts to unpack and explain international judicial behavior, this article focuses on panels' substantive review of domestic policy choices (Busch and Pelc 2010; Cole 2011; Danner and Voeten 2010; Lupu and Voeten 2012; Pelc 2014; Voeten 2008, 2013). The extent to which panels validate such choices effectively allocates decision-making authority between national actors on the one hand and international institutions on the other, and defines the parameters of the delegation relationship between states and the WTO (Oesch 2003; von Staden 2012: 1033). I argue that WTO members have relied primarily on diffuse rhetorical pressures to signal support for or criticism of the Organization's exercise of judicial authority, through regular contestation over dispute rulings and their impact on policy areas over which they continue to assert exclusive authority. Panels are sensitive, and under certain conditions responsive, to these collective pressures in order to ensure the continued political support of the membership as a whole.

To evaluate this argument, I examine how panels respond—through their substantive review of domestic laws and policies—to views expressed by member states on the WTO's exercise of dispute settlement authority. Employing methods of automated content analysis to construct original measures of the aggregate level of state support for the DSM (what I call its 'political capital') between 1995 and 2013, this article finds that dispute panels are politically savvy when it comes to reviewing government decisions. When WTO members are relatively more critical of the DSM, panels will signal more deference to a government's regulatory choices, effectively allocating authority over those policy spaces to states. Conversely, panels tend to validate fewer elements of a trade regulation only when they enjoy relatively greater support among members.

Yet WTO panels are not purely political actors and are very much aware of the dual political and legal environments within which they issue their judgments (Steinberg 2004). Dispute panels are accountable to a higher legal authority—the Appellate Body—and while not strictly bound by previous case law, some argue that a form of *de facto* precedent operates within the WTO (Bhala 1999*a,b*). In addition, for reasons of personal and professional self-interest, panelists

often do not want to see their decisions overturned. Indeed, this article finds further that the density of relevant Appellate Body case law moderates the influence of diffuse political pressures on dispute outcomes, suggesting that panels simultaneously seek to maximize their support with both their legal and their political audiences.

These findings have important implications for our understanding of the way in which authority is created and sustained within international institutions, and particularly international judicial bodies. The politics underlying how panels exercise their delegated authority also has implications for debates regarding sovereignty costs and the accountability of international courts more generally (Hathaway 2008; Lake 2007*a*). The ability of these bodies to affect the everyday lives of citizens is not trivial. Take, for instance, WTO rulings on policies enacted to protect human health, safety and the environment, areas long considered central domains of state authority. In the United States, such rulings have weakened consumer protection measures regulating the packaging of meat products and tuna labeling requirements under dolphin protection laws (Strawbridge 2013). These decisions have sparked heated criticism by politicians and public advocacy groups that the WTO is encroaching on the right of sovereign states to decide how best to protect their citizens.³

By identifying how panels accommodate political and legal considerations within their rulings on domestic regulatory choices, this article's findings suggest that the WTO's authority is neither completely dependent on nor entirely divorced from member state support. Governments seeking to retain exclusive sovereign authority over such policy decisions can actively—and *collectively*—use diffuse rhetorical means to signal that the DSM has overstepped the scope of judicial authority members are willing to support. This may in turn pre-empt more severe forms of political backlash that could undermine the stability of the international trade regime.

The remainder of this article proceeds in three parts. Section 2 develops the article's theoretical argument for why and when we would expect panels to signal deference to national authorities within their rulings. Section 3 outlines the empirical strategy used to test these theoretical expectations and discusses the primary findings of the analysis. The final section

³See, for example, Tom Miles and Doug Palmer, "WTO dents U.S. ban on clove cigarettes," *Reuters* (4 April 2012); Tim Carman, "Tuna, meat labeling disputes highlight WTO control," *The Washington Post* (10 January 2012).

concludes.

2. JUDICIAL REVIEW WITHIN THE WTO

WTO dispute settlement begins with a complainant (the government bringing the case) alleging that the trade measure(s) of a respondent (the ‘defendant’ government) violate WTO law. If the dispute is not settled during the consultations period, the complainant may request that a panel of three individuals adjudicate the dispute. Following oral and written arguments, the panel issues a final report that contains multiple findings on whether the challenged measures violate WTO law (Understanding on Rules and Procedures Governing the Settlement of Disputes (DSU), Articles 4-16). In reaching these findings, panels necessarily exercise a form of judicial review over the acts of states subject to their jurisdiction.

Panels’ exercise of this review function holds political and legal significance for the regime as a whole, and is viewed by governments, commentators, and WTO officials as an important—albeit sometimes controversial—issue.⁴ Yet the legal standard of review is flexible, not fully specified within WTO law, and not clearly developed by the Appellate Body, which gives panels considerable *discretion* in how to apply it to the the case before them. This provides a unique opportunity to examine the way in which legal and political factors simultaneously shape the contours of international trade authority.

Despite frequent warnings about the threat posed by judicial activism, international courts are not always interested in extending their authority or interpreting treaties expansively (Helfer and Alter 2013). The fact that panels can increase the scope of the DSM’s authority through their exercise of judicial review is not to suggest they always wish to do so or that they are always successful when they do.⁵ This is particularly the case given that the WTO describes itself as “member-driven,” and governments view themselves as the primary constituents of the Organization and its dispute settlement system (Footer 2006: 101; WTO 2011).

⁴The standard of review issue almost de-railed negotiations and represented “one of three or four issues that could have broken [them] apart” (Jackson 1994: 139). Ultimately, a compromise was reached by including an explicit standard within the Anti-Dumping Agreement (ADA) alone, in order to appease the United States (Oesch 2003: 72-80). The remaining WTO agreements, however, do not specify the standard of review to be applied.

⁵Similar to courts within new democracies where judges seek to protect themselves and their institutions from political reaction, the DSM operates within a political context where norms of judicial independence are underdeveloped and expectations that governments will comply with decisions are unclear (Helmke 2002: 291).

Given the environment in which they operate, I argue that panelists are strategic actors who are sensitive to various forms of political pressure, especially when the institutional legitimacy of the DSM appears fragile.⁶ A straightforward expectation derived from a view of panelists as strategic political actors is that they will exercise judicial restraint when they are likely to face sanctioning by states parties, particularly economically dominant ones; if this is the case, we would observe panels largely tilting their rulings in favor of the more powerful party of the disputing dyad. However, as discussed below, direct and unilateral sanctioning by the economically dominant members is difficult to achieve in practice, due to collective action problems and the institutional structure of the DSM. Instead, this article suggests that member states use diffuse rhetorical pressures to voice criticism of or support for panels' exercise of authority, a claim supported by government representatives within interviews.

Following from this, I argue that panels pay attention to their support among the membership *as a whole*, and not solely the largest trading partners. This is because panels anticipate how government responses to their decisions—both rhetorical and behavioral responses—will impact their ultimate goal: ensuring the survival and increasing the institutional legitimacy of the DSM. In other words, collective support for the WTO's judicial bodies represents one of the policy preferences I hypothesize panels seek to maximize through adjudication. If this is the case, we should see panelists not only paying attention to dips in their *collective* political support but also responding to these diffuse pressures through the content of their rulings. On the other hand, if panelists are pandering more to the largest traders, we would not observe any consistent relationship between collective political support and panel rulings.

These diffuse political pressures are not entirely determinative of a panel's exercise of authority. The WTO subordinates panels' judicial authority to that of the Appellate Body, which creates an additional set of legal constraints. If panels are strategic legal actors, we would expect them to anticipate the response of the higher judicial body to their ruling and strive to ensure that they will not be overturned. I argue further, however, that panels are playing to their two audiences—legal and political—simultaneously. If this is the case, we should see panelists

⁶Strategic approaches to judicial behavior have come to dominate both the American and the comparative literatures on courts. These approaches emphasize how judges seek to maximize personal policy preferences but take into account anticipated responses of other legal and political actors (Epstein and Knight 1998, 2000; Helmke 2002).

considering *both* the views of legal actors (the Appellate Body) and political actors (member states) when deciding how to rule. This section develops this theoretical argument and teases out a number of observable implications, drawing from interviews conducted with government representatives and WTO officials.

2.1. *The Political Capital of the DSM*

The claim that international judges are sensitive to diffuse forms of political influence is relatively under-theorized within the literature, even though many recognize that states frequently employ such “horizontal pressure” (Alford 2006: 220; Ehlermann and Ehring 2005: 819; Ginsburg 2005; Goldstein and Steinberg 2008: 273-4; Mavroidis and Cottier 2003: 2; Steinberg 2004). Anecdotal evidence of the political pressures exerted by member states abounds, but we lack a comprehensive understanding of their use and consequences over time. In the context of the WTO, commentators largely agree that states find it difficult if not impossible to engage in formal legislative response to rulings of the organization’s quasi-judicial bodies.⁷ Instead, as I argue below, governments rely primarily on diffuse political pressures and legal persuasion to influence panels’ exercise of their review authority. Members engage in a continuing ‘dialogue’ with the DSM by conveying (dis)satisfaction with its decisional behavior and by rhetorically challenging or supporting its institutional legitimacy (Alter 2008; Ginsburg 2005; Helfer 2006).

Meetings of governments representatives—particularly meetings of the body tasked with overseeing the dispute settlement system (the Dispute Settlement Body (DSB))—constitute the primary public forum within which WTO members express views on the DSM and seek to exert such political influence.⁸ The formal purpose of the DSB is to ensure the effective and efficient

⁷The WTO Agreement provides for two formal mechanisms of legislative response: amendment of the covered agreements (Article X, WTO Agreement) and adoption of authoritative interpretations (Article IX:2, WTO Agreement). Within the domestic context, mechanisms of legislative response or override play an important role in ensuring accountability between branches of government and minimizing the “countermajoritarian difficulty” of unaccountable judges overruling the “will of the legislature” (Bickel 1962: 16; Ely 1980: 5; Eskridge 1987). However, due to the practice of consensus decision-making within the WTO, member states have not been able to make effective use of these mechanisms. The amendment procedure is inefficient and understandably lengthy, and states have made relatively few efforts to adopt interpretations, with none adopted to date.

⁸This section draws from a series of interviews conducted with Member representatives and WTO Secretariat officials. A total of twenty-nine interviews were conducted between 13-17 January 2014 in Geneva, Switzerland. Three interviewees were officials within the WTO Secretariat; twenty-five interviews were conducted with current or former delegates representing their respective Members within DSB meetings; one interview was conducted with a representative from the Advisory Centre on WTO Law, an independent organization that provides legal

operation of the system, and both government representatives and the WTO Secretariat proudly note that the DSB is viewed as the least politicized and most efficient body within the WTO. But they also see DSB meetings as serving an important communicative role, by providing a public forum within which governments may send political signals to disputing parties, the membership as a whole, domestic audiences, and the DSM itself (Mueller-Holyst 2005: p. 25).⁹ The DSB typically meets once per month for regular sessions, with statements made on the record considered the official position of the government. In the view of a number of representatives, the formality of DSB meetings—the fact that statements are made publicly, officially and on the record—plays a critical role in fulfilling the DSB’s underlying purpose and also influences a government’s decision to make a statement. Not only do governments use these communicative acts on a regular basis, they do so with the explicit intention of signaling to the WTO’s judicial bodies their (dis)satisfaction with its exercise of authority.

Governments use DSB statements to convey to the DSM as a whole (whether future panelists, Appellate Body members, or the Secretariat) their (dis)satisfaction with its overall operation, jurisprudential developments or decisional outcomes.¹⁰ The quantity, type and nature of these expressed views represent the DSM’s political capital—or the aggregate level of state support for its exercise of authority—at a given moment in time. More critically, the majority of these statements do not represent ‘cheap talk’ by states. First, these statements typically necessitate considerable research, analysis and drafting by government officials prior to a DSB meeting. In other words, the content of a statement is pre-meditated and purposive. Second, these statements place on the formal, public record of an international political body a government’s ‘official’ view on a given issue. Because these views may impact a state’s bargaining position or diplomatic relationships within other forums, governments carefully and intentionally decide when and what views to express. Third, while a few statements may simply represent a losing

advice and assistance to developing and least-developed countries. Interviewed Members varied across relevant characteristics, including size, wealth, use of the dispute settlement system, and vocal participation within meetings of the DSB. The identities of all interviewees have been redacted and replaced with random numbers, to ensure interviewee confidentiality.

⁹Interview 4.3, Geneva, Switzerland (16 January 2014).

¹⁰The majority of representatives interviewed emphasized that they believe the Secretariat (both the Legal Affairs Division and the Appellate Body Division) is paying attention to their statements. Almost all interviewees emphasized that whether statements actually influence subsequent decisions is a separate question. Interview 2.3, Geneva, Switzerland (14 January 2014); Interview 4.3, Geneva, Switzerland (16 January 2014).

state complaining about or a winning state approving of a ruling, governments without a direct stake in a case often express views on the broader, systemic implications of a court's exercise of authority.

How might we expect panels to respond to such diffuse forms of pressure? I argue that both professional role orientations and self-interest motivate WTO adjudicators to promote the DSM's political support and ensure its survival, as these factors critically influence their own personal salary potential, professional prestige, and occupational ambition (Schauer 2000). As strategic actors interested in ensuring continued support for their exercise of authority, the WTO's judicial bodies pay attention to their collective political capital, as expressed within views of member states. Given the difficulty of employing unilateral sanctioning within the WTO, we would expect them to have an interest in cultivating it among the wider membership, and not solely in relation to the largest economies or most 'powerful' states within the regime.

The extent to which cultivating the collective political capital of the DSM drives judicial decision-making behavior varies across the two levels of WTO adjudication—the dispute panels and the Appellate Body. Appointed individuals serve as panelists on a part-time basis, in addition to their usual job.¹¹ On average, panelists have shorter time horizons than elected international judges, as they are appointed *ad hoc* for a given dispute and only have this 'one shot' to achieve their individual goals or motivations for agreeing to serve on a panel. Due to their shorter time horizons, we would expect panelists to be more sensitive than members of the Appellate Body (who enjoy relatively longer tenures) to short-term fluctuations in political support for the DSM.

Why would panelists care about or invest in cultivating the DSM's political capital, given that they are appointed *ad hoc* and hold other, permanent jobs? First, for reasons of professional self-interest, a number are likely motivated by re-appointment. A little over half of the panelists appointed since 1995 have sat on more than one panel, with a few sitting on as many as ten or eleven separate panels. Repeat panelists are likely more integrated in the organizational life of the WTO, and some even go on to become Secretariat officials or Appellate Body members. For

¹¹To date, the majority of appointed panelists were concurrent or previous government delegates to the WTO. Some were capital-based trade officials, with a few having served previously as WTO Secretariat officials or Appellate Body members. Even fewer have come from private practice or academia.

this reason, they seek to maintain a respected professional reputation and strive to issue reports that will not provoke widespread political backlash and sanctioning.¹²

Second, panelists integrated into the organizational life of the WTO likely are more sensitive to vacillations in political support for the dispute settlement system and more wary about judicial overreaching into sovereign authority. Those who are concurrently WTO representatives must regularly interact with the individuals representing the complainant or respondent, within meetings of committees and other political bodies. In addition, WTO delegates likely are more attuned to the need to provide home governments with some sort of political cover to implement adverse judgments. Tellingly, no panels have been composed with all three panelists lacking any prior experience within or interaction with the WTO.¹³

Third, even if a panelist is not motivated by re-appointment, she both: (a) possesses imperfect information on how the panel report will impact her long-term professional ambitions; and (b) must reach a collective agreement on the report's findings with two other panelists, who may themselves be motivated by re-appointment.¹⁴ More critically, even indifferent or lazy panelists can hedge their bets and engage in low-cost risk-averse decisional behavior. Panelists do not need to conduct extensive and time-consuming research to obtain information on the DSM's political capital because aggregate revealed preferences of governments are pre-assimilated by officials within the Legal Affairs and Rules Divisions of the WTO Secretariat, who are tasked

¹²Repeat panelists have an interest in cultivating support among the *wider* membership—and not always or necessarily with particular governments—because the majority of panelist appointments are made by the WTO Director General, and not by the disputing parties. To date, 60.8% of panels have been composed by the WTO Director General. Moreover, the selection procedure for panelists protects against the operation of a direct nationality bias, given that nationals of a disputing state (or third party) may not serve without the parties' agreement (Article 8.3, DSU). Since the inception of the WTO, this has only occurred three times, all in the context of EU member state panelists hearing disputes involving the European Union.

¹³Between 1995 and 2013, only seven panels (less than 5%) were composed with the majority of panelists having no prior experience within the WTO, while 62.5% of panels were composed entirely of individuals with some prior experience within or interaction with the WTO. In fact, around 20% of panels were composed of all WTO/GATT representatives and fully two-thirds of the panels consisted of a majority of delegates. At least one individual with prior WTO or GATT Secretariat experience has sat on 22% of panels, and 40% of panels have had at least one academic sitting on them. This data was obtained by collecting any *curriculum vitae* and information on professional and educational history of each individual panelist available on-line and from authoritative sources. This information was then coded for professional experience both pre- and post-panel appointment. Professional backgrounds that entailed acting as a WTO/GATT representative, an official within the WTO/GATT Secretariat or Appellate Body, or working within their home ministry responsible for international trade were coded as prior experience within or interaction with the WTO.

¹⁴The organizational norm of consensus decision-making also operates within the DSM, and there is considerable pressure for panel reports to reflect consensus among the three panelists, although individual panelists may—but very rarely do—anonously include a separate or dissenting view.

with assisting the dispute panels.

The institutional structure of the WTO, with regular meetings of the political body of the DSB, facilitates the assimilation of these revealed collective preferences. Secretariat officials within the Legal Affairs Division attend every meeting of the DSB and help transcribe the statements made on the record into formal minutes. In addition, they pay particular attention to aggregate revealed views on the DSM's exercise of authority.¹⁵ They effectively provide a low-cost way of transmitting these rhetorical signals to panelists, who are thereby able to pay attention to fluctuations in the DSM's political capital among the wider membership. Secretariat officials are incentivized to do so because they have relatively longer time horizons than individual panelists. Moreover, these officials seem to hold strongly internalized role perceptions, in that they subscribe to the WTO's self-identification as a Member-driven organization and view the Secretariat's role as serving the interests and needs of the member states. In the context of the dispute settlement system, this entails fulfilling its stated purpose: facilitating the settlement of disputes by drafting rulings in a way that will secure compliance.

To summarize, I hypothesize that panels—and the Secretariat officials supporting panels—not only pay attention to fluctuations in the DSM's political capital but will, under certain conditions, respond to these rhetorical pressures through the content of their rulings. The DSM's political capital and conversely diffuse political pressures are largely a function of the quantity and quality of recent signals of dissatisfaction with the judicial bodies' exercise of authority. It is not determined solely by the preferences of powerful states or by collective responses to extreme or controversial instances of judicial overreaching. Rather, the shaping of and contestation over the balance of authority between member states and the WTO's judicial bodies occurs on a near-monthly basis within the DSB. For these reasons, I expect panels to exercise greater judicial restraint—thereby shifting the balance of authority slightly away from member states and toward the WTO—when the DSM's support among the broader membership declines.

¹⁵Secretariat officials indicated that they pay more attention to DSB statements about procedural issues or government signals of systemic concerns about the operation of the DSM, than to individual complaints of the 'losing' party to a dispute. Interview 2.3, Geneva, Switzerland (14 January 2014); Interview 4.3, Geneva, Switzerland (16 January 2014).

2.2. *The Legal Context of Dispute Panels*

Judges—including international ones—are not purely political actors. Professional role orientations often motivate judges to consider how their rulings comport with the existing legal framework, which may also contribute to the continued legitimation of their exercise of authority. Compared to many other international judicial bodies, WTO dispute panels occupy a somewhat unique space as they effectively represent the court of first instance, with their decisions subject to appeal. Unlike the trial chambers of a number of international(ized) criminal courts, however, each panel is appointed *ad hoc*. As argued above, this not only incentivizes panelists to rely on Secretariat officials but also creates shorter time horizons, which we would expect to increase their sensitivity to short-term fluctuations in the DSM’s political capital. More so than many other international judicial bodies, then, we should see panels simultaneously playing to both a legal and a political audience.

Remarkably strong and highly internalized professional norms push international adjudicators and lawyers to follow the “discursive constraints (e.g., procedural rules, interpretive methodologies, and substantive norms)” of their profession (Helfer 2006: 266). In addition, the perceived legitimacy of a court derives partially from the degree to which its rulings are consistent and contain high-quality reasoning motivated by the applicable legal framework (Voeten 2014). Despite the absence of a formal rule of *stare decisis* in WTO dispute settlement (Pelc 2014), we would expect these professional norms motivate panelists to seek consistency across cases and indeed, a type of *de facto* precedent is said to operate within the WTO (Bhala 1999*a,b*; Palmetier and Mavroidis 2004).¹⁶ In fact, the Appellate Body has been fairly strict about reversing panel findings that explicitly go against its previous case law, at times explicitly chastising panels that do so.¹⁷ Panelists seeking to avoid appellate reversal of their rulings often look to

¹⁶Although the Appellate Body has emphasized that there is no formal doctrine of *stare decisis* in WTO dispute settlement, it has also indicated that panel decisions should create reasonable expectations that similar disputes will be decided similarly. Appellate Body Report, *Japan—Taxes on Alcoholic Beverages*, WT/DS8/AB/R, WT/DS10/AB/R, WT/DS11/AB/R; WT/DS8/R, WT/DS10/R, WT/DS11/R, 4 October 1996, pp. 13-14; Appellate Body Report, *United States—Final Anti-Dumping Measures on Stainless Steel from Mexico*, WT/DS344/AB/R, 20 May 2008, para. 161.

¹⁷The Appellate Body has also stressed that in order to “ensure ‘security and predictability’ in the dispute settlement system. . . absent cogent reasons, an adjudicatory body will resolve the same legal question in the same way in a subsequent case.” Appellate Body Report, *United States—Final Anti-Dumping Measures on Stainless Steel from Mexico*, WT/DS344/AB/R, 20 May 2008, para. 160.

and cite extensively from relevant Appellate Body case law. The existing legal framework and relevant jurisprudence thus should act as constraints on judicial choice and panel discretion. For this reason, I expect rulings on provisions litigated extensively at the appellate level will cluster and converge, whereas findings under provisions with little existing appellate case law will display greater variance (as panelists have a greater range of discretion within which to act).

3. EMPIRICAL STRATEGY

The previous section argued that a primary goal of panels is to increase support for the DSM's continued exercise of authority. If this is the case, we should see panels affording governments greater political cover to implement adverse rulings—by validating more elements of trade regulations within their rulings—as the DSM's institutional legitimacy declines. The previous section argued further that panelists simultaneously assess the degree of legal discretion available and the extent of collective political support for their exercise of authority. For this reason, I hypothesize that panels will first look to any available guidance provided by Appellate Body case law instead of the expressed views of members, as they likely are more concerned about scathing reversals than *post hoc* criticism by states. Still, panels do retain considerable discretion in many areas of WTO law. Thus I also expect that when they have relatively less direction from the Appellate Body, they will look primarily to members' preferences. In this way, panels are able to use the content of their rulings to balance the goal of legal consistency and coherence with the maximization of political support among a broad subset of the DSM's stakeholders (Kelemen 2001).

This section tests these expectations empirically. It first describes the measures used to proxy the outcome variable of interest—panel validation—and the primary explanatory variables—the DSM's political capital and the legal constraints faced by panels. It then outlines the models used to estimate the relationship between political capital and validation and discusses findings.

3.1. Data

3.1.1. *Panel Validation within Mixed Rulings*

Each case that comes before the WTO’s judicial bodies raises a number of distinct issues and claims that must be resolved. Of the 188 reports issued between 1995 and 2013, a panel found *no* breach of WTO rules across all claims in only ten instances (5.3%).¹⁸ Binary measures of dispute outcomes thus provide very little variation and moreover do not adequately capture the underlying extent to which panels signal to governments judicial restraint within a given ruling.¹⁹

Instead of focusing on case outcomes, I employ a measure of panel validation that seeks to capture the degree of judicial restraint a panel exercises within a ruling, under the assumption that panels are able to signal greater restraint when they make fewer breach findings within each dispute. Take, for example, disputes raised under the Anti-Dumping Agreement. This agreement creates substantive and procedural requirements for the imposition of remedies in response to the ‘dumping’ of imported products that cause material injury to a domestic industry. These disputes often raise claims regarding: whether the national investigating authority relied on sufficient evidence to initiate a dumping investigation; the authority’s use of facts available; the authority’s injury and/or causation analysis; and various procedural obligations, such as disclosure and notification requirements. A panel may find that an investigating authority’s injury and causation analyses conformed to WTO rules, but that the national authority failed to fulfill notification (procedural) requirements. Arguably, the latter finding—while still considered

¹⁸Some studies of WTO disputes code case outcomes along three values: win, loss and ‘mixed’ (Busch and Reinhardt 2006). Nearly 58% of the panel reports issued between 1995 and 2013 are mixed, but there is considerable variance in the mixed category in terms of the number and type of violation findings.

¹⁹Additionally, the strength of the legal case might be determinative of a finding of at least one violation. While this may be a concern for my measure of panel validation as well, the claims raised within the empaneled stage of a dispute have gone through a filtering process. During each stage of a dispute—the decision to request consultations, the consultations themselves, the decision to request establishment of a panel, and arguments before the panel—the complainant may discover that it has an incredibly weak legal basis for one of its claims. Governments have numerous incentives (cost, time, limited resources, and potential negative diplomatic effects) to drop claims they discover are weak and to focus on claims they believe have the strongest chance of succeeding. The consultations stage often helps clarify the legal basis of various claims and provides the complainant with critical information regarding the strength of its legal argument for each of its claims. In fact, a number of complainants raise claims within their requests for consultations that they do not subsequently pursue within their requests for panel establishment or requests for findings at the litigation stage. This suggests that the claims on which the panel ultimately rules are ones for which there does exist reasonable uncertainty over the legal outcome, thus providing the panel with considerable discretion over whether or not to find at least one violation.

a ‘violation’—is easier to implement than a complete restructuring of the investigating authority’s methodology for conducting its injury or causation analysis.²⁰ The treatment of discrete national laws or policy decisions directly connects to how a panel is allocating authority in practice. By validating some domestic choices—even if ultimately the respondent ‘loses’ the case—a panel can signal to states that it is sensitive to sovereignty concerns, judicial overreaching and the balance of authority within the regime.²¹

To construct this measure of panel validation, each dispute report issued between 1995 and 2013 was assigned a score that represents the proportion of discrete findings made by the panel within which it validated an element or aspect of the trade policy challenged.²² This VALIDATION SCORE is a continuous variable that ranges from zero to one. Reports receiving a score of one found no instances of breach, signaling complete validation of the trade policy under review, while those receiving a score of zero contain only breach findings across claims, signaling complete invalidation. Panels initially provided very little validation of government’s trade policy choices, but then began to signal increasingly greater judicial restraint—or provide greater political cover—over the first six years of the DSM’s operation. Following this initial sharp increase, however, average panel validation has fluctuated around the 0.5 mark, though with relatively constant variance over those years.

²⁰A proposal by the United States and Chile within the Dispute Settlement Understanding Review process supports this assumption, or at least that members are very much attuned to the discrete elements of dispute rulings and the relatively higher implementation costs imposed by a greater number of violation findings. See Communication from Chile and the United States: Flexibility and Member Control, WTO Doc. No. TN/DS/W/89 (31 May 2007). The assumption that the number and types of violation findings affect relative implementation costs was further supported through interviews conducted with Member representatives. Interview 2.2, Geneva, Switzerland (14 January 2014); Interviews 5.5, 5.7, Geneva, Switzerland (17 January 2014).

²¹Similarly, some scholars have disaggregated the U.S. administrative law *Chevron* analysis (the standard for judicial review of administrative actions) into its discrete steps, coding for whether the reviewing court deferred to the administrative agency at each step of the analysis (Czarnecki 2008). Even if a court ultimately overturns an administrative agency determination, it can choose to signal through its analysis a level of deference to some aspects of that decision.

²²Each panel report was coded for the number of claims raised by the complainant for which the panel undertook a legal and factual analysis—that is, claims under which the panel engaged in substantive review of a domestic measure and the national authority’s decisional process (see Appendix for coding instrument). While the number of issues raised and decided will vary by case and by Agreement provision, the percent of non-breach findings provides a standardized measure across all disputes.

3.1.2. *The DSM's Political Capital*

To proxy the political capital of the WTO's judicial bodies, I use measures of aggregate support for and criticism of the DSM's exercise of authority as voiced by member governments within the DSB. To construct these measures, I first compiled all individual statements made within the DSB since the WTO's inception.²³ I then used methods of automated text analysis to estimate the percentage of government statements within the DSB that express a critical, neutral or supportive view on the DSM's exercise of authority.²⁴

Within the 340 DSB meetings held between 1995 and 2013, governments have made a total of 9,833 statements, an average of 518 statements per year. States were the most vocal in absolute terms during the late 1990s, but DSB participation subsequently declined pretty sharply. After 2005, however, representatives began to speak more, with some minor fluctuations (see Figure 1). The notable rise in DSB participation after 2005 is largely attributable to an increase in discussion surrounding some governments' (non)implementation of dispute rulings, as the frequency of statements does not correspond with absolute levels of dispute settlement activity, at least in terms of the number of requests for consultations filed each year, which has declined steadily over the years (Figure 1).

Trends in aggregate statement sentiment over time (Figure 2) suggest that states began to change the way they used DSB statements following the first decade of the WTO's operation. After 2006, governments begin to make proportionally more statements estimated to fall within the 'Other' category—statements not explicitly about the DSM but that still represent engagement with the dispute system—and proportionally fewer neutral statements. Over the course of the first ten years of the DSM's operation, states (and particularly those that use the system regularly) may have begun to tacitly accept the authority of the WTO's judicial bodies and thus did not see a need to express a view (neutral or supportive) unless it related to an issue of

²³Official minutes of the 340 DSB meetings held between February 1995 and December 2013 were obtained from *WTO Documents Online* (<https://docs.wto.org/>), from which individual Member statements were extracted and coded for relevant meta-data.

²⁴I employed proportion classification methods using ReadMe (Hopkins and King 2010) to obtain final half-year estimates for the out-of-sample statements because it performed better than individual classification methods. To identify the sentiment of the view expressed, DSB statements were placed into one of four categories: Critical, Supportive, Neutral, and Other. The Other category captures those statements that do not reference or express a view on the DSM specifically (though may be about disputes more generally). See Appendix for description of coding scheme and validation tests.

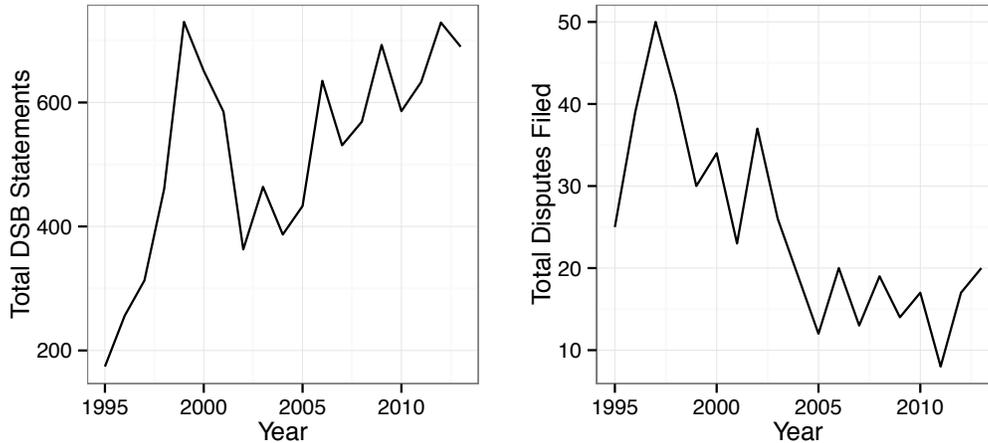


Figure 1: Total statements within the DSB (N=9833) and total complaints filed (requests for consultations)(N=464), by year. Source: WTO *Documents Online* (<https://docs.wto.org/>).

high importance. This would suggest that governments tend to express views on the DSM when they are (particularly) dissatisfied, which may partly account for the slight increase in the relative proportion of statements criticizing the WTO’s judicial bodies. In other words, statements before 2005 likely reflected the entirety of members’ ‘sincere’ views, as they were learning about and beginning to engage with the new system. After 2006, however, statements about the DSM may be becoming more selective and strategic.

Participation within DSB meetings varies fairly predictably across states, with the United States and the European Union the most vocal, having made 2,069 and 1,536 individual DSB statements respectively (compared to the next most vocal country, Japan, with a total of 615 individual statements). Expressed views on the DSM have also varied over time depending on how often a country uses the system. Figure 3 displays the proportion of statements made by states, grouped according to their use of the system, that are critical of the DSM.²⁵ The United States and European Union (high use) make proportionally fewer critical statements than all other governments, whereas countries that rarely or never use the dispute settlement system

²⁵DSM usage was calculated according to a Member’s participation as a party within empaneled disputes and thus reflects engagement with the DSM specifically, as opposed to the dispute settlement system as a whole. The high usage category only includes the U.S. and the EU, while the medium category includes the next ten most active users: Canada, the Republic of Korea, China, Japan, Mexico, Brazil, India, Argentina, Thailand and Australia. The low category includes all other Members that have participated as a party within at least one empaneled dispute (twenty-six Members in total), while non-use includes Members that have never directly participated in an empaneled dispute (though may have participated as a third party).

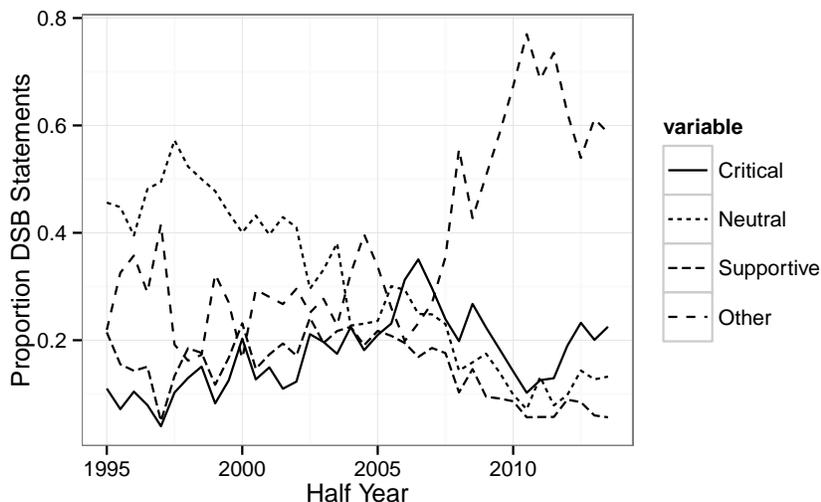


Figure 2: Estimates for proportion of total DSB statements by half-year, categorized by statement sentiment (Critical, Neutral, Supportive and Other). Total DSB statements = 9833.

have become increasingly critical of the DSM over time, reaching a high in 2006 before leveling out somewhat.

3.1.3. *Legal Constraints*

In order to take into account a panel’s range of legal discretion, I collected data on each panel report’s existing and relevant Appellate Body case law. The AB CASE LAW variable was constructed for each Agreement provision by first calculating the cumulative count (per year) of Appellate Body reports that made a finding under that provision. For every provision under which the panel made a substantive finding (breach or non-breach), the AB CASE LAW count for those provisions from the previous year were summed, and then normalized with the panel’s total (single) provision findings. While Appellate Body jurisprudence noticeably accumulated for some provisions over the years, there are other agreements and provisions that have not been litigated extensively, which accounts for greater variance in the AB CASE LAW measure over time.

Critically, the hypothesized direction of the Appellate Body’s influence on panel validation will depend on how closely it believes panels should review domestic policy choices. For the most litigated provisions at the appellate level (primarily those regarding national treatment and general exceptions to GATT rules, and a few provisions under the three trade remedy

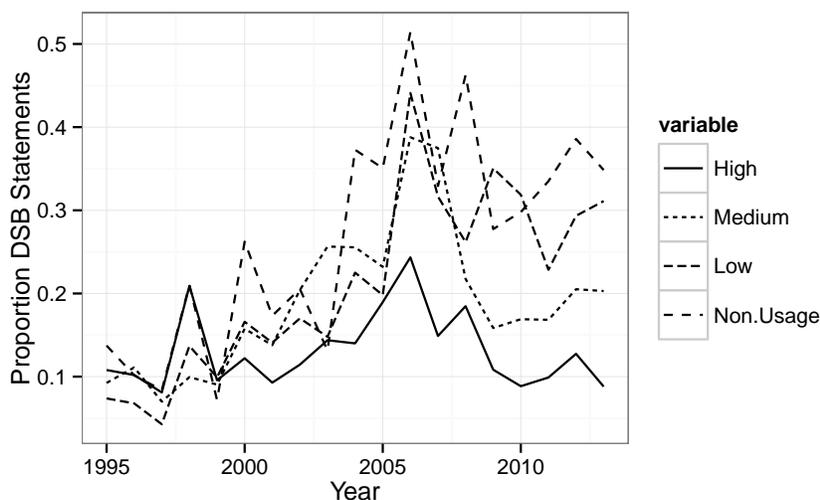


Figure 3: Yearly estimates of statements categorized as Critical, as proportion of total DSB statements, by member state’s use of the WTO dispute settlement system. High use category includes the U.S. and EU (Total DSB statements=3605). Medium use category includes the next ten most active users (as parties to an empaneled dispute) of the dispute settlement system (Total DSB statements=3671). Low use category includes remaining states that have been party to at least on empaneled dispute (Total DSB statements = 1713). Non-Usage category includes all member states that have never been party to an empaneled dispute (Total DSB statements=894).

agreements), the Appellate Body has tended to push panels to engage in more searching review of government agency decisions, though the extent to which this is the case varies over time.²⁶

As this section has demonstrated, the legal constraints and political pressures faced by panels have steadily increased, yet they have also varied across issue areas and groups of member states. This last point is key, as unlike previous examinations of the ‘strategic space’ of the Appellate Body that focus on the role of the most powerful members (Elsig and Pollack 2012; Steinberg 2004), I seek to identify the role of the full WTO membership in shaping the organization’s exercise of judicial authority. In the following sections, I use the measures of political capital and panel validation described above to evaluate the extent to which panels pay attention to and attempt to assimilate the preferences of this wider audience, while simultaneously accommodating the legal constraints imposed by the Appellate Body.

²⁶By December 2013, Article 11 of the DSU (establishing the ‘objective assessment’ standard of review to be applied by panels) had the highest cumulative AB CASE LAW count. This further supports the importance of studying panels’ exercise of judicial review within reports. Panels are not able, however, to make findings under this provision (as it governs panels’ obligations in reviewing the claims before them) and thus this case law count is not reflected in any of the scores for AB CASE LAW assigned to each panel report.

3.2. Research Design and Controls

Are WTO dispute panels responsive to diffuse political pressures? Do they tend to exercise greater restraint or provide governments with greater validation of trade measures when the DSM's political capital declines?²⁷ To answer these questions, I rely on the VALIDATION SCORE described above as the outcome variable of interest, a continuous variable that ranges from zero (no validation/all breach findings) to one (complete validation/no breach findings).²⁸

The primary explanatory variable concerns the political capital of the DSM, which I proxy with half-year estimates of the proportion of statements made by members within the DSB that were critical (POLCAP (CRITICISM)) or supportive (POLCAP (SUPPORT)). For each observation, I use the estimates for one year prior to the date of circulation of a panel report. Panel proceedings typically take around one year, so introducing this time lag captures the DSM's political capital right at the point when panelists are beginning their work.²⁹ If panels are responsive to collective political pressures, we would expect a positive association between the proportion of critical statements and validation. Conversely, panels will likely exercise less restraint as political support for the DSM increases. WTO officials suggested within interviews that the Legal Affairs Division of the Secretariat, which assists panels in dispute settlement, tends to pay more attention to critical statements about the DSM than other statements within the DSB.

²⁷Admittedly, my results may suffer from selection bias if there is something about the subset of disputes not settled bilaterally that makes greater validation within panel reports more likely. However, these empaneled disputes are those over which there is reasonable disagreement about the legal outcome of the dispute and considerable ambiguity over what the rules require, as most are filtered through WTO committees and subsequently, at the consultations stage, legal arguments are clarified and information asymmetries reduced. These types of disputes are precisely why the DSM was delegated authority in the first place and the areas in which the boundaries of the DSM's authority remain untested or vague. Given the interpretive discretion such ambiguity provides, we might expect that for these types of disputes, panels would be even more likely to engage in judicial lawmaking and expansive assertions of authority than they would for the more 'clear-cut' disputes that are able to be settled bilaterally.

²⁸The level of analysis is panel report. A few disputes involve multiple complainants that filed separate requests for consultations, resulting in multiple reports for that dispute. However, as each complainant often raises distinct claims, the validation scores for reports consolidated under the same dispute almost always differ. I exclude one report from the analysis because the panel concluded that the agreement provisions relied on by the complainant were inapplicable to the dispute and did not engage in any substantive review of domestic measures (Panel Report, *Brazil—Measures Affecting Desiccated Coconut*, WT/DS22/R, 17 October 1996). The resulting sample includes 187 observations.

²⁹The average length of time between panel constitution and circulation of the panel report is 373.6 days. Lagging the explanatory variables of interest one year thus captures the political capital of the DSM right around the time of panel constitution. In unreported findings, lagging the POLCAP variables by six months did not affect the main results.

For this reason, I expect POLCAP (CRITICISM) to be relatively more influential than POLCAP (SUPPORT), and so include these two variables separately. I estimate two models, one with POLCAP (CRITICISM) alone (Model 1) and one including POLCAP (SUPPORT) (Model 2).

Additionally, I control for a range of alternative determinants of panel validation. To account for the influence of party characteristics, I include two variables. First, I control for disputes in which either the United States or the European Union were the respondent (US/EU RESPONDENT). If power politics is determinative, as some studies of the DSM intimate (Elsig and Pollack 2012; Steinberg 2004), panels will provide greater validation to governments whose support is critical to the operation of the DSM. Panels have, on average, been slightly more willing to validate trade measures of the United States than the European Union, and even more willing than all other respondents combined.³⁰ Second, I get at the power of the complainant state, which would on average favor more breach findings, by including its logged GDP per capita (COMPLAINANT GDP).

The nature of review undertaken by panels is intimately related to their ability to validate discrete aspects of a trade policy and varies with the type of measure(s) challenged within a dispute. The extent to which a panel can and will scrutinize the basis for a statute or other legislation may differ systematically from the level of scrutiny applied to an administrative or executive regulation. In addition to legislative or administrative regulations, panels are often called upon to assess the WTO-compatibility of investigations undertaken by an administrative body, such as those conducted prior to the imposition of anti-dumping duties. These investigations represent a type of quasi-judicial proceeding that includes notice, participation, and transparency obligations. To account for the variable nature of panel review across types of measures, I include a factor variable (MEASURE TYPE) with three categories: legislation; executive/administrative regulation; executive/administrative investigation(s).

To control for politically sensitive disputes, I include a binary variable (POLITICALLY SENSITIVE AGREEMENTS) for reports that made findings under the Agreement on the Application of Sanitary and Phytosanitary Measures (SPS) (11 reports), the Agreement on Agriculture (AoA)

³⁰Notably, the average panel validation (within twelve reports) provided to China, the third largest trader, is only slightly higher than that for all other members. However, China has only been a WTO Member since 2001 and has been involved in fewer disputes absolutely (nineteen in total) than the United States or the European Union.

(14 reports), or the Agreement on Textiles and Clothing (ATC) (4 reports). These agreements seek to regulate internal policies traditionally viewed as core elements of domestic sovereignty, such as measures protecting the health and safety of citizens under the SPS (Guzman 2004; Letsas 2006; Shany 2006). Agreements regulating agriculture and textiles have also generated considerable contestation within the WTO, due to the strength of domestic interest groups lobbying for continued protectionism within these sectors (Davis 2003; Eagleton-Pierce 2013). Some WTO scholars suspect that the DSM is particularly careful about how it handles disputes under these contentious agreements, though one study found that panels are not any more likely to avoid ruling on those claims (under the AoA in particular) than any other claims (Busch and Pelc 2010). In addition, disputes that involve trade remedies (claims falling under the ADA, the Safeguards Agreement (SA), and the Subsidies and Countervailing Measures Agreement (SCM)) are viewed by many trade lawyers as distinct within the WTO, in that they often require panels to engage in a different type of review of administrative agency investigations.³¹ Moreover, the ADA is the only agreement that contains an explicit standard of review and claims are often brought under the ADA and SCM together. For this reason, I control for whether the report made findings under one or more of these three agreements (TRADE REMEDIES). A little less than half (45.5%) of all panel reports involve trade remedy disputes.

I also take into account panelists' experience with and integration into the life of the WTO by including two binary variables for panels with a majority of REPEAT PANELISTS and majority of WTO DELEGATES. Repeat panelists—whether selected by the disputing parties or the Director General—as well as individuals previously or are concurrently government representatives to the WTO are likely more integrated in its organizational life, and thus more attuned to fluctuations in the political capital of the DSM. Repeat panelists in particular tend to be well-respected by WTO officials and members governments alike, and can be relatively more certain that a controversial decision (or one expanding the DSM's authority) will be met with slightly less criticism than if made by an 'untested' panelist. In contrast, individuals with no prior panel experience have imperfect information about their responsibilities and are likely not as familiar with WTO law, the culture of the Organization, or anticipated responses of members. Addi-

³¹Within trade remedy disputes, panels effectively act as a court of second-instance, engaging in both substantive and procedural review of the quasi-judicial investigation undertaken by a domestic administrative agency.

tionally, research on “freshman” or acclimation effects in the American context suggests that new judges often follow the lead of their more experienced colleagues (Hagle 1993; Hettinger, Lindquist and Martinek 2003; Wood et al. 1998). In the WTO context, we might expect novice panelists to defer to their more experienced colleagues, namely repeat panelists and the Legal Affairs Division of the Secretariat, suggesting that panels with mostly repeat adjudicators will tend to signal less restraint. Nearly half of all panels (48%) were composed with a majority of repeat panelists, 68% of all panels were composed with a majority of WTO delegates. Finally, all models include a cubic year trend variable.³²

For all models, I rely on ordinary least squares regression with robust standard errors to estimate the effect of political capital on panel validation.³³ The base estimation (Model 1) includes all the control variables discussed above and the primary explanatory variable POLCAP (CRITICISM), while Model 2 additionally includes POLCAP (SUPPORT). As discussed previously, I expect the influence of criticism to be conditional on the range of legal discretion available to the panel. To test this, I include in Model 3 an interaction term for POLCAP (CRITICISM) and the density of Appellate Body case law (in the previous year) for the provision findings made within a report (AB CASE LAW). Summary statistics on all variables are reported in Table A1.

3.3. Results and Discussion

Across all models, criticism of the DSM (POLCAP(CRITICISM)) is positively correlated with a report’s validation score, supporting this article’s primary argument (Figure 4. See Table A2 for standard regression tables for all models). Panels do appear to provide government authorities with greater political cover (through non-breach findings) when members have been relatively more critical of the DSM’s exercise of authority. The influence of criticism is also substantively significant. When controlling for the level of political support (Model 2), a ten percent increase in criticism increases the average panel validation provided by seventeen percent (Figure 4, Model 2). Similarly in line with this article’s theoretical argument, support for the DSM is negatively

³²I do not include year or half-year fixed effects as the treatment variable of interest—political capital—does not vary across observations within the same half-year.

³³I use the term ‘effect’ as shorthand for ‘relationship between,’ as identifying a causal relationship between two variables requires further conditional independence assumptions that may not hold. All statistical analyses were conducted using “Zelig” for R. See Imai, King and Lau 2008. See also Kosuke Imai, Gary King, and Olivia Lau, “Zelig: Everyone’s Statistical Software”(2007), available at: <http://gking.harvard.edu/zelig>.

correlated with panel validation, although not significant. The substantive relationship is also less than that of criticism, with a ten percent increase in support associated with around a nine percent decrease in average panel validation, even when taking into account the relative level of member dissatisfaction with the DSM. These findings provide support for the argument that panels are not only paying attention to members' collective views on the WTO's judicial bodies, but also tend to rule more in favor of the defendant when challenges to their institutional legitimacy increase.

However, panels still exercise slightly greater restraint when ruling on the trade measures of the United States or the European Union alone. While disputes brought against one of these two countries are consistently associated with slightly higher validation scores, this relationship is substantively small, and not as significant when controlling for political support within bootstrapped simulations (Figure 4). To assess the relationship between a respondent's power or wealth and panel validation further, I control for the respondent's GDP per capita (logged) in a series of robustness checks. When I do so, the significance of the US/EU RESPONDENT variable washes out completely (see Table A3). Similarly, there does not appear to be a strong relationship between the wealth of the plaintiff countries and panel validation, though the direction is as expected (panels tend to find relatively more breaches in disputes with a powerful complainant).

In terms of the measure under review, panels tend to validate more aspects of a trade measure when legislation or statutes are challenged than when reviewing administrative regulations or investigations undertaken by executive agencies. In other words, panels appear much less willing to second-guess legislators as decision-makers than executive or administrative agencies, although this association never reaches traditional levels of significance. When reviewing the conduct of and determinations reached by administrative proceedings, panels tend to validate more aspects of the investigation, though these disputes tend to involve a greater number of claims on average and the relationship is not particularly strong across all model specifications.

As for the type of case decided, the indicator for politically sensitive agreements is consistently and significantly associated with lower validation scores. This could be for a number of reasons. First, members have widely divergent preferences on the desirability of regulating agriculture and textiles under the WTO. Panels may be taking advantage of such polarization

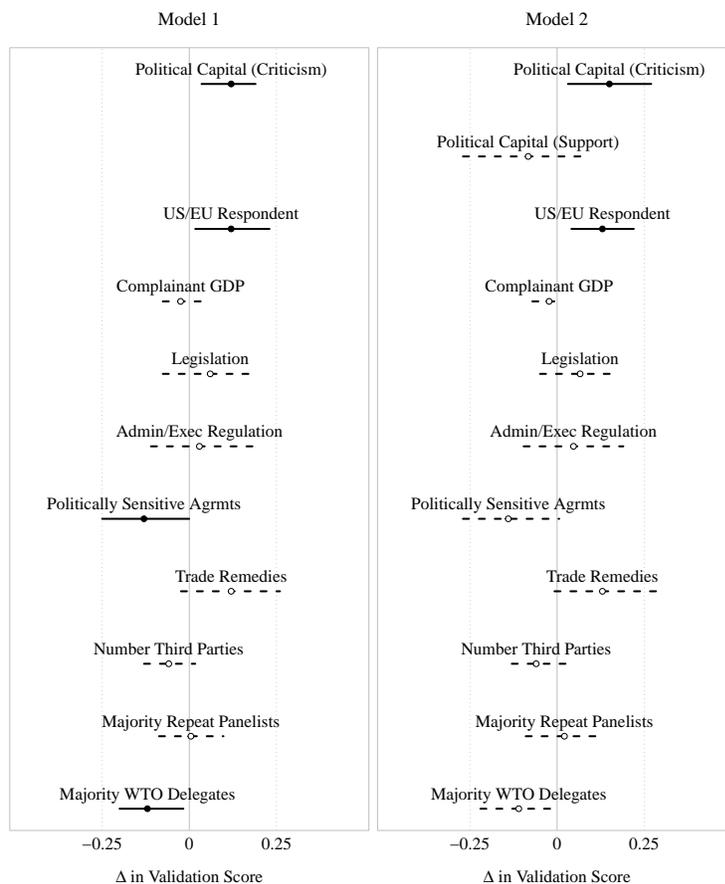


Figure 4: Estimated effects on panel validation, Models 1 and 2.

The circles are estimates of the expected change in panel validation as the indicated variable changes from 0 to 1 (for binary variables) or from their 25th to 75th percentile values (for continuous variables), and all other variables are held constant at their means. The lines are 90% confidence intervals. The circles and lines are solid when there is at least 90% confidence of a positive or negative effect on panel validation. Otherwise, circles are open and lines are dotted. Estimates obtained from simulated bootstrap parameters of ordinary least squares regression. N=182. Neither model includes Appellate Body case law; both models include a cubic year trend variable. See Table A2 for standard regression tables.

to assert greater institutional authority. Second, because these issue areas were so politically sensitive, the resulting agreements were likely drafted in more ambiguous terms than the less controversial WTO disciplines. Such vague and general treaty language affords panels greater room to maneuver and provides a legitimate justification for expansive interpretations. Finally, disputes under the ATC and AoA usually entail considerable domestic political ramifications,

with governments facing pressures from influential industries or interest groups. For this reason, even if the legal outcome of a dispute is clear or the government prefers to remove the discriminatory measure, it may not be settled at the consultations stage due to domestic political obstacles. A ruling by an external court, however, may help the respondent government overcome these obstacles by allowing it to claim that its hands are effectively tied (Simmons 2002). Panelists are likely aware of these domestic obstacles and could be seeking to facilitate compliance by providing the government with international legal cover.

In contrast to the politically sensitive agreements, rulings under trade remedy agreements are associated with slightly higher validation scores, which is not surprising, as the ADA contains an explicit standard of review intended to afford greater deference to the decisions of national investigating authorities. Furthermore, disputes involving anti-dumping duties and the standard of review applied by panels therein have been particularly controversial. One reason why anti-dumping disputes have sparked so much controversy is that the United States actively criticized a line of cases ruling against the Department of Commerce’s practice of ‘zeroing’ in anti-dumping investigations.³⁴ Following the Appellate Body’s ruling in *United States—Softwood Lumber V*,³⁵ a number of countries filed complaints against the U.S. Department of Commerce’s zeroing practice in investigations and reviews. To address the unique nature of this line of cases, I removed all disputes from the sample, following *United States—Softwood Lumber V*, that only challenged the United States’ zeroing practices.³⁶ Doing so for Model 2 strengthens the positive association for the US/EU RESPONDENT variable (see Table A3 for results), but does not alter the relationship between criticism and validation.

Panels with a majority of repeat panelists do not seem to exercise any more or less judicial restraint than those without. Similarly, the relationship between panels with a majority of WTO delegates and panel validation is weak and inconsistent across model specifications. The

³⁴Anti-dumping authorities calculate the margin of dumping for a product by computing the difference between normal value and export price for each model or type of a particular product, and aggregate the results. ‘Zeroing’ refers to the practice of omitting the calculations where export price was higher than normal value, thus inflating dumping margins.

³⁵*United States—Final Dumping Determination on Softwood Lumber from Canada*, WT/DS264/AB/R, 31 August 2004.

³⁶In five of these cases, the United States did not contest the claims or present counter-arguments. The jurisprudence became so settled around the use of a zeroing methodology in anti-dumping investigations that, absent counter-arguments by the U.S., panels only needed to decide whether the complaining party had provided sufficient evidence to demonstrate that the United States had zeroed in that particular instance.

direction of the relationship is opposite of that expected—WTO delegates appear to validate fewer aspects of a trade measure on average—yet controlling for this does not affect the impact of political capital on panel rulings. This could be indicative of the strong hand played by the Secretariat within the crafting of dispute judgments, or simply that the professional background of each individual panelist, on average, might not affect outcomes as much as some of the literature on international judicial behavior might suggest.

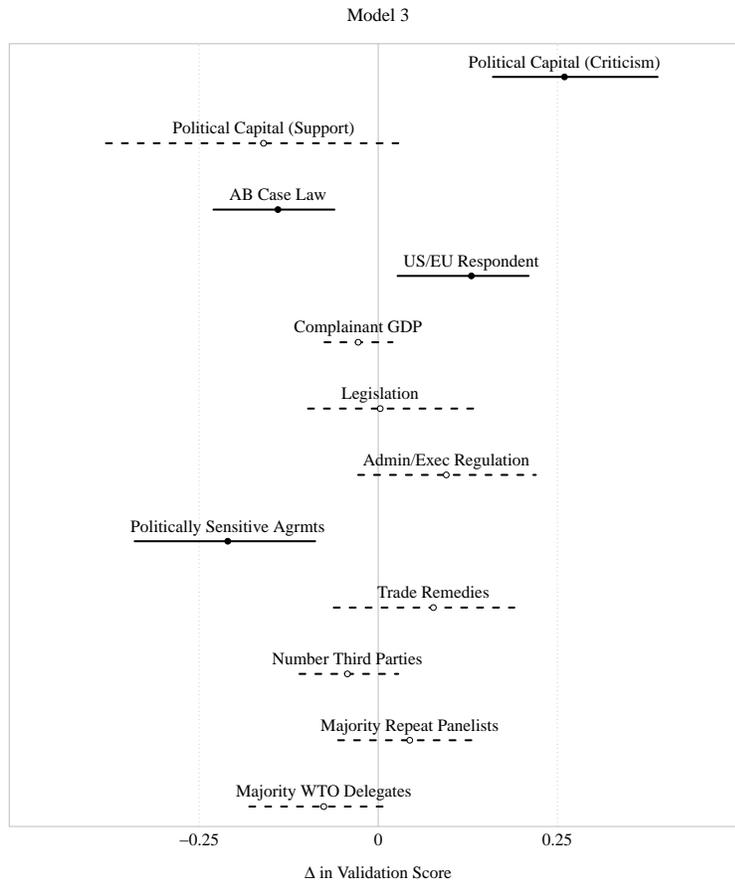


Figure 5: Estimated effects on panel validation, Model 3.

The circles are estimates of the expected change in panel validation as the indicated variable changes from 0 to 1 (for binary variables) or from their 25th to 75th percentile values (for continuous variables), and all other variables are held constant at their means. The lines are 90% confidence intervals. The circles and lines are solid when there is at least 90% confidence of a positive or negative effect on panel validation. Otherwise, circles are open and lines are dotted. Estimates obtained from simulated bootstrap parameters of ordinary least squares regression. N=182. Includes a cubic year trend variable. See Table A2 for standard regression tables.

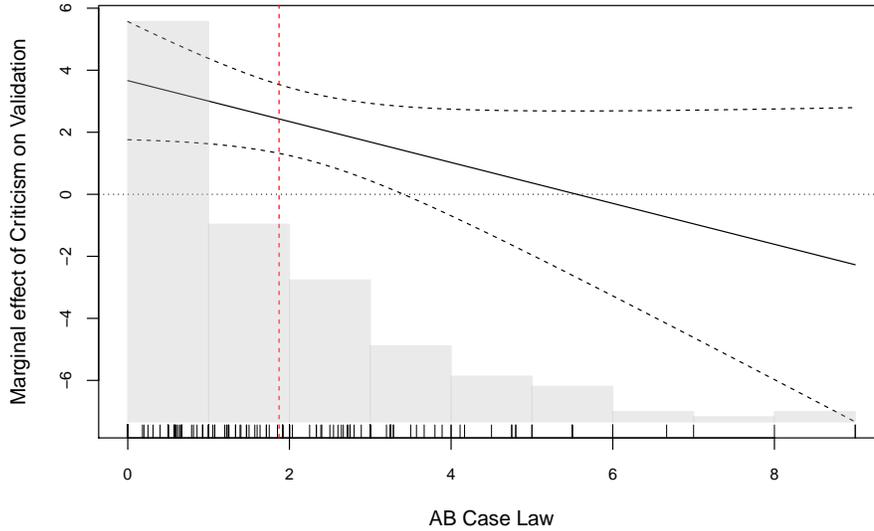


Figure 6: Relationship between Political Capital (Criticism) and Panel Validation, for different levels of Appellate Body case law (Model 3).

Findings are also consistent for Model 3, which includes an interaction term for POLCAP (CRITICISM) and AB CASE LAW (Figure 5).³⁷ When controlling for the conditional effect of political criticism on panel validation, a ten percent increase in political support for the DSM (POLCAP (SUPPORT)) is associated with a decrease in validation of seventeen percent. Similarly, criticism by members continues to have a strong association with greater validation when there is less existing Appellate Body case law on the rules being litigated.³⁸

However, the influence of these political pressures declines as appellate case law increases. As seen in Figure 6, the impact of criticism on validation is only significant when reports' AB CASE LAW scores fall below three, which includes seventy-five percent of the panel reports issued between 1995 and 2013.³⁹ This finding does suggest, though, that as appellate jurisprudence

³⁷See Table A4 for results of Model 3 robustness checks similar to those discussed above for Model 2. While the estimated coefficient of the interaction term for POLCAP (CRITICISM) and AB CASE LAW loses statistical significance when removing all disputes from the sample that challenged only the United States' zeroing practices (Models 3b and 3c), Figure A2 demonstrates that the marginal effect of POLCAP (CRITICISM) is still significant for most (lower) values of AB CASE LAW (compare Figure A2 with Figure 6).

³⁸Because this model includes an interaction term, the estimated coefficient for POLCAP (CRITICISM) is interpretable as the estimate effect of criticism when there is no existing Appellate Body case law. Only 22.5% of reports received an AB CASE LAW score of zero.

³⁹Recall that this measure is not just a count variable of Appellate Body findings, but is normalized by the number of (single) provision findings reached by the panel. An AB CASE LAW score of three thus reflects fairly high levels of existing appellate jurisprudence for reports that make findings under multiple provisions.

accumulates over time around frequently litigated provisions, panels could be expected to increasingly ignore the expressed views of members and instead look to the Appellate Body for guidance in their exercise of authority.

4. CONCLUSION

WTO dispute panels do not operate in a political vacuum, yet they are also not indiscriminately bending to the views of member states. As entities appointed *ad hoc*, panels do not enjoy the same degree of permanence or structural independence as the Appellate Body or permanent international courts. The Legal Affairs Division of the Secretariat creates a quasi-permanent foundation in that it provides the requisite institutional memory and legal assistance for panels to engage in consistent adjudication, yet Secretariat officials still very much see their role as one in the service of member states.⁴⁰ While the majority of panelists have been appointed by the WTO Director General—and not by member states—professional career incentives often push these individuals to attempt to satisfy both member governments and the Appellate Body. In this respect, panels inhabit a unique space in that their institutional design tasks them with finding an acceptable balance between legal rigidity and political realities. Panels establish this balance through their discrete findings that incrementally and subtly shape the allocation of authority over trade-related policies.

Although many claim that international institutions—and courts in particular—attempt to take into account state preferences and will sometimes independently tailor their activities to appease member states and preempt more severe forms of backlash (Alter 2009; Ginsburg 2005; Garrett, Kelemen and Schulz 1998; Steinberg 2004; Weiler 1994), we lack a clear understanding of how exactly international courts obtain information about governments' views. Existing research focuses on the information signaled by parties or third parties to a dispute (Busch and Reinhardt 2006; Busch and Pelc 2010), but as I have argued, panels are strategic actors interested in the views of a wider (more representative) subset of the membership. This article provides one of the first empirical attempts to ascertain how panels and governments 'communicate' by systematically examining the rhetorical politics surrounding the DSM's exercise of authority.

⁴⁰Interview 2.4, Geneva, Switzerland (14 January 2014).

I have found that panels, under certain conditions, are responsive to fluctuations in the level of member state support enjoyed by the Dispute Settlement Mechanism. The political capital of the DSM is not necessarily or always determined by the extent to which it has engaged in expansionist or activist judicial lawmaking, as courts can “spark controversy due to the domestic political consequences of their rulings, whether or not those rulings are expansionist” (Helfer and Alter 2013: 502). Yet the way panels respond to these diffuse political pressures is to signal greater deference to political concerns by providing government authorities with slightly more domestic political cover for adverse decisions and subtly reshaping the balance of authority between states and the WTO.

To be clear, the diffuse rhetorical pressures—and the political capital of the DSM—on which this article focuses are distinct from the degree of political controversy surrounding an issue area or dispute. This is most strikingly observed in the low levels of validation panels afford challenged trade measures under two of the most controversial WTO agreements—the Agreement on Agriculture and the Agreement on Textiles and Clothing. Members have widely divergent preferences on the scope and desirability of these agreements. Such polarization creates a situation in which panels face multiple principals with extreme views on dispute outcomes. This may embolden a panel to engage in expansive assertions of judicial authority as it provides plausible allies for its ruling. Because these agreements engender such broad disagreement among members, it may be even more important for panels to adopt a non-deferential stance and assert their authority, particularly if this helps the respondent government overcome domestic political obstacles or save face with influential interest groups. In this way, panels may be using the content of their rulings to ease the domestic political costs of incredibly charged policy changes, thereby acting as ‘partners in compliance’ with a government’s executive branch.

Panels effectively use their rulings under agreement provisions for which there exists considerably less Appellate Body case law to provide members with greater flexibility. In this way, panels help to address one of the primary institutional design concerns over how to balance greater legalism and predictability with the continued stability and adaptability of the regime (Busch and Pelc 2010; Koremenos, Lipson and Snidal 2001; Koremenos 2005; Rosendorff and Milner 2001). However, the moderating effect of the legal constraints imposed by the Appellate

Body suggests that for the most litigated provisions—national treatment under the GATT and review of trade remedy investigations by national authorities—the DSM is providing greater legalism and predictability at the expense of long-term adaptability. This could be due to the fact that the DSM’s scope of authority in these issue areas was acceptably negotiated and contested during the early cases. Greater legalism may not be problematic from an institutional design perspective, so long as this balance continues to serve the interests of members. Still, this does suggest that as appellate jurisprudence accumulates over time, panels may become less responsive overall to the expressed views of members. This paper’s findings imply, then, that if members wish to avoid or are unable to resort to more drastic forms of re-contracting for policy spaces over which they desire to retain exclusive political authority, they should engage in greater *collective* rhetorical pressure within the DSB. Additionally, these findings sound a cautionary note for panelists and the Secretariat to continue to pay attention to the political capital of the DSM, even when ruling under the most litigated WTO laws, in order to pre-empt more severe forms of political backlash that could undermine the stability of the international trade regime.

Appendix

Panel Validation Score

To construct the VALIDATION SCORE of panel reports, each report issued between 1995 and 2013 was first coded for the number of claims or issues raised by the complainant and for which the panel undertook a legal and factual analysis—that is, claims under which the panel engaged in substantive review of a domestic measure and thereby undertook to review the national authority’s decisional process. The denominator of VALIDATION SCORE was based on coding the number of Agreement provisions under which the panel made an explicit finding within the Conclusions and Recommendations section of its report, and supplemented if necessary with findings made within the report itself. The numerator of this measure is based on the number of Agreement provisions under which the panel made an explicit finding of consistency (or no violation).

This coding excludes provisions for which the panel explicitly refrained from making a finding, exercised judicial economy, or found that the claim fell outside its terms of reference. The number of issues or claims for which the panel exercised judicial economy or refrained from making a finding were noted separately, but not included within the construction of the measure, as this is one ‘issue avoidance’ technique available to panels that arguably signals a different type of judicial restraint (Busch and Pelc 2010), but that do not require panels to undertake a substantive review. That is, the judicial economy doctrine does *not* allow panels to diminish the scope of the challenge—merely the number of WTO-based reasons for finding a particular element of a trade measure incompatible with WTO rules. Once a particular element is found WTO-inconsistent on one legal theory, a panel may stop considering other asserted bases for the same claim. However, it still must continue examining all other elements of the laws being challenged.

This coding also excludes provisions cited regarding recommendations, such as specific recommendations under the Subsidies and Countervailing Measures Agreement (Article 4.7, SCM Agreement). Citation of a Dispute Settlement Understanding (DSU) provision is also excluded

unless the panel made a finding with respect to (no)violation of that DSU provision.

For affirmative defenses or exceptions, the provision cited is coded as an additional “Violation” finding if the panel rejected the defense or exception. If the panel accepted the defense or exception, it is coded as a “No Violation” finding. For example, if the panel found that a measure was inconsistent with GATT Article III:4, but was justified under GATT Article XX(d), GATT: III:4 is coded as a “Violation Finding” but GATT: XX(d) is coded as a “No Violation Finding.” Finally, a provision may be cited multiple times within one particular finding, but is coded only once if the subsequent references are merely elaborating the finding requested or describing the provision’s relationship with other provisions. The provision is coded more than once if an additional finding is made with respect to a different domestic measure or product at issue. For example, the following would be coded as one “Violation Finding” for GATT: XX(b): “The Panel finds that China may not seek to justify the application of export duties to forms of magnesium, manganese and zinc pursuant to Article XX(b) of the GATT1994. Even assuming *arguendo* that China could seek to justify the application of export duties, the Panel finds that China has not demonstrated that the application of export duties to forms of magnesium, manganese and zinc is justified pursuant to Article XX(b) of the GATT 1994.”

The final VALIDATION SCORE is a continuous measure that represents the proportion of explicit findings under Agreement provisions made by the panel within which it validated the trade measure, with 0 representing no validation and 1 representing complete validation.

Political Capital of the DSM

To construct the measures for the political capital of the DSM (POLCAP (CRITICISM) and POLCAP (SUPPORT)), all individual Member statements made within the DSB since the WTO’s inception were collected. Statements made on the record within DSB meetings (as transcribed within meeting minutes) are considered the verbatim official view of the member government. Official minutes of the 340 DSB meetings held between February 1995 and Decembers 2013 were obtained from *WTO Documents Online* (<https://docs.wto.org/>), from which individual Member statements were extracted. The unit of analysis for this dataset is statement made

within a DSB meeting, coded for relevant meta-data, including: Member’s name, date of statement, meeting number, whether the statement was made on behalf of a group of Members, the DSB Agenda item under which the statement was made, whether the statement was made about a specific dispute, and if so the dispute name and official WTO document number. Only statements made by Members were coded; statements made by the Chair or any other observers invited to speak have been excluded from the analysis. The dataset includes 9,883 individual statements in total.

Supervised methods of text classification were then employed to identify Members’ expressed views on the DSM. First, a coding scheme was developed to identify statements made by Members that explicitly express a view—criticism or appreciation—of the DSM and those not directly related to the WTO’s judicial bodies. While a large number of the statements are only one or two paragraphs long, some may span up to several pages. The coding scheme takes this into account by creating clear rules regarding categorization of statements that may signal both satisfaction and dissatisfaction. The original coding instrument consisted of six categories: (1) Strongly Critical; (2) Predominantly Critical; (3) Neutral; (4) Predominantly Supportive; (5) Strongly Supportive; and (6) Other/None of the Above.

Statements that fall within the *Strongly Supportive* category express strong support for the dispute settlement system as a whole, and panel and/or Appellate Body proceedings or reports. To fall within this category, the language employed in relation to the DSM must convey strong support, despite the inclusion of one or two indirectly critical comments. For example, Members may emphasize that a report is of a “high quality...setting a high standard for future panels,”⁴¹ or that the “sound legal reasoning underlying the Appellate Body’s conclusions made a significant contribution to the dispute settlement system.”⁴² This category also includes statements that exclusively express support for the DSM, even if the statement does not employ strongly supportive language.⁴³ Similarly statements that fall within the *Strongly*

⁴¹Statement by the representative of the United States, Dispute Settlement Body, Minutes of Meeting held on 17 February 1999, WT/DSB/M/55, p. 11.

⁴²Statement by the representative of Guatemala, Dispute Settlement Body, Minutes of Meeting held on 25 November 1998, WT/DSB/M/51, p. 19.

⁴³For example, Members often stress that the report under adoption was “generally well-reasoned” or that the Member “support[s] the Appellate Body’s interpretation” on a specific question, and uses the rest of the statement to recap the findings, without expressing any further view on the issue. See, for example, statement by the representative of the United States, Dispute Settlement Body, Minutes of Meeting held on 18 August 2003,

Critical category include both those that express strong criticism of the DSM despite the presence of indirectly supportive comments and exclusively critical statements. These statements often include phrases through which a Member conveys that it is “seriously concerned about the Report” or “was extremely disappointed [with the findings].”⁴⁴

Statements that fall within the *Predominantly Supportive* category express both criticism and support for the DSM, but overall—both qualitatively and quantitatively—convey greater support. For example, Members often state that they are “in general very satisfied with the Report of the Appellate Body” or that “these reports were on the whole excellent,” but that they “wished to register its concern regarding one element of the Report” or “wished to address two concerns.”⁴⁵ Such statements signal weaker support for the DSM and represent qualitatively different views than Strongly Supportive statements. Statements made in the context of adoption of an Appellate Body report that also include comments on the panel report often fall into this category. A Member may be supportive of the Appellate Body report, and in particular Appellate Body rulings that overturned panel findings with which the Member disagreed, yet still express criticism of the panel’s reasoning or findings. In such cases, the coding scheme provides that expressed views on the Appellate Body should be accorded more weight than expressed views on the panel report, because the Appellate Body is the standing judicial body of the WTO and its reports represent the final legal ruling for a particular dispute. In addition, AB reports tend to have more *de facto* precedential value than panel reports. These statements do not qualify as Strongly Supportive, however, given that the Member is also expressing dissatisfaction with the operation of the first tier of adjudication.

Similarly, statements that fall within the *Predominantly Critical* category express both criticism and support for the DSM but overall convey greater criticism.⁴⁶ Such statements include those that express criticism of the Appellate Body report but support for the panel’s findings.

WT/DSB/M/154, para. 40; Statement by the representative of Japan, Dispute Settlement Body, Minutes of Meeting held on 25 November 1998, WT/DSB/M/51, p. 19.

⁴⁴See for instance, statement by the representative of Brazil, Dispute Settlement Body, Minutes of Meeting held on 20 April 2004, WT/DSB/M/167, para. 71; Statement by the representative of Côte d’Ivoire, Dispute Settlement Body, Minutes of Meeting held on 11 December 2008, WT/DSB/M/260, para. 16-17.

⁴⁵Statement by the representative of Norway, Dispute Settlement Body, Minutes of Meeting held on 23 January 2007, WT/DSB/M/225, para. 81; Statement by the representative of the United States, Dispute Settlement Body, Minutes of Meeting held on 10 December 2003, WT/DSB/M/160, para. 8.

⁴⁶See for instance, statement by the representative of Hong Kong, China, Dispute Settlement Body, Minutes of Meeting held on 27 July 2000, WT/DSB/M/86, para. 71.

This category also includes statements that are not primarily about the DSM, or panel and/or Appellate Body proceedings, but nevertheless signal criticism of the effectiveness or operation of the system as whole. For example, Members may express frustration over the lack of implementation of DSB recommendations by another Member, which is not necessarily about the DSM itself and would be coded as Other. However, if the Member additionally highlights the systemic negative effects that long delays in implementation or continued non-compliance has for the operation of the DSM, the statement is coded as Predominantly Critical. Statements that express concern over the detrimental effects that continued non-compliance has for the interests of certain Members or groups of Members, such as developing countries, also fall within this category.

Statements classified as *Neutral* in some way reference the DSM, by mentioning the dispute or the report at hand, but without expressing criticism or support for it or the DSM.⁴⁷ Finally, the *Other* category ensures that statements not about the operation of the DSM are excluded from the analysis. This category includes statements that discuss other issues, such as the trading system in general but not the DSM in particular,⁴⁸ or refer to the actions or lack thereof of a Member, such as (non)implementation of DSB recommendations.⁴⁹ As mentioned previously, if a statement primarily about the actions of Members but also explicitly signals criticism of or support for the system as a whole or a panel/AB report, it is coded as Predominantly Critical or Predominantly Supportive.

For the training set, 527 statements were randomly selected, and two coders manually assigned each statement to a category. Random selection is preferable for supervised methods of

⁴⁷For example: “The representative of the United States said that references had been made to the reports of the panel and the Appellate Body adopted on 16 January 1998 in the case of the US-India dispute on the same matter. The United States still awaited India’s action to comply with the DSB’s recommendations concerning this issue. He therefore wished to ask India whether it could provide any information on steps to be taken towards compliance.” Statement by the representative of United States, Dispute Settlement Body, Minutes of Meeting held on 22 September 1998, WT/DSB/M/48, p. 19.

⁴⁸For example: “Multilateral trade rules could not, and should not force Members to allow a product that caused an alarming and unacceptable number of deaths to continue to be sold packaged like a sweet to attract new victims. This would undermine the credibility of the multilateral trading system. Those who depended on the WTO’s multilateral trading system felt obliged to defend it from the detrimental effect from an industry that had caused irreparable social and economic damage to countries and was an obstacle to sustainable development.” Statement by the Representative of Uruguay, WT/DSB/M/322, 28 September 2012, at para. 69.

⁴⁹For example, the following statement would fall within the Other category: “Norway would also prefer immediate implementation, but in view of the complexity of the matter some time was needed. He requested the United States to inform Members in which issue of the Federal Register the notice of the EPA would be published.” Statement by the representative of Norway, WT/DSB/M/19, 18 June 1996, at p. 3.

text analysis (Grimmer and Stewart 2013; Hopkins and King 2010) and in this case provided sufficient variation across four critical dimensions: (1) year in which statement made; (2) the Agenda Item under which the statement was made; (3) usage of the dispute settlement system by the Member making the statement (high, medium, low and non-usage of the DSM); and (4) geographical region of the Member country making the statement.

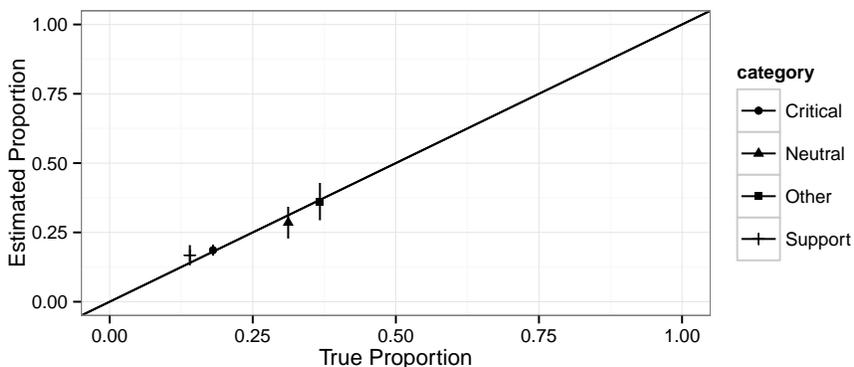


Figure A1: This graph displays the accuracy of ReadMe in recovering the distribution of critical, neutral, supportive, and other category statements within the DSB, using 10-fold cross-validation on the training set (527 statements in total). 95% confidence intervals are displayed as vertical lines. Estimates closer to the 45°line are more accurate.

Individual classification methods⁵⁰ and proportion classification methods using ReadMe, the method of automated text analysis developed in Hopkins and King (2010), were both employed. Due to initial poor model performance of individual classification methods, I collapsed the original six categories to four final classification categories: Critical, Supportive, Neutral, and Other. Because ReadMe performed better than individual classification methods, the former was used to obtain final half-year estimates for the out-of-sample statements. Figure A1 displays the accuracy of the ReadMe estimates using 10-fold cross-validation on the training set.

⁵⁰For individual classification, the RTextTools package in R was used, which provides basic preprocessing functionality, nine machine algorithms, and analytics to enable classification of large amounts of texts (Jurka et al. 2013).

Regression Results and Robustness Checks

Table A1: Summary Statistics

	N	Mean	St. Dev.	Min	Max
VALIDATION SCORE	187	0.367	0.343	0.000	1.000
POLCAP (CRITICISM)	187	0.158	0.068	0.040	0.351
POLCAP (SUPPORT)	187	0.153	0.059	0.049	0.242
AB CASE LAW	187	1.868	1.861	0.000	9.000
US/EU RESPONDENT	187	0.535	0.500	0	1
COMPLAINANT GDP (logged)	187	10.106	1.899	6.018	13.728
MEASURE TYPE	187	2.064	0.884	1	3
POLITICALLY SENSITIVE AGREEMENTS	187	0.155	0.363	0	1
TRADE REMEDIES	187	0.455	0.499	0	1
REPEAT PANELISTS	187	0.476	0.501	0	1
WTO DELEGATES	182	0.681	0.467	0	1

	Panel Validation		
	Model 1	Model 2	Model 3
POLCAP(CRITICISM)	0.098 (0.435)	0.525 (0.607)	2.140* (0.968)
POLCAP(SUPPORT)		-0.774 (1.023)	-1.271 (1.056)
AB CASE LAW			0.046 (0.058)
FINDINGS REQUESTED	0.004*** (0.001)	0.004*** (0.001)	0.003* (0.001)
US/EU RESPONDENT	0.090* (0.049)	0.100* (0.052)	0.111** (0.053)
COMPLAINANT GDP	-0.011 (0.013)	-0.012 (0.013)	-0.014 (0.013)
MEASURE TYPE (Legislation)	0.096 (0.064)	0.105 (0.068)	0.093 (0.066)
MEASURE TYPE (Admin investigation)	0.028 (0.083)	0.027 (0.083)	0.074 (0.079)
POLITICALLY SENSITIVE AGRMTS	-0.144* (0.073)	-0.144** (0.072)	-0.224*** (0.067)
TRADE REMEDIES	0.138* (0.076)	0.146* (0.078)	0.085 (0.076)
REPEAT PANELISTS (Majority)	0.011 (0.058)	0.016 (0.059)	0.045 (0.057)
WTO DELEGATES (Majority)	-0.104** (0.050)	-0.104** (0.051)	-0.061 (0.048)
POLCAP(CRITICISM)*AB CASE LAW			-0.666* (0.377)
Constant	0.213 (0.162)	0.292 (0.187)	0.224 (0.183)
Observations	182	182	182
R ²	0.231	0.236	0.325

Note: *p<0.1; **p<0.05; ***p<0.01

Table A2: Regression Results.

Ordinary least squares regression, with panel report PANEL VALIDATION as the dependent variable. Estimated coefficient values reported, with robust standard errors in parentheses. Model 2 adds POLCAP (SUPPORT) to Model 1's base estimation; Model 3 includes the interaction term between POLCAP (CRITICISM) and AB CASE LAW. All models include a cubic year trend.

	Validation Score		
	Model 2a	Model 2b	Model 2c
POLCAP(CRITICISM)	1.478** (0.608)	1.742*** (0.590)	1.707*** (0.597)
POLCAP(SUPPORT)	-0.829 (1.013)	-1.375 (0.989)	-1.316 (0.982)
US/EU RESPONDENT	0.038 (0.075)	0.109** (0.053)	0.052 (0.076)
RESPONDENT GDP (logged)	0.023 (0.020)		0.021 (0.020)
COMPLAINANT GDP (logged)	-0.011 (0.013)	-0.018 (0.014)	-0.018 (0.014)
MEASURE TYPE (Legislation)	0.104 (0.068)	0.089 (0.071)	0.088 (0.070)
MEASURE TYPE (Admin Investigation)	0.011 (0.084)	-0.0003 (0.087)	-0.015 (0.087)
POLITICALLY SENSITIVE AGRMTS	-0.160** (0.073)	-0.153** (0.070)	-0.167** (0.071)
TRADE REMEDIES	0.147* (0.078)	0.146* (0.080)	0.146* (0.079)
REPEAT PANELISTS (majority)	0.018 (0.059)	0.019 (0.060)	0.021 (0.060)
WTO DELEGATES (majority)	-0.100* (0.051)	-0.074 (0.052)	-0.070 (0.052)
Constant	0.084 (0.236)	0.393** (0.189)	0.203 (0.240)
Observations	182	173	173
R ²	0.242	0.253	0.259

Note: *p<0.1; **p<0.05; ***p<0.01

Table A3: Robustness Checks: Regression Results for Model 2

Ordinary least squares regression, with panel report VALIDATION SCORE as the dependent variable. Estimated coefficient values reported, with robust standard errors in parentheses. Model 2a includes respondent GDP in addition to US/EU RESPONDENT. Model 2b removes the ten U.S. anti-dumping zeroing cases from the initial Model 2 specification sample. Model 2c includes within Model 2b respondent GDP, in addition to US/EU RESPONDENT. All models include a cubic year trend.

	Validation Score		
	Model 3a	Model 3b	Model 3c
POLCAP(CRITICISM)	3.610*** (0.988)	3.651*** (0.999)	3.596*** (1.018)
AB CASE LAW	0.045 (0.060)	0.045 (0.062)	0.044 (0.064)
POLCAP(SUPPORT)	-1.604 (1.051)	-1.925* (1.044)	-1.862* (1.038)
US/EU RESPONDENT	0.061 (0.071)	0.108** (0.053)	0.057 (0.073)
RESPONDENT GDP (logged)	0.019 (0.020)		0.019 (0.021)
COMPLAINANT GDP (logged)	-0.014 (0.013)	-0.021 (0.013)	-0.021 (0.013)
MEASURE TYPE (Legislation)	0.092 (0.066)	0.091 (0.068)	0.090 (0.067)
MEASURE TYPE (Admin Investigation)	0.061 (0.079)	0.066 (0.084)	0.052 (0.083)
POLITICALLY SENSITIVE AGRMTS	-0.236*** (0.067)	-0.222*** (0.066)	-0.235*** (0.067)
TRADE REMEDIES	0.086 (0.075)	0.087 (0.078)	0.087 (0.077)
REPEAT PANELISTS (majority)	0.046 (0.056)	0.049 (0.056)	0.050 (0.056)
WTO DELEGATES (majority)	-0.058 (0.048)	-0.035 (0.049)	-0.032 (0.049)
POLCAP(CRITICISM)*AB CASE LAW	-0.657* (0.389)	-0.629 (0.414)	-0.620 (0.431)
Constant	0.057 (0.239)	0.313* (0.186)	0.144 (0.252)
Observations	182	173	173
R ²	0.329	0.322	0.327

Note: *p<0.1; **p<0.05; ***p<0.01

Table A4: Robustness Checks: Regression Results for Model 3

Ordinary least squares regression, with panel report VALIDATION SCORE as the dependent variable. Estimated coefficient values reported, with robust standard errors in parentheses. Model 3a includes respondent GDP in addition to US/EU RESPONDENT. Model 3b removes the ten U.S. anti-dumping zeroing cases from the initial Model 3 specification sample. Model 3c includes within Model 3b respondent GDP, in addition to US/EU RESPONDENT. All models include a cubic year trend.

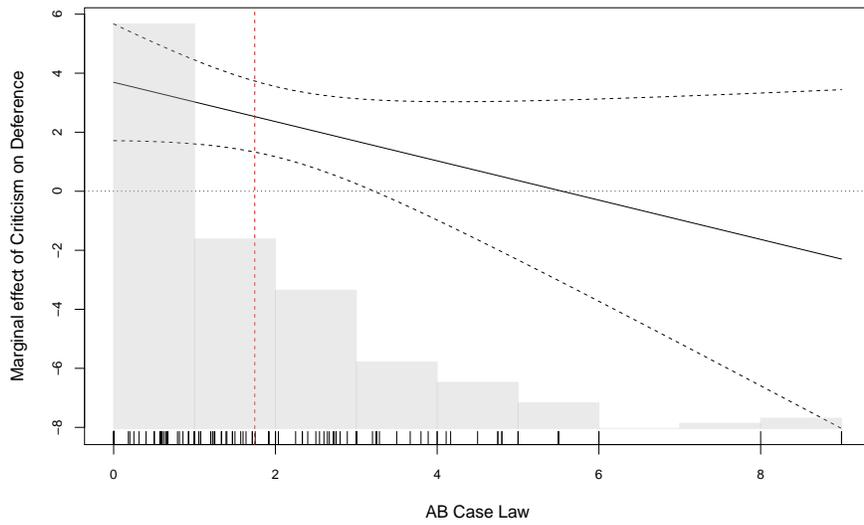


Figure A2: Relationship between Political Capital (Criticism) and Panel Validation, for different levels of Appellate Body case law (Model 3b).

References

- Alford, Robert P. 2006. "Reflections on US—Zeroing: A Study in Judicial Overreaching by the WTO Appellate Body." *Columbia Journal of Transnational Law* 45(1):196–220.
- Alter, Karen J. 2008. "Agents or Trustees? International Courts in their Political Context." *European Journal of International Relations* 14(1):33–63.
- Alter, Karen J. 2009. The European Union's Legal System and Domestic Policy: Spillover or Backlash? In *The European Court's Political Power: Selected Essays*. Oxford University Press.
- Alter, Karen J. 2011. "The Evolving International Judiciary." *Annual Review of Law and Social Science* 7:387–415.
- Alter, Karen J. 2014. *The New Terrain of International Law*. Princeton University Press.
- Barnett, Michael N. and Martha Finnemore. 2004. *Rules for the World: International Organizations in Global Politics*. Cornell University Press.
- Barnett, Michael N. and Raymond Duvall, eds. 2004. *Power in Global Governance*. Cambridge University Press.
- Becroft, Ross. 2012. *The Standard of Review in WTO Dispute Settlement: Critique and Development*. Edward Elgar Publishing Limited.
- Bhala, Raj. 1999a. "The Myth About Stare Decisis and International Trade Law: Part One of a Trilogy." *American University International Law Review* 14(4):845–956.
- Bhala, Raj. 1999b. "The Precedent Setters: De Facto Stare Decisis in WTO Adjudication: Part Two of a Trilogy." *Journal of Transnational Law and Policy* 9(1):1–151.
- Bickel, Alexander. 1962. *The Least Dangerous Branch: The Supreme Court at the Bar of Politics*. Bobbs-Merrill.

- Bohanes, Jan. 2002. "Risk Regulation in WTO Law: A Procedure-Based Approach to the Precautionary Principle." *Columbia Journal of Transnational Law* 40(2):323–389.
- Bradley, Curtis A and Judith G. Kelley. 2008. "The Concept of International Delegation." *Law and Contemporary Problems* 71(1):1–36.
- Busch, Marc and Krzysztof Pelc. 2010. "The Politics of Judicial Economy at the World Trade Organization." *International Organization* 64(2):257–279.
- Busch, Marc L. and Eric Reinhardt. 2006. "Three's a Crowd: Third Parties and WTO Dispute Settlement." *World Politics* 58(3):446–477.
- Cass, Deborah Z. 2001. "The 'Constitutionalization' of International Trade Law: Judicial Norm-Generation as the Engine of Constitutional Development in International Trade." *European Journal of International Law* 12(1):39–75.
- Christoforou, Theofanis. 2000. "Settlement of Science-Based Trade Disputes in the WTO: A Critical Review of the Developing Case Law in the Face of Scientific Uncertainty." *New York University Environmental Law Journal* 8(3):622–648.
- Cole, Wade M. 2011. "Individuals v. states: The correlates of Human Rights Committee rulings, 1979-2007." *Social Science Research* 40(3):985–1000.
- Croley, Steven and John H. Jackson. 1996. "WTO Dispute Procedures, Standard of Review, and Deference to National Governments." *American Journal of International Law* 90(2):193–213.
- Cronin, Bruce and Ian Hurd, eds. 2008. *The UN Security Council and the Politics of International Authority*. Routledge.
- Czarnezki, Jason J. 2008. "An Empirical Investigation of Judicial Decisionmaking, Statutory Interpretation, and the Chevron Doctrine in Environmental Law." *University of Colorado Law Review* 79(3):767–824.
- Danner, Allison and Erik Voeten. 2010. Who is running the international criminal justice system? In *Who Governs the Globe?*, ed. Deborah Avant, Martha Finnemore and Susan Sell. Cambridge: Cambridge University Press.

- Davey, William. 2007. Evaluating WTO Dispute Settlement: What Results Have Been Achieved through Consultations and Implementation of Panel Reports? In *The WTO in the Twenty-First Century*, ed. Yasuhei Taniguchi, Yasuhei Yanovich and Jan Bohanes. Cambridge University Press.
- Davis, Christina L. 2003. *Food Fights over Free Trade: How International Institutions Promote Agricultural Trade Liberalization*. Princeton University Press.
- Du, Michael. 2010. "Standard of Review under the SPS Agreement After EC-Hormones II." *International and Comparative Law Quarterly* 59(2):441–459.
- Eagleton-Pierce, Matthew. 2013. *Symbolic Power in the World Trade Organization*. Oxford University Press.
- Ehlermann, Claus-Dieter and Lothar Ehring. 2005. "The Authoritative Interpretation Under Article IX:2 of the Agreement Establishing the World Trade Organization: Current Law, Practice and Possible Improvements." *Journal of International Economic Law* 8(4):803–824.
- Ehlermann, Claus-Dieter and Nicolas Lockhart. 2004. "Standard of Review in WTO Law." *Journal of International Economic Law* 7(3):491–521.
- Elsig, Manfred and Mark A. Pollack. 2012. "Agents, trustees, and international courts: The politics of judicial appointment at the World Trade Organization." *European Journal of International Relations* pp. 1–25.
- Ely, John H. 1980. *Democracy and Distrust: A Theory of Judicial Review*. Harvard University Press.
- Epps, Tracey. 2008. *International Trade and Health Protection: A Critical Assessment of the WTO's SPS Agreement*. Edward Elgar Publishing Limited.
- Epstein, Lee and Jack Knight. 1998. *The Choices Justices Make*. Congressional Quarterly Press.
- Epstein, Lee and Jack Knight. 2000. "Toward a Strategic Revolution in Judicial Politics: A Look Back, A Look Ahead." *Political Research Quarterly* 53(3):625–661.

- Eskridge, William N. 1987. "Dynamic Statutory Interpretation." *University of Pennsylvania Law Review* 135(6):1479–1555.
- Footer, Mary E. 2006. *An Institutional and Normative Analysis of the World Trade Organization*. Martinus Nijhoff.
- GAO. 2003. "World Trade Organization-Standard of Review and Impact of Trade Remedy Rulings-Report to the Ranking Minority Member, Committee on Finance, US Senate."
- Garrett, Geoffrey, R. Daniel Kelemen and Heiner Schulz. 1998. "The European Court of Justice, National Governments, and Legal Integration in the European Union." *International Organization* 52(1):149–176.
- Ginsburg, Tom. 2005. "Bounded Discretion in International Judicial Lawmaking." *Virginia Journal of International Law* 45(2):631–674.
- Goldstein, Judith L. and Richard H. Steinberg. 2008. "Negotiate or Litigate? Effects of WTO Delegation on U.S. Trade Politics." *Law and Contemporary Problems* 71(1):257–282.
- Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–297.
- Guzman, Andrew T. 2004. "Food Fears: Health and Safety at the WTO." *Virginia Journal of International Law* 45(1):1–39.
- Guzman, Andrew T. 2009. "Determining the Appropriate Standard of Review in WTO Disputes." *Cornell International Law Journal* 42(1):45–76.
- Hagle, Timothy. 1993. "'Freshman Effects' for Supreme Court Justices." *American Journal of Political Science* 37(4):1142–57.
- Hathaway, Oona A. 2008. "International Delegation and State Sovereignty." *Law and Contemporary Problems* 71(1):115–150.
- Hawkins, Darren G., David A. Lake, Daniel L. Nielson and Michael J. Tierney, eds. 2006. *Delegation and Agency in International Organizations*. Cambridge University Press.

- Helfer, Laurence R. 2006. Why States Create International Tribunals: A Theory of Constrained Independence. In *International Conflict Resolution*, ed. Stefan Voigt, Max Albert and Dieter Schmidtchen. Tübingen: Mohr Siebeck.
- Helfer, Laurence R. and Karen J. Alter. 2013. “Legitimacy and Lawmaking: A Tale of Three International Courts.” *Theoretical Inquiries in Law* 14(2):479–504.
- Helmke, Gretchen. 2002. “The Logic of Strategic Defection: Judicial Decision-Making in Argentina under Dictatorship and Democracy.” *American Political Science Review* 96(2):291–230.
- Hettinger, Virginia A., Stefanie A. Lindquist and Wendy L. Martinek. 2003. “Acclimation Effects and Separate Opinion Writing on the U.S. Courts of Appeals.” *Social Science Quarterly* 84(4):792–810.
- Hooghe, Liesbet and Gary Marks. 2014. “Delegation and Pooling in International Organizations.” *Review of International Organizations* Forthcoming.
- Hopkins, Daniel and Gary King. 2010. “A Method of Automated Nonparametric Content Analysis for Social Science.” *American Journal of Political Science* 54(1):229–247.
- Horlick, Gary N. and Peggy A. Clarke. 1997. Standards for Panels Reviewing Anti-Dumping Determinations under the GATT. In *International Trade Law and the GATT and WTO Dispute Settlement System*, ed. Ernst-Ulrich Petersmann. Kluwer Law International.
- Imai, Kosuke, Gary King and Olivia Lau. 2008. “Toward A Common Framework for Statistical Analysis and Development.” *Journal of Computational and Graphical Statistics* 17(4):892–913.
- Jackson, John H. 1994. “Remarks on ‘The WTO and the New Dispute Settlement’.” *Proceedings of the Annual Meeting (American Society of International Law)* 88(131-139).
- Jackson, John H. 2001. “The WTO “Constitution” and Proposed Reforms: Seven “Mantras” Revisted.” *Journal of International Economic Law* 4(1):67–78.

- Jurka, Timothy P., Loren Collingwood, Amber E. Boydston, Emiliano Grossman and Wouter von Atteveldt. 2013. "RTextTools: A Supervised Learning Package for Text Classification." *The R Journal* 5(1):6–12.
- Kelemen, R. Daniel. 2001. "The Limits of Judicial Power: Trade-Environment Disputes in the GATT/WTO and the EU." *Comparative Political Studies* 34(6):622–650.
- Keohane, Robert O., Andrew Moravcsik and Anne-Marie Slaughter. 2000. "Legalized Dispute Resolution: Interstate and Transnational." *International Organization* 54(3):457–488.
- Koppell, Jonathan G.S. 2008. "Global Governance Organizations: Legitimacy and Authority in Conflict." *Journal of Public Administration Research and Theory* 18(2):177–203.
- Koremenos, Barbara. 2005. "Contracting around International Uncertainty." *American Political Science Review* 99(4):549–565.
- Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(4):761–799.
- Lake, David A. 2007a. "Delegating divisible sovereignty: Sweeping a conceptual minefield." *The Review of International Organizations* 2:219–237.
- Lake, David A. 2007b. "Escape from the State of Nature: Authority and Hierarchy in World Politics." *International Security* 32(1):47–79.
- Lake, David A. 2010. "Rightful Rules: Authority, Order, and the Foundations of Global Governance." *International Studies Quarterly* 54(3):587–613.
- Letsas, George. 2006. "Two Concepts of the Margin of Appreciation." *Oxford Journal of Legal Studies* 26(4):705–732.
- Lupu, Yonatan and Erik Voeten. 2012. "Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights." *British Journal of Political Science* 42(2):413–439.

- Mavroidis, Petros C. and Thomas Cottier. 2003. The Role of the Judge in International Trade Regulation: An Overview. In *The Role of the Judge in International Trade Regulation: Experience and Lessons for the WTO*, ed. Thomas Cottier and Petros C. Mavroidis. University of Michigan Press.
- Mueller-Holyst, Bozena. 2005. The role of the Dispute Settlement Body in the dispute settlement process. In *Key Issues in WTO Dispute Settlement: The First Ten Years*, ed. Rufus Yerxa and Bruce Wilson. Cambridge: Cambridge University Press.
- Oesch, Matthias. 2003. *Standards of Review in WTO Dispute Resolution*. Oxford University Press.
- Palmeter, David N. and Petros C. Mavroidis. 2004. *Dispute Settlement in the World Trade Organization: Practice and Procedures*. 2d ed. Cambridge University Press.
- Pelc, Krzysztof. 2014. "The Politics of Precedent in International Law: A Social Network Application." *American Political Science Review* 108(3):547–564.
- Peters, Birgit and Johan Karlsson Schaffer. 2013. "Introduction: The Turn to Authority beyond States." *Transnational Legal Theory* 4(3):315–335.
- Raab, Dominic and Hans Bevers. 2006. "The International Criminal Court and the Separation of Powers." *International Organization Law Review* 3.
- Romano, Cesare P.R. 2011. "A Taxonomy of International Rule of Law Institutions." *Journal of International Dispute Settlement* 2(1):241–277.
- Rosendorff, B. Peter and Helev V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55(4):829–857.
- Schauer, Frederick. 2000. "Lecture: Incentives, Reputation, and the Inglorious Determinants of Judicial Behavior." *University of Cincinnati Law Review* 68(3):615–636.
- Shany, Yuval. 2006. "Toward a General Margin of Appreciation Doctrine in International Law?" *European Journal of International Law* 16(5):907–940.

- Simmons, Beth A. 2002. "Capacity, Commitment, and Compliance: International Institutions and Territorial Disputes." *Journal of Conflict Resolution* 46(6):829–856.
- Smith, James McCall. 2003. "WTO Dispute Settlement: The Politics of Procedure in Appellate Body Rulings." *World Trade Review* 2(1):65–100.
- Spamann, Holger. 2004. "Standard of Review for World Trade Organization Panels in Trade Remedy Cases: a Critical Analysis." *Journal of World Trade* 38(3):509–555.
- Steinberg, Richard. 2004. "Judicial Lawmaking at the WTO: Discursive, Constitutional, and Political Constraints." *American Journal of International Law* 98(2):247–275.
- Stone Sweet, Alec. 1997. The New GATT: Dispute Resolution and the Judicialization of the Trade Regime. In *Law above Nations: Supranational Courts and the Legalization of Politics*, ed. Mary L. Volcansek. University of Florida Press.
- Stone Sweet, Alec and Thomas L. Brunell. 2013. "Trustee Courts and the Judicialization of International Regimes: The Politics of Majoritarian Activism." *Journal of Law and Courts* 1(1):61–88.
- Strawbridge, Jamie. 2013. "U.S. Implementation of Adverse WTO Rulings: A Closer Look at the Tuna-Dolphin, COOL, and Clove Cigarettes Cases." *ASIL Insights* 17(23).
- Sykes, Alan O. 2006. Domestic Regulation, Sovereignty, and Scientific Evidence Requirements: A Pessimistic View. In *Trade and Human Health and Safety*, ed. George A. Bermann and Petros C. Mavroidis. Cambridge University Press.
- Voeten, Erik. 2008. "The Impartiality of International Judges: Evidence from the European Court of Human Rights." *American Political Science Review* 102(4):417–433.
- Voeten, Erik. 2013. International Judicial Independence. In *Interdisciplinary Perspectives on International Law and International Relations: The State of the Art*, ed. Jeffrey Dunoff and Mark Pollack. Cambridge University Press.

- Voeten, Erik. 2014. International Judicial Behavior. In *Oxford University Handbook of International Adjudication*, ed. Cesare P.R. Romano, Karen J. Alter and Chrisanthi Avgerou. Oxford University Press.
- von Bogdandy, Armin, Rüdiger Wolfrum, Jochen von Bernstorff, Philipp Dann and Matthias Goldmann, eds. 2010. *The Exercise of Public Authority by International Institutions: Advancing International Institutional Law*. Springer.
- von Staden, Andreas. 2012. “The democratic legitimacy of judicial review beyond the state: Normative subsidiarity and judicial standards of review.” *I-CON* 10(4):1023–1049.
- Wagner, J. Martin. 2000. “The WTO’s Interpretation of the SPS Agreement Has Undermined the Right of Governments to Establish Appropriate Levels of Protection Against Risk.” *Law and Policy in International Business* 31(3):855–859.
- Weiler, J.H.H. 1994. “A Quiet Revolution: The European Court of Justice and its Interlocutors.” *Comparative Political Studies* 26(4):510–534.
- Wood, Sandra L., Linda Camp Keith, Drew Noble Lanier and Ayo Ogundele. 1998. “‘Acclimation Effects’ for Supreme Court Justices: A Cross-Validation, 1888–1940.” *American Journal of Political Science* 42(2):690–97.
- WTO. 2011. “Understanding the WTO.” Geneva, Switzerland: .
- Zürn, Michael, Martin Binder and Matthias Ecker-Eckhardt. 2012. “International authority and its politicization.” *International Theory* 4(1):69–106.