**British Journal of
Political Science**

Legalization and Compliance: How Judicial Activity Undercuts the Global Trade Regime

Jeffrey Kucik[1]* [ID], Lauren Peritz[2] and Sergio Puig[1]

[1]University of Arizona, Tucson, Arizona, USA and [2]University of California, Davis, California, USA
*Corresponding author. Email: jkucik@email.arizona.edu

Abstract

The crisis facing the World Trade Organization illustrates the trade-off between legalization and compliance in international legal systems. Dispute bodies can sometimes "overreach" in their rulings, leading to resistance from member states. This article looks at one form of legal overreach: the extension of legal precedent. We argue that extending previous decisions can reduce the flexibility that states include deliberately in their agreements. We utilize original data on individual applications of precedent in the World Trade Organization's Appellate Body decisions from 1995 to 2015 and on policy responses to those decisions. We find strong evidence that extending precedent reduces on-time compliance. It also leads to longer delays before members comply. The results speak to the life cycles of international organizations, as well as questions of design and cooperation.

The global trade regime's legal backbone was broken on December 11, 2019. The World Trade Organization's (WTO's) Appellate Body (AB) had been a significant step forward in international trade law (see, for example, Goldstein and Martin 2000; Hudec 2000). However, citing years of frustrations with the dispute system, the United States began vetoing AB member appointments. By the end of 2019, the AB had too few members to hear appeals, effectively ending the appeals process. What led to this crisis?

To many, the crisis at the WTO is an example of growing disillusionment with decades of neo-liberal policies (see Johnson 2019; Wallach 2019). However, the WTO's problems predate the recent "populist backlash." The crisis reflects deeper, structural issues within the organization. The WTO illustrates the limits of legalization in international dispute resolution. Institutions are "legalized" when states delegate authority to a system of binding, precise rules (Abbott et al. 2000; Goldstein et al. 2000). In theory, highly legalized institutions can benefit states. Governments share an interest in tying one another's hands to the law (Hathaway 2005). Yet, attempts to enforce international law can backfire. States are protective of their sovereignty and they resist delegating authority to dispute bodies with binding authority (see, for example, Downs, Rocke and Barsoom 1996).

Many international courts (ICs) reflect this balance between sovereignty and enforcement. ICs are often designed so that rulings do not set legal precedent (Cohen 2014; Verdier and Voeten 2014). Precedent means treating prior legal decisions as authoritative and obligatory. Precedent is discouraged in ICs precisely because states value flexibility in their agreements and are wary of binding rulings that may alter states' obligations (Alter 2008; Vaubel 2006).

Despite their design, many ICs rely heavily on precedent. That is because courts have incentives to apply the law consistently over time (see, for example, Helfer 1993; Lupu and Voeten 2012). Consistency increases the predictability of rulings, adds authority to those rulings, and may even induce higher levels of voluntary compliance (Franck 2006). This is especially true of bodies that decide issues on appeal, where judges wish to protect the court's legitimacy by confirming prior decisions. Given these incentives, ICs regularly invoke previous rulings—even if doing so might overstep their mandate.

We argue that ICs risk political resistance from governments when they rely too heavily on precedent. The backlash is particularly sharp when ICs *extend* precedent. Extensions constitute increases in legalization. Extensions occur when ICs apply prior decisions to areas of the law not implicated in the previous case, effectively widening the binding coverage of the ruling. In this way, extensions can expand countries' obligations and reduce the flexibility that governments prefer.

Extending precedent is one of the problems facing the WTO, which contains one of international law's most legalized dispute systems (Jackson 1997; Young 1995). Like many ICs, the WTO's dispute system proscribes precedent. Yet, two-thirds of all WTO AB reports include at least one precedent extension. These extensions expand the binding scope of WTO law, gradually refining states' obligations and removing leeway in the rules. In response, some WTO members have accused the AB of "legal overreach." They argue that the AB makes obligations stricter and more precise than states originally intended or agreed to. One way governments signal their dissatisfaction is by refusing to comply with AB decisions.

We explore the relationship between precedent and compliance using original data. We gather new data on over 5,500 precedent applications in AB reports and a uniquely detailed record of respondents' compliance with disputes from 1995 to 2015. We define compliance as tangible policy changes respondents make to conform to WTO rulings, typically by reversing or modifying trade barriers. The results show a significant, negative correlation between precedent extensions and compliance. The findings are robust to various estimation techniques and model specifications. We infer that the *content* of rulings—as distinct from a dispute's political or economic stakes—helps to determine compliance rates.

This article contributes to emerging literature on the life cycles of international organizations (Gray 2018). Research shows that some organizations induce very little cooperation because governments gradually stop investing in commitments that run counter to their interests. We demonstrate how the authority of—and faith in—ICs can deteriorate over time. At the WTO, systematic noncompliance reveals declining support and threatens the organization's future.

Secondly, our findings speak to a debate in the literature on ICs, including the European Court of Justice. Some international judges are reluctant to issue adverse rulings when there is a serious threat of backlash from the membership (Carrubba, Gabel and Hankla 2008; Garrett, Kelemen, and Schulz 1998). With respect to the WTO, we show that threat is credible: judges who turn a blind eye to the political context of their decisions may see countries ignore their rulings.

Thirdly, our findings resonate with the literature on institutional design that emphasizes the importance of flexibility. International agreements are deliberately "incomplete" (Koremenos 2005), and research demonstrates that ICs ruling on a more precise body of law can avert many potential conflicts (Johns 2015). However, we demonstrate how a court's strong reliance on precedent, which makes the law more refined and precise, can backfire. Rather than facilitating the early resolution of disputes in the shadow of the court (see, for example, Busch and Reinhardt 2000; Poletti et al. 2015), legalization may erode flexibility and expand states' obligations, driving down their willingness to comply.

Background and Theory

Current events highlight the limits of legalization. Recent resistance to the International Criminal Court, the European Union (EU), and the WTO arises partly because IC efforts to uphold

binding rulings conflict with states' interest in protecting their sovereignty and their desire for flexible obligations.

Interstate dispute systems illustrate the core tension. Research shows that dispute settlement has a variety of benefits, including monitoring state behavior, clarifying the law (Chaudoin 2014; Keohane, Moravcsik, and Slaughter 2000), and promoting agreement durability (see, for example, Gilligan, Johns and Rosendorff 2010; Rosendorff 2005). However, governments are also less likely to comply if their hands are tied too tightly. That is why agreement designers often place limits on ICs' authority (Maggi and Staiger 2011). A common limit is proscribing precedent.

Precedent in International Law

Precedent means treating previous decisions as authoritative. Precedent "do[es] more than simply resolve a dispute"; it also "creates a body of law for lower courts to apply and for the high court itself to follow in the future" (Lester and Bacchus 2019, 1). Many ICs are designed so that rulings *do not* establish precedent (Shahabuddeen 2007). Famously, Article 59 of the Statute of the International Court of Justice (ICJ) states that "[t]he decision of the Court has no binding force except between the parties and in respect of that particular case."[1]

States constrain ICs because they are wary of rigid obligations. International agreements are delicate political bargains that contain deliberately ambiguous, flexible language, which helps states overcome obstacles to compliance (Horn, Maggi and Staiger 2010; Koremenos, Lipson, and Snidal 2001). Given their desire for flexibility, governments worry that precedent gives ICs the ability to upset the political balance through decisions that increase binding commitments under the law.

Yet, ICs have strong incentives to rely on previous rulings (see, for example, Guillaume 2011). Courts value consistency and candor in the law. For example, the ICJ tends "to follow the reasoning and conclusions of earlier cases,"[2] and similar behavior is commonplace in international adjudication. Reading the law consistently over time can increase ICs' legitimacy (Franck 1988). It can also help clarify states' obligations. Since international law is so often ambiguous (Linos and Pegram 2016), holding consistent positions on disputed issues can refine the rules. Moreover, adjudicators view ambiguity as a legal problem to be solved through stricter interpretations of the rules. International judges are trained to produce consistent, coherent rulings, which means adjudicators typically view precedent as the preferred way to apply law across disputes and issue areas (Cohen 2013; Pauwelyn 2015).

Applying precedent often means *following* prior rulings—that is, directly applying a previous decision to a similar issue in a subsequent case. ICs can also go further, sometimes *extending* precedent to widen the scope of prior rulings. Through extensions, ICs apply a binding interpretation of states' obligations under the law to an issue area (or context) not identified in the prior ruling. ICs extend precedent for the reasons stated earlier: they value legal coherence that clarifies states' obligations under the law (Ginsberg 2005). In doing so, judges ensure the law covers a broader range of circumstances. Evidence shows the European Court of Justice has promoted its own authority through expansive rulings (Mattli and Slaughter 1998). We argue that extending precedent also affects respondent states' willingness to comply, as illustrated by behavior at the WTO.

Precedent at the WTO

The WTO's AB fits the criteria for a highly legalized IC: WTO members (1) delegate authority to a legal body whose (2) decisions add precision to highly specific commitments with (3) binding force (Johns 2015; Steinberg 2004). However, the WTO's founding members placed limits on

---

[1]Statute of the ICJ, art. 59, June 26, 1945, 59 Stat. 1055, 33 U.N.T.S. 933.

[2]Land and Maritime Boundary Between Cameroon and Nigeria (Cameroon v. Nigeria), Preliminary Objection, 1998 I.C.J. 275, ¶ 28, ¶ 31 (June 11).

dispute rulings. Article IX:2 of the WTO Agreement[3] and Article III.2 of the Dispute Settlement Understanding proscribe precedent (Bhala 1999). Under the rules, WTO disputes are independent events, judged on their own merit, and the rulings are binding on disputing parties only.

The dispute system's design was a deliberate decision to ensure that rulings did not alter states' rights and obligations.[4] Like many international policy domains, trade is contentious. Cooperation is vulnerable to time-inconsistency problems, wherein governments recognize the long-run benefits of multilateral trade liberalization but may feel strong domestic pressures to shirk their agreements during hard times (Davis and Pelc 2017; Goldberg and Maggi 1999). States respond to this problem by creating incomplete contracts, including flexibility provisions that provide leeway to adjust policies (Bello 1996; Milner and Rosendorff 1997). States then interpret legal provisions according to their own needs. Exploiting ambiguity in WTO rules, states use trade protection to assuage domestic stakeholders while maintaining the position that those policies are consistent with their obligations.

Unfortunately, flexibility use creates disagreement among members. The majority of WTO disputes involve the exercise of flexibility provisions, including antidumping and safeguards. In response, the WTO's dispute body prefers to reduce ambiguity in the law, closing controversial "loopholes" that states rely upon to alleviate domestic protectionist pressures. They do this by using precedent (see, for example, Kucik 2019; Pelc 2014). Panel and AB reports regularly cite past decisions, as do litigants in their legal arguments (Shaffer, Elsig, and Puig 2016). The WTO itself states that "it is very likely that the panel or the Appellate Body will repeat and follow" a persuasive interpretation of a decision.[5] In other words, previous decisions are instructive—and treated as binding.

The WTO's history reflects its reliance on precedent. In 25 years, the AB has never entirely overruled a previous decision.[6] In fact, the AB frequently *extends* previous decisions. Some 10 per cent of the WTO's total precedent applications (575 of 5,576) were extensions. These extensions have a political consequence: less or delayed compliance with rulings.

How Precedent May Affect Compliance

Extending trade law precedent is controversial. Extensions can result in "legal overreach," whereby decisions alter the rights and obligations of members under the WTO agreements.

As noted, the WTO is the product of difficult, lengthy negotiations, during which states crafted their commitments carefully (Finger and Nogués 2001). This is why the Dispute Settlement Body's (DSB's) reliance on precedent can be problematic. Extensions remove the benefits of flexible language, leaving states less room to maneuver within the rules. As Goldstein and Steinberg (2008, 269–70) note: "the [AB] has engaged repeatedly in a form of lawmaking by which it has given specific meaning to ambiguous treaty language [that had been] intentionally left vague by the negotiators." In effect, the reduction of flexibility increases *legalization*, expanding the scope of governments' obligations. Thus, extending precedent can upset the delicate political balance that sustains compliance with trade law.

---

[3]"The Ministerial Conference and the General Counsel shall have the exclusive authority to adopt interpretations of this Agreement and of the Multilateral Trade Agreement." Marrakesh Agreement Establishing the World Trade Organization, April 15, 1994, Final Act Embodying the Results of the Uruguay Round of Multilateral Trade Negotiations 9, 33 I.L.M. 1125, 1148 (1994).

[4]In the "Understanding on Rules and Procedures Governing Settlement of Disputes," WTO Agreement, Annex 2, April 15, 1994, art 3.2, 33 I.L.M. 1226 (1994), the negotiators clarified that "recommendations and rulings of the [Dispute Settlement Body] cannot add to or diminish the rights and obligations provided in the covered agreements."

[5]WTO (no date) "Legal effect of panel and Appellate Body reports and DSB recommendations and rulings." *Dispute Settlement Training Module: Chapter 7*.

[6]The AB has only overruled decisions of GATT panels.

Leaving governments less leeway to overcome domestic resistance can result in delayed compliance—or outright refusal to comply—with international rulings. For instance, states might retain a trade barrier that is beneficial to key domestic stakeholders long after the AB has determined it violates WTO rules. Respondents often try to delay compliance (Brewster 2011). Delays prolong the period of immunity during which respondents protect domestic groups while avoiding retaliation from complainant states. Beyond that, states might argue that AB decisions apply the rules incorrectly—and flout the ruling entirely. While not all states oppose the extension of precedent all the time, on average, we expect that when legal rulings chip away at the flexibility states value, those rulings will provoke backlash.

These dynamics played out in disputes like *US—Stainless Steel (Mexico)*, which the Office of the US Trade Representative (USTR) cites as a turning point in the AB's application of precedent. In that case, Mexico disputed US antidumping practices, including zeroing—a controversial methodology used to determine whether imports are being "dumped" in the importer's market. Mexico argued that previous antidumping rulings were binding, stating: "there is an expectation that panels will respect prior AB rulings on the same issues … which expressly requires panels to promote the systemic values of 'security and predictability.'"[7]

The United States pushed back, arguing that previous rulings were only suggestive, not binding. The United States stated that "[precedent] was never agreed to by the WTO Members in the Uruguay Round" and that "it raises fundamental and deeply troubling questions about the AB's role within the WTO system."[8]

The AB eventually sided with Mexico and, in so doing, extended precedent. One of the cited cases was *US—Softwood Lumber V*, which addressed whether the Agreement on Anti-dumping prohibited zeroing. In that previous case, the AB decided that the antidumping did not clearly ban the zeroing methodology. Yet, in *US—Stainless Steel (Mexico)*, the AB changed course, writing: "[General Agreement on Tariffs and Trade] 1994 has to be interpreted now in conjunction with the relevant provisions of the *Anti-Dumping Agreement*." This includes Article 2.4.2, which, the AB decided, did prohibit zeroing in both the original investigations of damages and in the "sunset reviews" that countries use to decide whether to extend antidumping measures.

This ruling constitutes an extension of precedent. In *US—Stainless Steel (Mexico)*, the AB revised a previously ambiguous decision and restricted the methods countries may use to determine whether dumping occurred. In so doing, the AB tightened the obligations and limited the flexibility the US government claimed under the original text.[9]

Prior to the *US—Stainless Steel (Mexico)* ruling, the United States generally complied with WTO decisions (69 per cent), including disputes about antidumping and steel. In this case, however, the AB ruling resulted in delayed, limited compliance. The United States missed its compliance deadline of April 30, 2009. Instead, it took until April 8, 2013, to reach a partial settlement. The delay was largely due to the content of the decision. US authorities and interest groups argued that the AB violated its own design by extending precedent. The American Iron and Steel Institute (AISI) stated: "any fundamental changes to the rights and obligations of WTO members should arise only out of negotiations and mutual agreement, *rather than out of interpretation in the course of dispute settlement* … the Department should not adopt any change to [zeroing] methodologies" (AISI comment, February 18, 2011, 3, emphasis added). The AISI urged the Department of Commerce to ignore the AB ruling because it extended WTO obligations outside proper channels of treaty renegotiation.[10] The AISI's concern is specifically about the loss of flexibility. US authorities had previously relied on ambiguity in the rules to argue

---

[7]WT/DS344/R, December 20, 2007, Annex A-1, p. A-3.

[8]Para, 10 AB-2008-1 Appellee Submission of the United States of America, *US—Stainless Steel (Mexico)*.

[9]*US—Stainless Steel (Mexico)* is also notable because it introduced the controversial "cogent reasons" approach, which reiterated the AB's commitment to precedent.

[10]See "Comments of the American Iron and Steel Institute on Regulation 0625-AA87, 75 FR 81, 533-36 (December 28, 2010)," submitted to Ronald K. Lorentzen on February 18, 2011.

that zeroing was permissible. Here, the AB had made antidumping obligations more stringent. The implication, in the AISI's view, was that AB rulings "created obligations that do not exist … and hinder the ability of the US government to adequately remedy injury to domestic producers."[11] This eventually led the Department of Commerce to refuse portions of the AB decision and the United States achieved only partial compliance.

US—Stainless Steel (Mexico) shows the content of decisions matters, not just the direction of the ruling. This is apparent when comparing rulings within policy areas. The United States faces frequent litigation over antidumping and yet it complies with 38 per cent of antidumping rulings when precedent is not extended. Conversely, in decisions that do extend precedent, the United States complies only 18 per cent of the time.

Other members express similar dissatisfaction. India felt that the AB unjustly expanded obligations under the Agreement on Textiles in *US—Wool Shirts and Blouses*. Japan and Hong Kong both worried about precedent extensions in *EC—Asbestos*. Chile, Colombia, and Australia raised concerned over the AB's ruling in *Turkey—Rice*. Even the EU, which generally has a favorable view of precedent, protested the *EC—Export Subsidies on Sugar* decision partly because the AB extended *Canada—Dairy*. Canada echoed the EU's worries.[12]

We do not argue that all members resist precedent. What these examples suggest is that precedent can elicit political resistance, helping explain respondent governments' (non)compliance with adverse legal rulings. We propose two observable implications:

Hypothesis 1 (H1): Respondents are less likely to follow stated deadlines when precedent is extended.

Hypothesis 2 (H2): Conditional on respondents ever complying, the extension of precedent should be associated with longer delays.

We acknowledge several alternative hypotheses. First, one might suspect that noncompliance is simply an artifact of the frustrations of losing a dispute. However, the loss rate for respondents at the WTO is highly consistent (over 90 per cent), whereas on-time compliance rates across the WTO are just under 50 per cent. As a result, simply losing a case cannot explain why the odds of on-time compliance are fifty–fifty.

As for the United States, it faces more litigation and more total losses than does any other member. However, the United States' loss rate is slightly lower than other countries (87 per cent), and its on-time compliance rate is close to the WTO average (48 per cent). If the United States resisted all rulings that cut against its interests, it would comply at far lower rates.[13]

Secondly, and related, compliance is not merely an artifact of a dispute's "stakes." Research shows that volume of disputed trade varies, with many cases involving very little trade (Bown and Reynolds 2015). Moreover, compliance rates are actually higher when more trade is at stake (Kucik and Peritz 2021). Therefore, while we control for economic stakes in the following analysis, it is unlikely to be the principal determinant of compliance.

In terms of political stakes, it is important to note that disputes only arise when complainants are willing to incur litigation costs. There were over 12,000 protectionist policies around the world between 2009 and 2020. However, there were only 220 WTO disputes filed. The scarcity of disputes shows how WTO members are highly selective in initiating complaints. In other words, the political stakes are always high and are unlikely to determine compliance in one instance and not another.

---

[11]See "Comments of the American Iron and Steel Institute on Regulation 0625-AA87, 75 FR 81, 533-36 (December 28, 2010)," submitted to Ronald K. Lorentzen on February 18, 2011.

[12]Descriptions of these examples are available in Online Appendix C.

[13]The United States plays a central role in any story about the WTO. In the analysis, we control for whether the respondent in a case was the United States.

Thirdly, noncompliance is not the only way states can express dissatisfaction. Governments could simply stop filing disputes or abandon a treaty. Yet, those behaviors are relatively rare and countries still rely on the WTO to adjudicate disagreements. This is consistent with what we observe in the European Court of Justice, where frustrations have not prevented countries from utilizing the court. European states sometimes choose not to abide by legal decisions but continue to support the legal system in general (Carrubba, Gabel and Hankla 2008; Garrett, Kelemen and Schulz 1998). States often choose noncompliance, rather than abandoning a system outright, as a way to signal frustration (Pauwelyn 2005).

Research Design

We collected original data to test our hypotheses, including new data on the application of precedent within AB decisions and a detailed record of compliance with each WTO ruling. The data represent, to our knowledge, the most comprehensive account of the *content* of WTO rulings,[14] as well as the behavioral responses to those decisions.[15] Our data represent a major contribution to the study of WTO dispute settlement. The sample covers rulings made between 1995 and 2015, totaling 415 disputes.[16] Those disputes include 158 rulings against respondents.[17] The data include one row per dispute *d*.

Measuring Precedent

Precedent involves the application of prior rulings to a current dispute *d*. From 1995 to 2015, 71.4 per cent of AB reports applied precedent. There were approximately 1,400 citations of previous disputes, totaling around 5,600 individual applications of precedent. For example, the AB ruling on *United States—Anti-dumping Measures on Oil Country Tubular Goods* (DS282) cited nine previous disputes,[18] for a total of 43 individual precedent applications.

We code each citation for the nature of precedent applied. There are four possibilities. When past precedent clearly applies,[19] a court may *follow* that precedent through a simple, direct application of prior readings. Alternatively, it may *narrow* precedent through a refinement of those past readings. When precedent does not clearly apply, the court can *distinguish* the current reading, explaining why precedent does not fit the case at hand, or it may *extend* precedent by applying a past reading to the analysis of an obligation in a new circumstance. In the previous example of DS282, of the forty-three applications, twenty-five followed, nine distinguished, six extended, and three narrowed previous readings.

Coding precedent requires careful attention. Decisions that follow prior rulings are generally apparent. However, applications that narrow or extend past decision requires understanding the implication of prior decisions. We determine that precedent is *extended* when a previous reading of a rule creating an obligation is applied to a different area of the law. Coders scrutinized each citation at least twice for meaning and compared decisions to increase intercoder reliability. We provide examples of the coding in Online Appendix B.

---

[14]Other work on precedent looks at the citation of networks (Pelc 2014). Our data go a step further, coding the direction of the AB's application of each citation.

[15]Busch and Reinhardt (2003) consider compliance with nine WTO disputes between the United States and Europe. Davey (2005) looks at fifty-eight disputes during the first ten years of WTO operation. Neither verifies whether the measures achieved compliance.

[16]Our sample ends at a point when: (1) disputes have had enough time to move through the process; and (2) there is enough time to observe states' decisions to comply.

[17]Only half of all disputes end with a formal ruling. There were 175 rulings in the sample period, and respondents won 17 of them. That is consistent with the overall loss rate for respondents at the WTO, which is about 90 per cent. It leaves 158 compliance decisions in the sample period.

[18]These include DS18, DS26, DS69, DS166, DS213, DS244, DS265, DS268, and DS302.

[19]In other words, the case involves issues addressed in a previous dispute.

The majority of applications *follow* a prior ruling. The AB follows its own precedent 77 per cent of the time (see Table 1). Our particular interest is in those decisions that *extend* prior rulings, which occurs in about 10 per cent of total applications. To measure the degree to which an AB ruling provokes respondent resistance, we take the logged count of the total *Extensions$_d$* in the AB report.[20] Looking now at the dispute d level, sixty-one out of the ninety AB reports that applied precedent contained extensions: that is two-thirds of the rulings took a prior interpretation of WTO treaties and expanded its application in the current dispute.

Extensions occur across a wide range of issues. Table 2 shows the frequency of total citations and extensions for the ten most disputed agreements. Customs valuations aside, precedent is extended at similar rates across issues. We might anticipate that extensions are more common in especially controversial areas of the law—for example, antidumping or subsidies. However, precedent is not extended any more frequently in those contentious areas, suggesting that the AB is not applying precedent strategically in an attempt to stress particularly controversial issues.

Measuring Compliance

We measure compliance as tangible policy reforms by the respondent after an adverse ruling. Policy change is not the only way to comply with the WTO. However, it is the best measure of whether international legal decisions shape state behavior—that is, whether governments made a deliberate choice to bring policies "into conformity."

Our compliance data adapt recent work by Peritz (forthcoming) and is discussed in Online Appendix B. We measure two outcomes. The first is whether states meet the deadline for compliance specified in the relevant DSB report. Reports identify a date, which varies case-to-case, by which a respondent is supposed to adopt the DSB's recommendations. The average window of time that respondents enjoy is 335 days (eleven months) from the AB report. *On-Time Compliance$_r$* is a dichotomous indicator of whether a respondent r implemented policy reforms before that formal deadline. Policy reforms are those resulting from a panel or, when appealed, the AB ruling. Secondly, we measure time until compliance. Many members eventually reform their policies well after the stated deadlines. *Time to Compliance$_r$* is the logged number of days from the AB report until the date that respondents implement reform.

Respondents met their compliance deadline on seventy-five of 158 occasions (48 per cent).[21] The average time to compliance is 485 days (SD = 586), or about sixteen months. Table 3 shows how compliance rates vary across the ten members who faced the most adverse rulings. The United States complies right at the WTO average, while the EU is lower (25 per cent) and countries like Korea (86 per cent) and Canada (64 per cent) meet their compliance obligations more frequently.

Respondent countries face precedent extensions at relatively consistent rates.[22] The data do not support a story in which the AB judges routinely target certain countries for extensions. Instead, WTO members are exposed to legal overreach on a case-by-case basis, which helps to drive compliance in the dispute at hand.

Control Variables

We control for a number of confounding factors. First, compliance may vary by a dispute's "stakes." Existing work shows that third-party participation proxies for the broader interest of the WTO membership (Busch and Reinhardt 2003, Johns and Pelc 2016). We include the logged

---

[20]Our coding treats all extensions as equally important. In practice, extensions on some issues may be more important than others. However, there is no exogenous way to assign a weighting scheme that quantifies "importance."

[21]We do not analyze GATT decisions because the transition from the GATT to the WTO represents a substantive break in the degree of legalization.

[22]The full sample mean is 10.7 per cent (SD = 7.25 per cent). Rates of exposure to extensions are within one standard deviation for nine of the ten most frequent respondents.

Table 1. Distribution of precedent applications by type

Form of precedent	Number of applications	Percentage of total applications	Percentage of AB reports
Follow	4,246	77	98
Extend	575	10	68
Narrow	374	7	68
Distinguish	323	6	50

Notes: This table records the frequency of precedent applications by type. The final column reports the share of AB reports containing at least one application of a given type. It should be noted that many disputes contain different kinds of applications within one report.

Table 2. Distribution of extensions by issue area

Agreement	Total citations	Extensions	Share
Customs valuation (Article VII of GATT [1994])	28	7	0.25
Safeguards	192	32	0.17
Protocol of accession	81	13	0.16
Services (General Agreement on Trade in Services)	116	15	0.13
Import licensing	123	14	0.11
Anti-dumping Agreement (Article VI of GATT [1994])	303	34	0.11
GATT (1994)	1,615	178	0.11
Subsidies and countervailing measures	403	43	0.11
Technical barriers to trade (TBT)	271	28	0.10
Agriculture	438	45	0.10

Note: This table shows the share of total precedent applications that were extended for the ten most frequently cited GATT/WTO agreements. The median is 0.11 (SD = 0.04).

Table 3. Compliance by member (ten most frequent respondents)

Respondent	Total adverse rulings	On-time compliance	Compliance rate (%)	Average days to compliance	Extension rate (%)
US	58	26	44.8	527	14
EU	28	7	25.0	379	9
Canada	11	6	54.5	739	12
China	10	7	70.0	314	10
South Korea	7	6	85.7	239	6
Argentina	6	3	50.0	177	0
Japan	6	1	16.7	1,171	5
Mexico	6	3	50.0	389	8
India	5	5	100.0	314	12
Australia	3	1	33.0	261	15
Total (WTO)	158	72	45.6	484.4	10.7

Note: This table reports on-time compliance rates by member, highlighting the ten most frequent respondents in WTO disputes. The extension rate measures the share of precedent applications a respondent country encountered that *extended*, rather than followed, narrowed, or distinguished, existing legal decisions.

count of *Third Parties*$_d$ to control for the membership's interest, as well as the political "weight" nonlitigants may bring to bear on the compliance decision.

Compliance may also vary by the amount of disputed trade, though it turns out that evidence on this is limited (Chaudoin et al. 2016). The average amount of disputed merchandise is only USD66 million (Bown and Reynolds 2015), implying that neither dispute filings nor the compliance decisions are driven mainly by trade stakes. Dispute settlement is sometimes "political theater" (Allee and Huth 2006; Davis 2012). Nevertheless, economic ties may matter. The threat of retaliation by a complainant against a respondent who fails to comply may be one factor

(Blonigen and Bown 2001).[23] We include a variable (*Trade Share_{r,c}*) that measures the share of respondent exports that go to the complainant *c*'s market. Political and economic pressure on the respondent may also be a function of relative market size since larger respondents can better absorb retaliation. We control for the share of total participant gross domestic product (GDP) accounted for by the respondent (*GDP Share_{r,c}*).

Certain issues might also be more contentious. Antidumping is especially controversial, accounting for 112 of the first 500 WTO cases. Therefore, we include a dichotomous control for whether the dispute is about *Antidumping_d*. Another way to capture the intensity of litigation around a given issue area is to consider the total number of previous decisions on disputed issues. This is distinct from precedent.[24] The number of past rulings simply captures the number of times a panel or the AB considered a particular issue. We include a measure of the logged count of times the AB issued *Past Rulings_d* on the legal issues named in dispute *d*.

Likewise, the respondent government's policy reform process could affect compliance. Some WTO rulings concern administrative measures, while others address legislative measures. The latter tend to be more difficult to reform because more actors must consent to those changes. Our regressions include a dummy variable for rulings on legislative measures.

In terms of political stakes, some trade disputes garner more attention than others (Shaffer 2003). For example, private sector involvement raises the political stakes of the dispute, potentially prolonging the standoff between litigant governments. Ryu and Stone (2018) provide data on the number of stakeholder firms involved on each side of the first 415 WTO disputes, that is, *Firms (Compl.)* and *Firms (Resp.)*. For the subset of disputes in which the United States is a litigant, they also measure annual firm lobbying expenditures (in millions of US dollars), that is, *Lobbying_{c,t}* and *Lobbying_{r,t}*, respectively. Given the difficulties measuring private sector influence, we believe these data are the best measures available to date.[25] Our robustness checks use additional controls, as discussed in Online Appendix A.

Analysis and Results

Our analysis first addresses H1, which anticipates a negative correlation between precedent extensions and on-time compliance. We then address H2, which posits a positive correlation between precedent extensions and delays in compliance.

On-Time Compliance

Our baseline analysis uses our measure of on-time compliance. We estimate a linear odds regression (Model 1) and a logistic regression (Model 2), with standard errors clustered by "repeated filings."[26] As the substantive results are very similar, we rely on linear models in Table 4 for their ease of interpretation. Our results are also consistent when clustering standard errors by respondent or year (see Table A2 in the Online Appendix). The full logistic model results are in the Online Appendix (see Table A3).

Table 4 reports the estimates from regressing on-time compliance on precedent extension and provides consistent support for H1. Model 1 provides a reasonable fit to the data.[27] The controls

---

[23]Controlling for disputed trade also significantly reduces our sample since about one-third of all cases are "nonmerchandise" disputes. These disputes are important to include in the sample because they often involve highly contentious policies.

[24]Our measures of precedent and the accumulation of past rulings are uncorrelated (0.149).

[25]Ryu and Stone (2018) match Fortune Global 500 firms to WTO disputes in which they have a stake and to their political contributions using public disclosure data.

[26]It is common for WTO members to file separate, simultaneous complaints against a respondent concerning the same measure or practice. In practice, these grouped disputes are combined into a single legal proceeding and subsequent compliance decisions are closely related.

[27]$F(7, 124) = 3.08$, $p < 0.005$.

Table 4. On-time compliance

	1	2	3	4	5	6	7	8	
			Country-year FEs					Selection: appeals	Outcome: on-time compliance
	OLS	Logit		OLS	OLS	OLS	OLS		
Extensions[a]	−0.14*	−0.65*	−0.22*			−0.14*	−0.13*		−0.15*
	(0.06)	(0.30)	(0.09)			(0.06)	(0.06)		(0.06)
Follows[a]				−0.05					
				(0.03)					
Narrows[a]					−0.06				
					(0.08)				
Third Parties[a]	−0.11	−0.54	−0.00	−0.12	−0.12	−0.09	−0.11	0.14	−0.07
	(0.07)	(0.35)	(0.17)	(0.07)	(0.07)	(0.07)	(0.07)	(0.21)	(0.07)
Trade Share	−0.03	−0.14	−0.09	−0.02	−0.02	−0.03	−0.03	0.04	−0.03
	(0.02)	(0.09)	(0.07)	(0.02)	(0.02)	(0.02)	(0.02)	(0.08)	(0.02)
GDP Share	−0.44**	−2.14**	−0.45	−0.43**	−0.41**	−0.41**	−0.40*	−0.17	−0.44**
	(0.15)	(0.79)	(0.51)	(0.14)	(0.14)	(0.14)	(0.15)	(0.53)	(0.16)
US Respondent	0.26*	1.30*		0.23*	0.21*	0.23*	0.25	0.44	0.27*
	(0.13)	(0.65)		(0.10)	(0.10)	(0.11)	(0.13)	(0.41)	(0.11)
AD Dispute	−0.26*	−1.27*	−0.37*	−0.25*	−0.24*	−0.27**	−0.28*	−0.60	−0.45**
	(0.11)	(0.57)	(0.15)	(0.10)	(0.10)	(0.10)	(0.11)	(0.38)	(0.13)
Past Rulings[a]	0.11**	0.55**	0.05	0.11**	0.11**	0.11**	0.12**	−0.02	0.12**
	(0.04)	(0.19)	(0.14)	(0.03)	(0.03)	(0.04)	(0.03)	(0.12)	(0.04)
Leg. Measure	−0.08	−0.40	−0.17	−0.08	−0.09	−0.09	−0.09		−0.17
	(0.09)	(0.44)	(0.24)	(0.08)	(0.08)	(0.08)	(0.09)		(0.09)
GATT Dispute						−0.07			
						(0.13)			
Subsidy Dispute						0.03			
						(0.08)			
Agric. Dispute						−0.13			
						(0.12)			
Firms (Compl.)							0.02		
							(0.02)		
Firms (Resp.)							−0.02		
							(0.02)		
Appeal Share								4.29**	
								(0.46)	
Systemic Issues								0.52	
								(0.33)	
Article XXII								−0.04	
								(0.37)	
Constant	1.06**	2.79*	1.05**	1.03**	1.02**	1.10**	1.05**	−3.19**	1.07**
	(0.22)	(1.19)	(0.22)	(0.22)	(0.22)	(0.24)	(0.24)	(0.93)	(0.29)
Observations	158	158	158	158	158	158	158	392	107
R^2	0.16	–	0.15	0.15	0.14	0.17	0.17	0.28	
Inv. Mills ratio								0.24* (0.10)	

Notes: [a] Log units. Robust standard errors in parentheses. AD = antidumping; OLS = ordinary least squares; FEs = fixed effects. ** $p < 0.01$; * $p < 0.05$.

behave as expected, with lower compliance in antidumping cases and larger markets. We report the substantive effects across the interquartile range (IQR) of extensions. The point prediction at $Extension_d = 0$ is 0.51 (0.41, 0.62).[28] At $Extension_d = 2$, which is the equivalent of roughly seven extensions of precedent, the prediction is 0.23 (0.03, 0.43). More extensions of precedent cut the compliance rate in half. In the logistic regression Model 2, the corresponding point predictions are 0.49 and 0.29, respectively, meaning that more extensions cut the compliance rate by over two-fifths.[29]

---

[28]Parentheses include the 95 per cent confidence intervals.
[29]We calculate predicted effects holding all other variables at their means.

The relationship holds when using maximum likelihood estimation (Model 2) and when including respondent-year fixed effects (Model 3), which control for unobserved factors that might drive compliance decisions. Models 4 and 5 confirm that precedent *extensions* have a uniquely strong correlation with noncompliance. Model 4 looks at the logged count of decisions that *follow* precedent, which is negatively signed but misses significance at the 5 per cent level. Model 5 looks at the *narrowing*—that is, limiting the scope—of past decisions. Narrowing is insignificant. These tests demonstrate that extensions are distinct from other forms of precedent.

Model 6 in Table 4 considers the importance of issue area. In addition to our control for anti-dumping, it includes separate dichotomous indicators for whether the dispute involves the most frequently litigated issues: agriculture, GATT (1994), subsidies, and technical barriers to trade. One possibility is that extensions are more likely in controversial areas of the law, particularly if the AB is trying to reaffirm a past stance with which members already failed to comply. The core results hold when controlling for additional issues (and Online Appendix D shows that precedent extensions are uncorrelated with issue area). Model 7 shows that the number of stakeholder firms involved in the dispute does not predict compliance.

Other robustness checks are included in the Online Appendix (see Tables A3 and A4). We control for: the total amount of disputed trade; a respondent's past compliance behavior; and the outcomes of previous disputes between the litigants in dispute d.[30] As not all potential confounding variables are observable, we also conducted a sensitivity analysis to quantify the extent of omitted variable bias that would be required to undermine our substantive conclusions (see Online Appendix A8). Our findings remain robust.

As with any study of legal disputes, there are potential drivers of nonrandom selection. It should be noted that selection into disputes is not a concern for our inquiry. Our goal is explaining whether given the *existence of a ruling*, respondents comply. We do not compare (non)compliant behavior to instances where there was no formal dispute.[31] Nor do we compare our results to instances in which there was no legal ruling.[32] Instead, we focus on selection into *appeals*, the stage in which extensions may occur, thereby altering governments' willingness to comply.

We utilized a Heckman model to address selection into AB rulings. It is possible that particularly contentious disputes—that is, those where compliance may be lower, all other things being equal—are more likely to wind up before the AB. We implement the Heckman correction where the selection equation uses a dichotomous indicator of whether there was an *Appeal$_d$* and the outcome equation is our measure of on-time compliance. We rely on several identifiers. First, we code the percentage of times that the panel chair in dispute d had previous decisions appealed. Some panelists may have their decisions appealed more frequently than others. However, this should not directly influence the compliance decision because litigants have influence over who serve as panelists (Brutger and Morse 2015). If anything, litigants may prefer panelists with favorable records, implying a "bias" *toward* compliance. We also include whether the dispute was filed under Article XXII, which shapes whether members can join as third parties, thereby influencing the likelihood that a dispute reaches an early settlement (see, for example, Johns and Pelc 2014). Whether the dispute addressed systemic issues concerning the broader membership is also a predictor.[33] Model 8 in Table 4 reports the Heckman estimates, and our baseline correlation holds.

---

[30]If a chief executive can comply with a ruling without legislative approval, we would expect higher compliance rates because this effectively bypasses legislatures, which, on average, are more prone to capture by special interests (Peritz forthcoming; Rickard 2010).

[31]For a thorough investigation of this selection process, see Davis (2012).

[32]A description of these alternative stages of selection is available in the Online Appendix materials. However, given that we confine our inferences to cases with rulings, they are unlikely to bias our estimates.

[33]Disputed issue area is not correlated with the probability of appeal.

Across these models, there is a strong, negative correlation between extended precedent and on-time compliance.

Compliance Delay

The second hypothesis is that governments delay their compliance in response to extended precedent. The linear models in this section use as the dependent variable a logged number of days from the request for consultations until compliance. The sample size is smaller because we only look at cases where compliance occurred.[34]

Model 9 in Table 5 reestimates our baseline model with the delay variable. It shows a positive, significant correlation between extended precedent and delayed compliance. Across the IQR, this corresponds to over a one-year difference in time until compliance. At the twenty-fifth percentile ($Extensions_d = 0$), the predicted length to compliance is 181 days.[35] At the seventy-fifth percentile ($Extensions_d = 2$), the predicted length is 612 days.[36] This predicted 413-day (fourteen-month) delay provides extra time for governments to negotiate with domestic interest groups while dragging their feet on compliance. The average dispute is about two years long from the initial request for consultations to its final resolution. Extending precedent effectively makes disputes about 52 per cent longer. Model 10 shows that the results are not driven by domestic stakeholders, as proxied by the number of firms involved.

Model 11 reruns our Heckman selection correction using our delay variable in the outcome equation. Conditional on selection into an appeal, states take longer to comply with adverse rulings when those AB decisions extend precedent. Our replication materials reestimate models 2–6 with our delay variable, and those results are consistent.

We evaluate another implication of H2: that states become frustrated with the system based on the accumulation of experiences with AB rulings. To do so, we rely on Cox proportional hazard models, which offer an alternative way to estimate compliance delay. The "failure event" is compliance with the ruling. Our duration models use yearly observations, the same control variables, and robust standard errors clustered by respondent. As earlier, the "duration" is measured relative to the date in the AB report.

Models 12, 13, and 14 in Table 6 examine precedent accumulation over time, finding that extensions correlate with longer delays until compliance. Accumulation is the rolling sum of applications against each respondent (rather than just the applications in dispute d). Model 12 measures accumulation of precedent extensions in all disputes with an AB decision against the respondent. Model 13 looks at the subset of AB decisions where the United States was a litigant. In these instances, targeted firm lobbying expenditures are matched to each side of the dispute. Model 13 suggests a negative effect of extensions and shows that respondent side lobbying tends to prolong disputes, though incomplete data calls these estimates into question. Model 14 confirms that precedent *extensions* matter more than other forms.

The survival model results are merely suggestive. Duration models hinge on assumptions about: (1) proportional hazard rates; and (2) the inevitability of the terminating event. However, the WTO legal process itself generates nonproportional hazards: the DSB imposes deadlines for compliance that governments adjust to, and so the "risk" of compliance is likely to change abruptly over time. Moreover, not all WTO disputes can be resolved decisively.

For these reasons, we infer only that a respondent's history of facing precedent extensions tends to induce longer delays. This is substantively important. Governments are not only

---

[34]There is no obvious "end date" to disputes lacking compliance. Therefore, we cannot observe the days until noncompliance.

[35]The 95 per cent confidence interval in days is (109, 302). The point prediction in logged terms is 5.20 (4.69, 5.71).

[36]The 95 per cent confidence interval in days is (258, 1,394) and in logged terms is 6.42 (5.59, 7.24).

Table 5. Compliance delay

	9	10	11	
	OLS	OLS	Selection: appeals	Outcome: on-time compliance
Extensions[a]	0.55**	0.53*		0.53*
	(0.20)	(0.20)		(0.26)
Third Parties[a]	−0.05	−0.04	0.06	0.07
	(0.26)	(0.25)	(0.24)	(0.31)
Trade Share	0.28*	0.31*	0.04	0.32**
	(0.13)	(0.13)	(0.09)	(0.10)
GDP Share	1.22	1.20	−0.41	1.17
	(0.74)	(0.75)	(0.62)	(0.72)
US Respondent	−1.04*	−1.03*	0.52	−1.84*
	(0.48)	(0.48)	(0.47)	(0.53)
AD Dispute	0.66	0.59	−0.45	2.17**
	(0.54)	(0.55)	(0.44)	(0.61)
Case Law[a]	−0.31*	−0.32*	−0.02	−0.42*
	(0.14)	(0.15)	(0.14)	(0.70)
Legislative Measure	0.91*	0.90*		0.70
	(0.35)	(0.36)		(0.38)
Firms (Compl.)		−0.03		
		(0.05)		
Firms (Resp.)		0.10		
		(0.10)		
Appeal Share			4.38**	
			(0.51)	
Systemic Issues			0.35	
			(0.36)	
Article XXII			0.12	
			(0.44)	
Constant	2.04	1.93	−3.21**	2.23
	(1.35)	(1.35)	(1.06)	(1.25)
Observations	111	111	359	74
R^2	0.23	0.24	0.40	
Inv. Mills ratio			−0.83* (0.40)	

Notes: [a] Log units. Robust standard errors in parentheses. AD = antidumping; OLS = ordinary least squares. ** $p < 0.01$; * $p < 0.05$.

responding to the immediate application of precedent; they are also sensitive to a history of encountering AB overreach. The results suggest that the ill effects of legalization may spill over from one dispute to the next, souring a state's experience with the WTO system.

Role of the United States

Finally, we acknowledge that the United States is the most vocal—though not the only—critic of the AB. It also faces the most litigation. Our baseline models control for US involvement, and the results hold when using country-year fixed effects. Models 15 and 16 in Table 7 include an interaction term between *Extensions* and whether the United States is the respondent. That interaction is insignificant for both on-time compliance and compliance delay. The estimates show that the United States complies more often and at shorter delay than the WTO average when there are no precedent extensions. Extensions also have a larger impact on the United States than the rest of the members, though that estimate is not statistically significant. This is due to the mixed attitudes of the membership at large. Many members besides the United States have expressed opposition to precedent.

Certainly, members do not oppose precedent universally. The EU, which is the second most frequent respondent, has a more favorable view of precedent. However, precedent divides opinion, creating tensions within the organization. That is precisely why the AB's precedent use helps explain the differential compliance rates among countries—and why the AB fell into crisis.

Table 6. Duration: Cox proportional hazard models

	12	13	14
Accumulation of Prior Extensions[a]	−0.28**	−0.12	
	(0.08)	(0.07)	
Accumulation of Prior Precedent[a]			−0.17
			(0.11)
Third Parties[a]	−0.60**	−0.52**	−0.71**
	(0.19)	(0.16)	(0.20)
Trade Share	−0.23**	−0.13	−0.22**
	(0.08)	(0.07)	(0.09)
GDP Share	−1.48**	0.46	−1.46**
	(0.32)	(0.80)	(0.36)
US Respondent	1.57**	0.34	1.30**
	(0.22)	(0.61)	(0.23)
AD Dispute	−1.10**	−1.35**	−1.23**
	(0.39)	(0.22)	(0.37)
Case Law[a]	0.70**	0.35**	0.72**
	(0.11)	(0.14)	(0.11)
Legislative	−0.45	−0.51	−0.50
	(0.36)	(0.65)	(0.38)
Complainant Side Lobbying[b]		0.04	
		(0.04)	
Respondent Side Lobbying[b]		0.04*	
		(0.01)	
Observations	307	130	307
Missing observations	0	62	0
R^2	0.32	0.26	0.30
Disputes w/appeal	98	60	98
Disputes w/appeal and compliance	79	51	79
Sample	All AB	US litigant	All AB

Notes: [a] Log units. [b] Millions USD. Robust standard errors in parentheses. ** $p < 0.01$; * $p < 0.05$.

Table 7. Role of the United States

	15	16
	On-time compliance	Compliance delay
Extensions × US Respondent	−0.15	0.11
	(0.12)	(0.43)
Extensions[a]	−0.06	0.49
	(0.09)	(0.25)
US Respondent	0.33*	−1.09
	(0.14)	(0.57)
Third Parties[a]	−0.10	−0.06
	(0.07)	(0.25)
Trade Share	−0.03	0.28*
	(0.02)	(0.13)
GDP Share	−0.44**	1.22
	(0.15)	(0.75)
AD Dispute	−0.26*	0.65
	(0.11)	(0.54)
Case Law[a]	0.11**	−0.30*
	(0.04)	(0.15)
Legislative Measure	−0.06	0.90*
	(0.09)	(0.37)
Constant	1.03**	2.07
	(0.24)	(1.38)
Observations	158	111
R^2	0.17	0.23

Notes: [a] Log units. AD = antidumping. Robust standard errors in parentheses. ** $p < 0.01$; * $p < 0.05$.

Conclusions and Implications

The WTO began life as an ambitious step forward in the legalization of international economic relations. Now, the AB is accused of overreach by adopting a strong reliance on precedent that often extends the scope of prior rulings. This exceeds, in the view of some members, the AB's mandate. The result is lower, and slower, compliance.

The results support an emerging literature on international organization life cycles. International adjudicatory bodies are designed with specific limits on their authority. Proscribing precedent is one such form of control. However, ICs can sometimes drift from their original mandates and push the boundaries of these limits. The recent pushback against legalization in a variety of contexts shows that states guard their sovereignty closely and resist perceived instances of overreach by adjudicatory bodies. Evolution can upset the delicate balance states strike when delegating authority to international organizations.

Our study brings novel data to bear on the dynamics of adjudicatory overreach and compliance, which remains a central concern in the study of cooperation. Our findings demonstrate how stricter enforcement of trade law can backfire (Downs, Rocke, and Barsoom 1996). That is what we see at the WTO. Strong reliance on precedent increased the legalization of the dispute system and restricted the policy leeway that members guard so preciously. When precedent *deepens* states' obligations, whether or not it increases the precision of the law, there are clear costs to legalization (Johns 2015).

Finally, there is a temptation to believe that the WTO's is a *sui generis* story driven by the United States. Given public testimony from other WTO members, the dissatisfaction with precedent is more general. Nevertheless, the United States has more to lose from the WTO's judges claiming the authority to "make law." From the US point of view, this AB practice is especially harmful when a state needs flexibility to support industries affected by the trade practices of lower-cost producer states. US opposition has shone a spotlight on a legal regime whose judicial activities have crept beyond their mandate and, in doing so, incentivized countries to delay or avoid compliance, tacitly undercutting the regime. Our findings point to the need for reform to the WTO dispute settlement system—reforms that might institutionalize and reintroduce the flexibility that is lost through precedent extensions.

References

Abbott KW et al. (2000) The concept of legalization. *International Organization* **54**(3), 401–419.

Allee TL and Huth PK (2006) Legitimizing dispute settlement: international legal rulings as domestic political cover. *American Political Science Review* **100**(2), 219–234.

Alter KJ (2008) Delegating to international courts: self-binding vs. other-binding delegation. *Law and Contemporary Problems* **71**(1), 37–76.

Bello JH (1996) The WTO dispute settlement understanding: less is more. *American Journal of International Law* **90**(3), 416–418.

Bhala R (1999) The myth about stare decisis and international trade law (part one of a trilogy). *American University of International Law Review* **14**(4), 845–956.

Blonigen BA and Bown CP (2001) Antidumping and retaliation threats. *Journal of International Economics* **60**(2), 249–273.

Bown C and Reynolds KM (2015) Trade flows and trade disputes. *Review of International Organizations* **10**, 145–177.

Brewster R (2011) The remedy gap: institutional design, retaliation and trade law enforcement. *George Washington Law Review* **80**(1), 102–158.

Brutger R and Morse JC (2015) Balancing law and politics: judicial incentives in WTO dispute settlement. *The Review of International Organizations* **10**(2), 179–205.

Busch ML and Reinhardt E (2000) Bargaining in the shadow of the law: early settlement in GATT/WTO disputes. *Fordham International Law Journal* **24**, 158–172.

Busch ML and Reinhardt E (2003) The evolution of GATT/WTO dispute settlement. In Curtis JM and Ciuriak D (eds), *Trade Policy Research*. Ottawa: Department of Foreign Affairs and International Trade, 143–183.

Carrubba C, Gabel M, and Hankla C (2008) Judicial behavior under political constraints: evidence from the European Court of Justice. *American Political Science Review* **102**(4), 435–452.

Chaudoin S (2014) Audience features and the strategic timing of trade disputes. *International Organization* **68**(4), 877–911.

Chaudoin S, Kucik J, and Pelc K (2016) Do WTO disputes actually increase trade? *International Studies Quarterly* **60**(2), 294–306.

Cohen HG (2013) Lawyers and precedent. *Vanderbilt Journal of Transnational Law* **46**(4), 1025–1040.

Cohen HG (2014) Theorizing precedent in international law. In Bianchi A, Peat D and Windsor M (eds), *Interpretation in International Law*. Oxford: Oxford University Press, 268–289.

Davey W (2005) The WTO dispute settlement system: the first ten years. *Journal of International Economic Law* **8**(1), 17–50.

Davis CL (2012) *Why Adjudicate? Enforcing Trade Rules in the WTO*. Princeton, NJ: Princeton University Press.

Davis CL and Pelc KJ (2017) Cooperation in hard times: self-restraint of trade protection. *Journal of Conflict Resolution* **61**(2), 398–429.

Downs GW, Rocke DM, and Barsoom PN (1996) Is the good news about compliance good news about cooperation? *International Organization* **3**, 379–406.

Finger JM and Nogués JJ (2001) *The Unbalanced Uruguay Round Outcome: The New Areas in Future WTO Negotiations*. Washington, DC: The World Bank.

Franck TM (1988) Legitimacy in the international system. *The American Journal of International Law* **82**(4), 705–759.

Franck TM (2006) The power of legitimacy and the legitimacy of power: international law in an age of power disequilibrium. *American Journal of International Law* **100**(1), 88–106.

Garrett GR, Kelemen D and Schulz H (1998) The European Court of Justice, national governments, and legal integration in the European Union. *International Organization* **52**(1), 149–176.

Gilligan M, Johns L and Rosendorff BP (2010) Strengthening international courts and the early settlement of disputes. *Journal of Conflict Resolution* **54**(1), 5–38.

Ginsberg T (2005) Bounded discretion in international judicial lawmaking. *Virginia Journal of International Law* **45**(3), 631–674.

Goldberg P and Maggi G (1999) Protection for sale: an empirical investigation. *American Economic Review* **89**(5), 1135–1155.

Goldstein J and Martin LL (2000) Legalization, trade liberalization, and domestic politics: a cautionary note. *International Organization* **54**(3), 603–632.

Goldstein J and Steinberg R (2008) Negotiate or litigate? Effects of WTO judicial delegation on US trade politics. *Law & Contemporary Problems* **71**, 257–282.

Goldstein J et al. (2000) Introduction: legalization and world politics. *International Organization* **54**(3), 385–399.

Gray J (2018) Life, death, or zombie? The vitality of international organizations. *International Studies Quarterly* **62**(1), 1–13.

Guillaume G (2011) The use of precedent by international judges and arbitrators. *Journal of International Dispute Settlement* **2**(1), 5–23.

Hathaway OA (2005) Between power and principle: an integrated theory of international law. *The University of Chicago Law Review* **72**(2), 469–536.

Helfer LR (1993) Consensus, coherence and the European Convention on Human Rights. *Cornell International Law Journal* **26**(1), 133–166.

Horn H, Maggi G and Staiger RW (2010) Trade agreements as endogenously incomplete contracts. *American Economic Review* **100**(1), 394–419.

Hudec RE (2000) Broadening the scope of remedies in WTO dispute settlement. In F Weiss (ed.), *Improving WTO Dispute Settlement Procedures*. London: Cameron May Publishing.

Jackson JH (1997) *The World Trading System: Law and Policy of International Economic Relations*, 2nd edn. Cambridge: MIT Press.

Johns L (2015) *Strengthening International Courts: The Hidden Costs of Legalization*. Ann Arbor, MI: University of Michigan Press.

Johns L and Pelc K (2014) Who gets to be in the room? Manipulating participation in WTO disputes. *International Organization* **68**(3), 663–699.

Johns L and Pelc K (2016) Fear of crowds in World Trade Organization disputes: why don't more countries participate? *The Journal of Politics* **78**(1), 88–104.

Johnson K (2019) How Trump may finally kill the WTO. *Foreign Policy*, December 9.

Keohane RO, Moravcsik A and Slaughter A-M (2000) Legalized dispute resolution: interstate and transnational. *International Organization* **54**(3), 457–488.

Koremenos B (2005) Contracting around international uncertainty. *American Political Science Review* **99**(4), 549–565.

Koremenos B, Lipson C and Snidal D (2001) The rational design of international institutions. *International Organization* **55**(4), 761–799.

Kucik J (2019) How do prior rulings affect future disputes. *International Studies Quarterly* **63**(4), 1122–1132.

Kucik J and Peritz L (2021) How do third parties affect compliance in the trade regime?. *The Journal of Politics* **83**(3), 1184–1189.

Kucik J, Peritz L, Puig S, (2022) Replication data for: Legalization and Compliance: How Judicial Activity Undercuts the Global Trade Regime, https://doi.org/10.7910/DVN/11V2P6, Harvard Dataverse, VI.

Lester S and Bacchus J (2019) Of precedent and persuasion: the crucial role of an appeals court in WTO disputes. *Free Trade Bulletin* **74**, 1–6.

Linos K and Pegram T (2016) The language of compromise in international agreements. *International Organization* **70**(3), 587–621.

Lupu Y and Voeten E (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.

Maggi G and Staiger RW (2011) The role of dispute settlement procedures in international agreements. *The Quarterly Journal of Economics* **126**(1), 475–515.

Mattli W and Slaughter A-M (1998) Revisiting the European Court of Justice. *International Organization* **52**(1), 177–209.

Milner HV and Rosendorff BP (1997) Democratic politics and international trade negotiations: elections and divided government as constraints on trade liberalization. *Journal of Conflict Resolution* **41**(1), 117–146.

Pauwelyn J (2005) The transformation of world trade. *Michigan Law Review* **104**(1), 1–66.

Pauwelyn J (2015) The rule of law without the rule of lawyers? Why investment arbitrators are from Mars, trade adjudicators from Venus. *American Journal of International Law* **109**(4), 761–805.

Pelc K (2014) The politics of precedent in international law: a social network application. *The American Political Science Review* **108**(3), 547–564.

Peritz L (forthcoming) *Delivering on Promises: The Domestic Politics of Compliance in International Courts*. Chicago, IL: University of Chicago Press.

Poletti A, De Bievre D and Chatagnier J (2015) Cooperation in the shadow of WTO law: Why litigate when you can negotiate. *World Trade Review* **14**(S1), S33–S58.

Rickard SJ (2010) Democratic differences: electoral institutions and compliance with GATT/WTO agreements. *European Journal of International Relations* **16**(4), 711–729.

Rosendorff BP (2005) Stability and rigidity: politics and design of the WTO's dispute settlement procedure. *American Political Science Review* **99**(3), 389–400.

Ryu J and Stone RW (2018) Plaintiffs by proxy: a firm-level approach to WTO dispute resolution. *Review of International Organization* **13**, 273–308.

Shaffer G (2003) *Defending Interests: Public–Private Partnerships in WTO Litigation*. Washington, DC: Brookings Institution Press.

Shaffer G, Elsig M, and Puig S (2016) The extensive (but fragile) authority of the WTO Appellate Body. *Law and Contemporary Problems* **79**, 237–314.

Shahabuddeen M (2007) *Precedent in the World Court*. Cambridge: Cambridge University Press.

Steinberg RH (2004) Judicial lawmaking at the WTO: discursive, constitutional, and political constraints. *American Journal of International Law* **98**(2), 247–275.

Vaubel R (2006) Principal–agent problems in international organizations. *The Review of International Organizations* **1**, 125–138.

Verdier P-H and Voeten E (2014) Precedent, compliance, and change in customary international law: an explanatory theory. *American Journal of International Law* **108**(3), 389–434.

Wallach L (2019) The WTO is dying. What should replace it? *New York Times*, November 28.

Young MK (1995) Dispute resolution in the Uruguay Round: lawyers triumph over diplomats. *International Lawyer (ABA)* **29**(2), 389–410.