**Legalization and Compliance:**

**How Judicial Activity Undercuts the Global Trade Regime**

**APPENDIX MATERIALS**

This appendix reports technical materials noted in the manuscript. We present supplementary information about our data and we provide analyses that probe the durability of the reported results.

Our extended analysis uses a variety of different estimators and model specifications. The results are highly consistent. We find a strong correlation between precedent extensions and (non)compliance. States are more likely to miss their compliance deadlines when precedent is extended in an Appellate Body ruling. And, even if governments eventually do bring their measures into conformity, they delay compliance much longer if AB reports extend precedent.

This document is laid out as follows. In part A, we start with a table of descriptive statistics, showing the distribution of the variables used in this analysis (Table A1). Table A2 then reports alternative estimations of our baseline model (Model 1, Table 4 in the manuscript). Table A3 repeats the core analysis using logistic regression models in lieu of linear specifications. Table A4 reports a series of on-time compliance estimations with control variables focused on the economic stakes and timing of the dispute. These additional controls supplement those reported in the manuscript. Table A5 shows that the results are consistent when substituting our measure of compliance delay in for the measure of on-time compliance. Table A6 reports additional duration model estimates, which we ran specifically for the tests of compliance delay. Table A7 reruns our baseline models with a variety of additional control variables. Across the estimations, we find highly consistent results, giving us additional confidence in the validity of the estimates reported in the main manuscript. In the final portion of part A, we provide a sensitivity analysis to demonstrate the durability of our findings. All of the estimations reported in the manuscript and in this appendix are available in our replication materials.

In part B, we provide details about the coding of our precedent data and our compliance data. The precedent data were generated through a painstakingly careful legal analysis of WTO Appellate Body reports and therefore form a central contribution of the research. The compliance data were generated from an extension of Peritz (2018) and Kucik and Peritz (forthcoming), as discussed below.

Part C includes additional evidence on countries (other than the US) who exhibit political resistance to precedent in Appellate Body rulings.

Finally, Part D addresses the leading alternative hypothesis—that compliance is all about a dispute’s stakes.

**PART A.**

*A1. Descriptive Statistics*

Table A1 summarizes the descriptions of our variables. It reports the concept each variable operationalizes as well as the mean, standard deviation, minimum and maximum values. To understand the units, notes that many of the variables are dichotomous indicators, such as the dummies for disputed issue area. Many other variables, including our measures of precedent, are logged. This is to correct for the skewed nature of the data and, substantively, for the fact that not all 1-interger increases in the value of the variable are equally important. Moving from 0 to 1 in terms of whether precedent was extended may matter more than moving from 14 to 15 total extensions. See A1 below.

*A2 and A3. Alternative Baseline Estimations*

Table A2 reports additional specifications of our baseline model (Model 1, Table 4) in the manuscript. Model 1 is a linear estimation with standard errors clustered by dispute (repeated here as Model A1.

We rely on linear probability models because of their ease of interpretation. To ensure the results are not driven by model dependence we rerun the estimator using logistic regression (Model A2) and ordered logistic regression (Model A3) to ensure that the results are consistent in a maximum likelihood estimation setting. The ordered logit has an additional benefit. It allows us to break compliance down into none (0), partial (1), and full (2).

Model A4 then runs a linear estimation with an alternative approach to clustering. Everything reported in the manuscript clusters by dispute. This is because some disputes are, essentially, ‘duplicate’ filings, where multiple complainants file separately against the respondent on the same issue. These are essentially the same grievance, and outcomes and compliance behavior are going to be tightly related. However, it is also true that a lot of litigation at the WTO takes place between the same pairs of members—especially the EU and US. As a result, we run Model A4 (Table A2), which clusters the SEs by a unique ID for each litigant pair.

Finally, Model A5 restricts the sample to only those instances where a panel report was appealed. In the main analysis, we include panel rulings that were not appealed. However, this implicitly means that our precedent variables take on a value of zero in two situations: (1) when an appeal does not apply precedent, or (2) when there is no appeal. In Model A5 (Table A2), we look only at appeals to avoid this conflation.

Table A3 replicates the Table 4 analysis using logistic regression throughout. Model A6 employs a logit model and uses country-year fixed effects. Models A7 and A8 use the different forms of precedent. Model A9 controls for common issue areas under dispute. Model A10 incorporates our proxy for political stakes: the number of stakeholder firms involved on the complainant and respondent sides. The choice of linear or logistic regression does not change our substantive results.

The results are consistent across all of these specifications.

*A4. Additional Controls: Economic Stakes and Time Trend*

Table A4 includes two models that introduce additional controls. One of them is total trade between the parties to the dispute, including third parties (Model A10). The other is whether the dispute is a merchandise case (Model A11). These provide two additional proxies for the economic stakes of the dispute. When we analyze trade stakes, we find that they do not affect the baseline correlations.

In the manuscript, we stress that disputes—and compliance decisions—are unlikely to be purely about the economic stakes. This could seem counterintuitive, but existing work establishes two strong patterns in WTO dispute settlement. First, many disputes involve relatively modest amounts of trade. A full 15 percent of WTO disputes involve less than $1 million in trade per year. And, the average amount of merchandise in question is only $66 million, a relatively small sum in context of the tremendous annual volume of trade between WTO members (Bown and Reynolds 2015). Those figures imply that neither dispute filings—nor the compliance decision—are predominantly driven by trade stakes. Second, it has been widely shown that disputes are about politics, not just economics. Governments use dispute settlement to achieve multiple goals, including signaling resolve to one another and/or trying to appease domestic interests.

Models A13 and A14 consider timing. As the WTO accrues a track record of panel and appellate body rulings, there is a greater corpus of potential precedent upon which AB judges can draw. To control for precedent invocation opportunities, we include a simple count of years from 1995 to the request for consultation and to the year of the AB ruling. Neither time trend alters our results.

*A5. Parallel Analysis with Compliance Delay*

We re-estimated the key models from the main manuscript using our measure of compliance delay—i.e., the logged count of total days from the AB report until a member eventually complies. We omitted these models from the manuscript in the interest of space, including only the core specifications in Table 5.

Model A15 (Table A5) includes a control for the year the dispute was initiated, again measured as a simple count from 1995 to the request for consultations. Model A16 includes a de-meaned measure of extended precedent. Model A17 looks at precedent applications that follow past rulings while Model A18 looks at applications that narrow the scope of precedent. Model A19 includes the additional variables for the most-contested areas of the GATT/WTO Agreements—namely, GATT (1994), subsidies, and agriculture. Finally, model A20 controls for the number of firms involved on each side of the dispute.

*A6. Additional Duration Estimations*

This section reports additional results from the duration analysis. Our baseline estimations look at compliance applications per dispute. However, it is possible that today’s compliance decision is a function of the total past experience—that is, political grievance—that respondents have accumulated over time. The core virtues of the duration approach are that (1) it allows us to deploy an estimator of survival given our interest in compliance delay, and (2) it provides a useful way of capturing the accumulation of precedent over time by respondent.

Table A6 reports some additional models capturing this. Precedent accumulation is measured over time for each respondent. This allows us to evaluate whether a respondent that has been subjected to more applications of precedent in previous AB rulings is less likely to comply in a subsequent dispute. For example, this allows us to ask: if the US was subjected to five extensions of precedent in 1999 and an additional seven in 2000 while Canada was subjected to nine and 17 in the corresponding years, is the US more likely to comply with an adverse ruling in 2001 than Canada, all else equal? Again, we consider all uses of precedent and extensions in particular, expecting the latter to have a more pronounced effect in delaying compliance.

Our duration models use yearly observations, the same control variables discussed above, and robust standard errors clustered by the respondent. Table A6 presents the estimated coefficients from a series of Cox model specifications of the duration of the dispute with time to compliance in years, using annual coding of the variables. Model A21 examines time to compliance from the complainant’s request for consultations for all uses of precedent (a) and just extensions of precedent (b). Model A22 looks at the period after the panel ruling.

Across all the model specifications, any application of precedent in previous disputes against a given respondent is associated with a longer duration to compliance (or no compliance at all) in subsequent disputes. And the association between *extensions* of precedent in previous disputes and noncompliance is even stronger.

*A7. Other Control Variables*

We consider several other control variables to probe the robustness of the results in Table A7. We account for:

* whether at least one disputant was either the European Union or US (“great powers”);
* whether the dispute was over remedy measures; the country’s past record of compliance;
* whether the respondent country has a legal system featuring civil versus common law tradition;
* and finally we re-run our analysis omitting the “zeroing” disputes that have proven so incendiary.

None of these control variables affects the significance of our main findings.

We also evaluate the robustness of our baseline results with respect to each respondent country. Section 4.3 in the main manuscript evaluates the importance of the United States and demonstrates that although the country is a crucial driver, it is not the only one contributing to backlash against AB overreach. We probe the contributions of all other respondent countries. We re-run the baseline model (table 4(1)) on the subsample that leaves out disputes for a given respondent country and calculate the coefficient on precedent extension. Figure A1 plots those coefficients (and corresponding 95% confidence intervals) for each subsample. The horizontal line at zero denotes no effect. The figure shows that beyond the US, only China is an influential country. Its omission somewhat diminishes the statistical significance of precedent extension variable in the regression from the p<0.05 to p<0.10 level. None of the other countries are substantially driving the negative correlation between precedent extension and compliance.

*A8. Sensitivity Analysis*

As with any observational study, there is a risk that unobserved confounding variables could introduce significant bias to our analysis. In this part of the appendix, we conduct a sensitivity analysis to quantify the risk that omitted variable bias (OVB) could overturn our research conclusions. There are many approaches to sensitivity analysis for regression models. Political science research is recently advancing these approaches, particularly for international relations (Chaudoin, Hays, and Hicks 2018). We leverage an elegant set of sensitivity analysis tools by Cinelli and Hazlett (2020). Their approach does not require assumptions about the functional form of the treatment assignment (i.e. distribution of key explanatory variable) nor the the unobserved confounders, and allows for researchers to apply substantive knowledge to “benchmark” the degree of possible OVB.

Our analysis probes the sensitivity of our baseline regression model (table 4(1)) to unobserved confounding variables. Our coefficient estimate on precedent extensions is = -0.14 (0.06). In our tests, we ask: how severe would the OVB have to be to “break” the reported correlation between precedent extensions and on-time compliance?

Cinelli and Hazlett propose a ‘robustness value’ (RV) that provides a convenient reference point to assess the overall robustness of a regression coefficient to unobserved confounding. If the confounders’ association to the treatment (i.e. precedent **Extensions**) and the outcome (i.e. on-time **Compliance**), measured in terms of partial are both less than the RV, then such confounders cannot explain away the observed effect. Our tests uncover a **RV** of **0.17**. This means that if an unobserved confounder that is orthogonal to all the covariates we do control for were to explain more than 17% of the residual variance of both the treatment and the outcome, that confounder would be strong enough to bring the point estimate to zero. Conversely, unobserved confounders that explain less than one-sixth of the residual variance of both Extensions and Compliance would not be strong enough to overturn the substantive results.

Cinelli and Hazlett also propose a benchmarking approach in which the OVB is quantified in terms of the explanatory power of a known variable. Here, we choose our benchmark variable as the United States Respondent indicator. Our coefficient on the estimate of US Respondent is a strong predictor of on-time compliance. Our tests show that if the unobserved confounder was as strong a predictor as US Respondent, then the estimated effect of precedent Extensions could diminish to = -0.091 (0.062). If the OVB were *twice* as strong a predictor as US Respondent, then the estimated effect of precedent extensions would diminish further to = -0.038 (0.063). The OVB would need to be nearly *three times* as strong a predictor as US Respondent to be severe enough to overturn our results.

Both tests indicate that our substantive results are durable. OVB would have to be quite severe to undermine the association between precedent extensions and delayed/non-compliance. The unobserved confounders would have to supply more explanatory power than US Respondent in order to cut in half the estimated effect of Extensions.

As a final way of probing the sensitivity of our results, we use Cinelli and Hazlett’s graphical tools for a bivariate contour plot. This plot, shown in Figure A2, parameterizes the unobserved confounder in terms of partial values. The horizontal axis shows the partial R2 of the hypothetical confounder with our explanatory variable of interest: precedent Extensions. The vertical axis shows the partial R2 of the hypothetical confounder with our outcome variable: on-time Compliance.

At the origin of the plot, we have the estimated point effect of Extensions on Compliance (). Moving away from the origin, we have contour lines with different point estimates decreasing in magnitude. The red dashed line shows a hypothetical of no effect of Extensions on Compliance, i.e. . This is the contour line for which our substantive findings are “broken”. The plot also shows the benchmarking relative to the US Respondent variable (1x, 2x and 3x as strong). To break our substantive results, the unobserved confounder would have to be quite strong in its correlations with *both* Extensions and Compliance, nearly three times the strength of US Respondent. The plot reveals that the sign of the point estimate is still relatively robust to confounding with such strengths, although the magnitude would be reduced to 65% (at 1x) and 27% (at 2x) of the original estimate.

**PART B.**

*B1. Precedent Data*

In this section of the Appendix, we describe our data on precedent in Appellate Body rulings. Our data records each citation that an AB report makes to past AB decisions and codes the way a prior decision is applied.

We start with identifying all references that AB reports made to previous AB decisions using the Trade Law Guide online jurisprudence citator.[[1]](#footnote-1) The sample covers appeals made in the first 450 disputes, amounting to 138 AB reports. Our data covers 20 years of substantive AB activity, from the start of WTO operations in 1995 to the end of 2015.[[2]](#footnote-2)

Each citation is associated with a specific subject matter—such as trade quotas, national treatment-taxation, etc.—and the WTO Agreements involved. We identified over 1,400 case references in AB reports for almost 5,600 individual applications of precedent. (The average number of prior cases referenced per dispute is 15.)

We then coded each reference using the categories: **follow**, **distinguish**, **expand**, or **narrow**. We focus on cited provisions that establish obligations rather than exceptions; we do not include exceptions to existing obligations under WTO treaties. The coding was conducted by two research assistants with international trade law background. To ensure reliability, each of them mapped all citations independently. The two coding sheets were compared, and outstanding gaps were resolved. Most references to prior decisions are made to cases under the same covered agreement.[[3]](#footnote-3)

Coding precedent was a challenging task. In some cases, deviations from precedent can be nuanced, leading to confusion as to how exactly a prior interpretation is being applied. Decisions that follow prior rulings are generally apparent because the application of precedent is often accompanied by a presentation “We emphasized in that Report […] that, in examining the issue” and shelved summarily. However, the narrowing and extending of past decisions often involves a longer discussion by the AB and deeper engagement with prior decisions. In such cases, the coding requires a good understanding of the implication of the finding. To ensure the quality of the codingeach citation was scrutinized at least twice for meaning by trained coders and substantial discussions followed complicated cases. We describe each form of precedent and provide examples to illustrate our coding.

*Following and* *distinguishing*

The AB’s citation patterns show a strong norm of **following** precedent. The AB follows its own precedent roughly 77 percent of the time. It is not surprising that most decisions are followed given that ICs have an interest in consistency. Areas of consistency include TRIPS (intellectual property) and GATS (services), two relatively newer areas of trade law where clarifications of the rules are especially important.

Some AB reports formally **distinguish** an invoked precedent, but this is rare at only 6 percent of the time. Distinguishing occurs when the AB explains specifically why the rationale of a prior case *does not* apply to the case at hand. These decisions occur most often with respect to GATT (1994) and to the Agreements on Subsidies and Technical Barriers to Trade.

For example, in the case *Chile – Price Band System*,[[4]](#footnote-4) Argentina relied on a prior AB’s ruling in its challenge of Chile’s Price Band System (PBS), by which tariff rates for certain products were assessed based on whether the price fell below a lower price band or rose beyond an upper price band. Argentina brought a claim under the first sentence of Article II:1(b) challenging the PBS as “ordinary customs duties” in excess of Chile’s bound rates. The panel found against Chile and ruled that the duties imposed under the PBS are “other duties or charges” that are prohibited under Article II:1(b)’s second sentence. Before the AB, Argentina argued that because the structure of Article II:1(b) (dealing with non-discrimination with respect to charges that can be levied on imports) is similar to that of Article III:2 of the GATT 1994 (dealing with non-discrimination with respect to taxation). As a result, they claimed that the decision in *Canada – Periodicals* was relevant. [[5]](#footnote-5) In that previous, case the AB determined that the relationship between the first and second sentences of Article III:2 of the GATT 1994 was systemic. One could move from an examination of the first sentence of that article to the examination of the second sentence as the two sentences are “part of a logical continuum.” However, rejecting Argentina’s position, the AB distinguished the application of such decision in *Canada – Periodicals.* In that later ruling, the AB decided that the first and second sentences of Article II:1(b) prescribed distinct obligations with respect to the disputed measures.

*Extending and narrowing*

Most important to our study, the AB sometimes **extends** precedent. Through extensions, the AB expands the authority of those prior decisions beyond their initial confinements. The AB extends its own precedent 10 percent of the time, It does so most frequently with respect to the GATT of 1994, principally because the vast majority of disputes cite some portion of that agreement, most commonly Articles II and III (dealing with non-discrimination). There were 822 applications of GATT between 1995 and 2015, and 104 of them were extensions—or, 12 percent. Other areas of WTO law extended with some regularity include the agreements on agriculture, subsidies, and anti-dumping.

One consequential example of precedent extension comes from the anti-dumping cases on “zeroing.” This is the US practice of setting negative dumping margins to zero to cancel out foreign trade practices. In this series of disputes, the AB has steadily expounded its prohibition of zeroing, with each subsequent ruling incrementally expanding the scope of the prior ones (Vermulst and Ikenson 2007).

There are numerous other instances in which we observe the AB take an aggressively expansionary stance on the use of precedent. For example, in *US – Oil Country Tubular Goods Sunset Reviews*, the body indicated that prior reasoning and rulings should be followed, stating that “to rely on the Appellate Body’s conclusions in earlier disputes is not only appropriate, but is what would be expected from panels, especially where the issues are the same.”[[6]](#footnote-6) Following that decision, in *US –Stainless Steel (Mexico)*,the AB relied on past reports to come to the conclusion that AB reports serve as precedent “absent cogent reasons.”[[7]](#footnote-7) However, it is difficult to see how the AB came to this conclusion based on the cited reports. The AB cited *Japan – Alcoholic Beverages II* and honed in on the statement that adopted reports are “an important part of the GATT *acquis*.”[[8]](#footnote-8) Yet, that same AB report had previously also been understood to demonstrate that previous reports were “not binding” on future adjudicators.[[9]](#footnote-9)

In *US – Lamb* the AB expanded the scope of review of trade remedies investigations by WTO panels relying on a prior decision in *EC – Hormones*. The AB indicated that on Safeguards, a panel's assessment involves both a formal aspect (whether the competent authorities have evaluated "all relevant factors") and a substantive aspect (whether the competent authorities have given a reasoned and adequate explanation for all their determination).[[10]](#footnote-10) While not particularly controversial at the time, it expanded the panel’s authority to conduct a more invasive review of domestic authorities’ determinations (Horn and Mavroidis 2003).

In *US – Gambling* the AB was tasked with interpreting the GATS and their application to the United States’ ban on internet gambling.[[11]](#footnote-11) Referring to prior decision dealing with GATT, the AB found the ban of the U.S. was a quantitative restriction rather than a qualitative restriction; essentially that it acted as a “prohibited market access restriction simply because they have the effect of a prohibition on certain cross-border supplies on gambling services” (Mercurio and Krajewski 2013). In effect, the AB made no effort to reference the dividing line between market access and domestic regulation, which is crucial to the distinction between GATT and GATS—importing the application of prior interpretations of GATT.

In *Argentina – Financial Services* the AB dipped into its GATT jurisprudence to interpret the “likeness” of services and service suppliers.[[12]](#footnote-12) Specifically, it relied on *EC – Asbestos* to analogize how “likeness” is interpreted consistent between GATT and GATS,[[13]](#footnote-13) and that the “criteria for assessing ‘likeness’ traditionally employed as analytical tools in the context of trade in goods are relevant for assessing Finally,in *Argentina – Financial Services* the AB dipped into its GATT jurisprudence to interpret the “likeness” of services and service suppliers. Specifically, it relied on *EC – Asbestos* to analogize how “likeness” is interpreted consistent between GATT and GATS,[[14]](#footnote-14) and that the “criteria for assessing ‘likeness’ traditionally employed as analytical tools in the context of trade in goods are relevant for assessing the competitive relationship of services and service suppliers.”[[15]](#footnote-15) Therefore, the AB used the interpretations and reasoning set out in GATT decisions to reason by analogy to decide this GATS dispute.

The last category of precedent is **narrowing**, in which the AB dilutes the force of prior readings. AB has narrowed its precedent less frequently—7 percent of the time. This may be due to pressures of members to change course over the interpretation. For example, the scope of a finding in *Canada – Dairy*[[16]](#footnote-16) case was whittled down by the subsequent decision in *US – FSC* case.[[17]](#footnote-17) In the former ruling, essentially grants or pay-outs could constitute subsidies under the agriculture agreements.In the latter case, the AB ruled as follows: “We held, in *Canada – Milk*, that ‘export subsidies’ under the *Agreement on Agriculture* may involve, not only direct payments, but also ‘revenue foregone.’ We believe, however, that in disputes brought under the *Agreement on Agriculture*, just as in cases under Article 1.1(a)(1)(ii) of the *SCM Agreement*, it is only where a government foregoes revenues that are ‘*otherwise due*’ that a ‘subsidy’ may arise.”[[18]](#footnote-18) Effectively, this ruling modifies the decision in *Canada – Dairy* because it limits subsidies under the the Agreement on Agricultureto revenues foregone. This had the effect of limiting the breadth of WTO obligations under the subsidies rules.

*B2. Compliance Data*

Our compliance data are an extension of data from the following sources:

* Peritz, Lauren 2018 *Database on Compliance in WTO Disputes*, U.C. Davis, available: <http://www.laurenperitz.com/data---details.html>
* Peritz, Lauren. N.d. Delivering on Promises: The Domestic Politics of Compliance in International Courts (under contract with University of Chicago Press).

Our data cover the first 415 WTO disputes with adverse rulings (1995 to 2016). Adverse rulings are those legal decisions in which the respondent government violated its WTO obligations and is required to reform its trade policy. The coding includes whether the respondent government complied with the ruling, the timing of the compliance efforts, and whether those measures achieved full, partial or non compliance. With respect to timing, we differentiate between on-time compliance, delayed compliance and non-compliance.

WTO disputes end with **on-time compliance** whenever respondent governments completed their policy reforms before the established deadline—the expiration of the ‘reasonable period of time’ (RPT). The typical ruling features a 12- to 18-month window between the ruling (adoption of the panel report) and the compliance deadline. In cases that go to appeal, the reasonable period of time is counted starting at the adoption of the AB ruling. Whether the panel ruling or the AB ruling (sometimes modifying the panel report) is the final legal decision, the end date of interest is the implementation of policies that correct the violation.

WTO disputes end with **delayed compliance** whenever the respondent governments completed their policy reforms after the established deadline. Many disputes feature delayed compliance when governments have political incentives to delay or encounter hurdles from divided authority within government (e.g. more veto points obstructing policy reform).

Sometimes governments negotiate extensions to the ‘reasonable period of time’ in order to postpone the adoption of compliant policies. This effectively buys the respondent time to remain in violation of WTO rules without activating the remedies procedure. In our coding, we use the first established deadline. We did not adjust for extensions in the RPT because those constitute legalized forms of delayed compliance; compliance measures that are implement after the first RPT expires are coded as delayed compliance.

Finally, we identify disputes with **no compliance**. Governments did not reform their trade policy or practice to conform to the adverse WTO ruling. The following examples illustrate the coding.

Timely compliance:

Dispute DS 276 (US v. Canada – Grain) illustrates timely compliance. In this case, the United States sued Canada over it grain export policies. The WTO panel determined that Canada had discriminated against foreign producers, thereby breaching its trade commitments. Canada was ordered to amend the Canada Grain Act to ensure equal treatment between domestic and foreign producers. The appellate body report was adopted on Sept. 24, 2004 and the reasonable period of time to comply was set to expire Aug. 1, 2005. Canada promptly amended the legislation on May 19, 2005 (Bill C-40) to comply before the deadline. In coding this case, we were able to easily locate the legislation in question and the passage of the amendment that brought Canada into compliance.

Delayed compliance:

An example of slightly delayed compliance can be found in dispute DS 332 (EU v. Brazil -- Retreaded Tires). In this case, the EU sued Brazil over its ban on retreaded tires—which Brazil had argued was justified on the basis of health and safety provisions. The WTO ruled against Brazil, determining that the trade policy was a violation of its obligations. The AB report was adopted on Dec. 17, 2007 and the reasonable period of time to comply was set to expire Dec. 17, 2008 (i.e. 12 months). However, Brazil did not reverse its ban by that deadline. There was an executive request to the Secretate of Foreign Trade and the Ministry of Development, Industry and International Commerce (Portaria SECEX 25/2008, SECEX 24/2009, researcher’s translation from Portuguese). Then the executive sought a ruling from Brazil's highest court. Brazil’s Supreme Court ruled that the importation of retreaded tires violated the constitution and lower courts would no longer be able to issue injunctions regarding efforts to overturn the ban. This brought about compliance with the WTO ruling effective August 28, 2009.

An example of significantly delayed compliance comes from Dispute 397 (China v. EU -- Fasteners). In this case, China sued the EU over antidumping measures. The WTO Panel and Appellate Body determined that the EU had breached its trade obligations and the AB report was adopted on July 28, 2011. The RPT was set to expire Oct. 12, 2012 (14.5 month window). The Commission implemented regulation 278/2016 on February 26, 2016, repealing the anti-dumping duties imposed on Chinese iron and steel fasteners. This, along with a Regulation earlier passed by the European Parliament and Council (Regulation 476/2015 of March 11, 2015) brought the EU into full compliance. Thus the EU delayed compliance until February 26, more than three years after the deadline.

Non-compliance:

A clear example of noncompliance following a prolonged WTO legal process comes from dispute DS 312 (Indonesia v. Korea – Paper) over antidumping measures. This dispute began in 2004. Four years later, after an adverse ruling and continued resistance on the part of Korea, the WTO compliance panel finally found that the government had failed to correct the initial violations. The Korean Trade Commission simply refused to adjust its original anti-dumping decision (2003-No.23) and Indonesian officials expressed dissatisfaction with this outcome. For some time, Korea issued periodic *pro forma* updates to the WTO with no change in status. The last mention of the dispute in WTO records occurs in late 2009; at that point Indonesia dropped the dispute after Korea's persistent noncompliance. We consider this noncompliance because the activity on the dispute ceased and there was no evidence of any compliance measures. By contrast, ongoing disputes with lengthy delays continue to feature status updates from the respondent government to the WTO.

Supporting documentation for the coding discussed above is available directly from the *Database on Compliance in WTO Disputes.*

**PART C.**

*C1. The Appellate Body in Context*

Here, we provide additional background on the WTO Appellate Body to give context to its precedent extension activity. The WTO AB was created specifically to sharpen the global trade regime’s legal teeth. Under the GATT, dispute settlement hinged on positive consensus. All interested states had to consent to the legal process (and adopt the eventual panel decision). This system became untenable as global trade markets grew in size and legal complexity. In response, the WTO’s founders created a more legalized process under the DSU. This included doing away with positive consensus and creating the AB, which was, at that time, widely viewed as a significant step forward in international economic cooperation. Members agreed to delegate authority to a dispute settlement process under which panel and AB decisions were legally binding on states.

That history is important because our argument is not that states disagree with the AB’s authority *per se*. Rather, they are dissatisfied with how the AB’s authority evolved—that is, how it expanded—over time. Arguments from states about the AB’s legal “over-reach” are not based simply on losing a case. Rather, as we show in the examples below, the criticisms come from concerns that the IC has gone too far beyond its original mandate.

To the AB’s critics, precedent is, in a sense, a form of “judicial activism.” The court is expanding states’ obligations through its decisions. This fundamentally upset the delicate political bargain imbedded in the GATT/WTO Agreements. And, its ability to do so is partly because it has a certain degree of independence from states. Positive consensus no longer limits the dispute process. Instead, the AB is imbued with the power to make binding decisions.

*C2. Additional Examples of Resistance to Precedent*

With this context, it is clear that many countries have objected to the AB’s heavy reliance on precedent. The United States is not the only country to express concerns. Here we provide a brief description of other high-profile cases where members voiced dissatisfaction with the content of the AB ruling.

**DS33** *US – Wool Shirts and Blouses*

India opposed the AB’s interpretation of the burden of proof standard under Article 6 of the Agreement on Textiles and Clothing. In particular, it claimed that the AB had previously (in DS24) used a narrow interpretation, but in this case used a more expansive interpretation. In India’s view, the ruling significantly altered the “well-established” practice regarding which litigants bore the burden of proof when justifying barriers.[[19]](#footnote-19)

**DS58** *US – Shrimp*

Thailand felt as though the AB’s use of *US-Gasoline* to analyze Article XX(g) “was uncalled for” and claimed that this newfound analysis would “permit members to discriminate against products on the basis of non-product related processes and production methods,”[[20]](#footnote-20) which would invite a sharp increase in unilateral action instead of multilateral cooperation.

India expressed concern for the AB’s use of *US-Gasoline* to solidify the two-tiered approach to interpreting Article XX, stating that it “was worrisome as the AB had previously noted that each case was to be considered on its own merit . . . .”[[21]](#footnote-21) India worried that this amounted to an implicit amendment “or [an] authoritative interpretation of the existing agreement.” In essence, the AB used *US-Gasoline* to “expand the scope of Art. XX(g) well beyond the intention of its drafters.”[[22]](#footnote-22)

**DS121** *Argentina – Footwear (EC)*

Argentina expressed its concern for the AB’s expansion of precedent relating to safeguards. Here, the AB upheld the panel’s invocation of Article 3.1 of the Agreement on Safeguards to justify exceeding the reference set out in Article 7.2 DSU. Specifically, Argentina felt that the AB’s ruling in this case “contradicted the ruling in *EC-Bananas* and seemed to grant panels broad inquisitorial powers that had not been envisaged by the members when they negotiated the DSU.”[[23]](#footnote-23)

**DS135** *EC – Asbestos*

Japan expressed its worry about the possible abuse of the *Japan-Alcoholic Beverages* interpretation of GATT Art. III. Hong Kong also joined this concern “regarding the propriety of such an interpretation and was concerned about the effect that such an interpretation could have on the balance between Art. III and XX of the GATT.”[[24]](#footnote-24)

**DS265** *EC – Export Subsidies on Sugar*

The EC raised concern about the AB’s use of Article 9.2(b)(iv) of the Agreement on Agriculture. It claimed that the “AB had gone even further than in the *Canada-Dairy* case, and had redrawn the boundary between domestic support and export subsidies in a manner which could not have been anticipated by any member at the time of the conclusions of the WTO Agreement.”[[25]](#footnote-25) Canada also found it disappointing that AB “updated” *Canada-Dairy* “without further considering its flaws” and believed that the AB “had broadened the scope of a ‘cross-subsidization’ standard that could be found nowhere in the Agreement on Agriculture.”[[26]](#footnote-26)

**DS282** *US – Anti-Dumping Measures on Oil Country Tubular Goods*

Mexico and the EC lamented the use and interpretation of *US-Corrosion-Resistance Steel Sunset Review*. Mexico “argued that by the interpretation in the instant case, what used to be an obligation to determine likelihood of dumping ‘in the review’ had been changed to a right to ‘cure’ a WTO-inconsistent likelihood determination during a reasonable time of implementation.”[[27]](#footnote-27) The EC went on further to express that “the conclusion that Art. 3 ADA did ‘not normally apply’ which had been reached in the *Argentina-OCTG* case and renewed in the dispute in question had added to the earlier finding in *US-Corrosion-Resistant Steel Sunset Review* that Art. 2 did not normally apply either to sunset reviews created legal uncertainties.”[[28]](#footnote-28)

**DS334** *Turkey – Rice*

Chile, Colombia and Australia, found it troubling that the AB condemned the panel “for its refusal to follow the AB’s past ruling on zeroing.”[[29]](#footnote-29) This, the US argued, and other countries agree, “purported to create a new legal effect for AB reports, one that would appear to grant to the AB the very authority to issue authoritative interpretations of the covered agreements that was reserved by the WTO Agreement exclusively to the members.”[[30]](#footnote-30) This explicit condemnation by the AB was taken by some members to suggest an effort to transform the WTO dispute settlement system into a common law system.

**PART D.**

An alternative to our hypothesis is the possibility that compliance is a function of a dispute’s economic and political stakes. We address the purely economic stakes in our robustness checks. There, we emphasize that many disputes involve relatively little trade. Moreover, disputed trade volumes are a poor predictor of compliance behavior. Instead, disputes are largely political events. The legal implications of decisions, according to our story, are the source of political resistance to many AB rulings.

Some readers might worry that the AB chooses to extend precedent in some areas and not others. Certain areas of the law are disputed more often, implying that they are more controversial, “higher-stakes” issues. For example, there have been 135 antidumping cases, which is over 20% of all disputes. We might reasonably expect that AD disputes are (i) more likely to reach the appeal stage and (ii) more likely to see extensions in the AB rulings.

If the AB extended precedent more often in these cases, it would be problematic for our research design because extensions would be associated with (non-)compliance simply because of the controversial nature of the issue area. Our argument is different. We theorize that the AB has an interest in legal coherence and, therefore, has strong incentives to apply precedent—even extending precedent in situations where it wishes to assert the WTO’s authority. We do not argue that these incentives vary by issue area per se. Instead, the AB is trained and conditioned to apply legal interpretations consistently across disputes—i.e., issue areas. In short, it processes disputes in a way that favors candor to the GATT/WTO Agreements.

The observable implications of our story are two-fold. First, the AB extends precedent in roughly equal proportions across issue areas. Table D1 shows the distribution of disputes, by issue area, for all WTO cases (column 1). It also shows the distribution of disputes, by issue area, for disputes that reached the AB (column 2). Comparing these two columns, there is no suggestion that key areas of the law are more likely to reach the appeals stage. The two columns are very similar.

This pattern is important for our story. Our data shows that the occurrence of AB reports by issue area is proportionate to the breakdown of issue areas all WTO disputes. AB judges are not handed some issues more than others as a portion of disputed issues. They get proportionate opportunities to weigh in. Looking at AD, for example, the fact that 21% of the AB’s caseload involves antidumping is perfectly consistent with the observation that AD cases constitute 22% of all disputes. AD, a notoriously controversial area of the law, is not overrepresented in AB rulings.

Second, we looked at the rates at which precedent gets extended by issue area (Table D2). Here, again, there is no clear evidence that the most controversial areas of the law are more likely to see extensions. Some areas of the law get cited more frequently than others, but the share of applications extended are rather consistent across area. We can again look at issues like antidumping, where precedent is extended (11.2%) close to the mean for all issue areas (10.4%). There is little evidence that issue area is a strong predictor of precedent extension.

Given the data, we infer that our precedent measures are not simply proxies for the controversial nature of certain GATT/WTO rules. The distributions of AB rulings and precedent extensions, by issue area, are consistent with patterns seen across the WTO’s total caseload.

**References**

Chaudoin, Stephen, Jude Hays and Raymond Hicks. “Do we really know the WTO cures cancer?” *British Journal of Political Science* 48, no. 4 (2018): 903-928.

Cinelli, Carlos and Chad Hazlett, “Making Sense of Sensitivity: Extending Omitted Variable Bias” *Journal of the Royal Statistical Society Series B,* 82 part.1 (2020): 39-67.

Horn, Henrik, & Mavroidis, Peter. “US – Lamb: United States – Safeguard Measures on Imports of Fresh, Chilled or Frozen Lamb Meat from New Zealand and Australia: What Should be Required of a Safeguard Investigation?” *World Trade Review*, 2 no. S1 (2003):72-114.

Mercurio, Bryan C. and Markus Krajewski, *The Regulation of Services and Intellectual Property: Volume III.* New York: Routledge, 2013.

Vermulst, Edwin and Daniel Ikenson, “Zeroing Under the WTO Anti-Dumping Agreement: Where Do We Stand?” *Global Trade and Customs Journal* 2 (2007): 231-242.

**Appendix Figures and Tables**

**Figure A1:** Coefficient Plot Omitting Each Country to Evaluate Influential Respondents



*Note:* Analysis is based on model specification of Table 4(1) in main manuscript. Coefficient on precedent extension and 95% confidence intervals calculated using subsamples that omit each respondent country, one at a time. Horizontal line denotes zero (null) effect.

**Figure A2:** Sensitivity Contour Plot of Point Estimate for Regression of On-time Compliance (outcome) on Precedent Extension (treatment)



*Note:* Analysis is based on model specification of Table 4(1) in main manuscript. Benchmark variable is US Respondent.

**Table A1. Descriptive Statistics**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| VARIABLE | Concept | Mean | Std. Dev. | Min. | Max. |
|  |  |  |  |  |  |
| On-time Comply | Compliance | 0.46 | 0.50 | 0 | 1 |
| Compliance Delay | Compliance | 5.35 | 1.77 | 0 | 8.16 |
|  |  |  |  |  |  |
| Extensions | Precedent | 0.41 | 0.64 | 0 | 2.30 |
| Follow | Precedent | 1.13 | 1.26 | 0 | 3.76 |
| Narrow | Precedent | 0.30 | 0.49 | 0 | 1.94 |
|  |  |  |  |  |  |
| Third Parties | Stakes | 1.83 | 0.73 | 0 | 3.22 |
| Trade Share | Stakes / Mkt Size | 10.64 | 2.08 | 2.34 | 13.61 |
| GDP Share | Mkt Size | 0.57 | 0.36 | 0.01 | 0.99 |
| Case Law | Issue Area | 1.61 | 1.44 | 0 | 4.23 |
| US Respondent | Power | 0.37 | 0.48 | 0 | 1 |
| Leg. Measure | Type of Measure | 0.37 | 0.48 | 0 | 1 |
| Firms (Complainant) | Stakes | 1.39 | 2.25 | 0 | 9 |
| Firms (Respondent) | Stakes | 1 | 1.66 | 0 | 13 |
|  |  |  |  |  |  |
| Anti-dumping | Issue Area | 0.22 | 0.42 | 0 | 1 |
| GATT | Issue Area | 0.86 | 0.35 | 0 | 1 |
| Subsidies | Issue Area | 0.25 | 0.44 | 0 | 1 |
| Agriculture | Issue Area | 0.15 | 0.35 | 0 | 1 |
|  |  |  |  |  |  |
| Appeal Share | Appeal Selection | 0.31 | 0.42 | 0 | 1 |
| Article XXII | Appeal Selection | 0.64 | 0.48 | 0 | 1 |
| Systemic | Appeal Selection | 0.36 | 0.48 | 0 | 1 |
|  |  |  |  |  |  |

**Table A2. Alternative baseline estimations**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **(A1)** | **(A2)** | **(A3)** | **(A4)** | **(A5)** |
|  | *On-time* | *On-time* | *On-time* | *On-time* | *On-time* |
| VARIABLES | *Comply (0/1)* | *Comply (0/1)* | *Comply*  *(0/1/2)* | *Comply (0/1)* | *Comply*  *(0/1)* |
|  |  |  |  |  |  |
| *Extensions†* | -0.14\* | -0.65\* | -0.62\* | -0.14\* | -0.15\* |
|  | (0.06) | (0.30) | (0.29) | (0.06) | (0.06) |
| *Third Parties†* | -0.11 | -0.54 | -0.54 | -0.11 | -0.09 |
|  | (0.07) | (0.35) | (0.35) | (0.07) | (0.09) |
| *Trade Share* | -0.03 | -0.14 | -0.16 | -0.03 | -0.03 |
|  | (0.02) | (0.09) | (0.10) | (0.02) | (0.02) |
| *GDP Share* | -0.44\*\* | -2.14\*\* | -2.25\*\* | -0.44\*\* | -0.41\* |
|  | (0.15) | (0.79) | (0.78) | (0.14) | (0.20) |
| *US Respondent* | 0.26\* | 1.230\* | 1.32\* | 0.26\* | 0.25 |
|  | (0.13) | (0.65) | (0.65) | (0.10) | (0.14) |
| *AD Dispute* | -0.26\* | -1.27\* | -1.26\* | -0.26\*\* | -0.40\*\* |
|  | (0.11) | (0.57) | (0.57) | (0.10) | (0.12) |
| *Case Law†* | 0.11\*\* | 0.55\*\* | 0.55\*\* | 0.11\*\* | 0.13\*\* |
|  | (0.04) | (0.19) | (0.19) | (0.03) | (0.04) |
| *Leg. Measure* | -0.08 | -0.40 | -0.48 | -0.08 | -0.17 |
|  | (0.09) | (0.44) | (0.43) | (0.08) | (0.10) |
| /cut1 |  |  | -3.12\*\* |  |  |
|  |  |  | (1.21) |  |  |
| /cut2 |  |  | -3.03\* |  |  |
|  |  |  | (1.20) |  |  |
| Constant | 1.06\*\* | 2.79\* |  | 1.06\*\* | 1.06\*\* |
|  | (0.24) | (1.19) |  | (0.24) | (0.28) |
| Clustered s.e. by | Dispute | Dispute | Dispute | Litigant Pair | Dispute |
| Model | OLS | Logit | Ordered Logit | OLS | OLS |
| Sample | Full | Full | Full | Full | AB Ruling |
| Observations | 158 | 158 | 158 | 158 | 108 |
| R-squared | 0.16 |  |  | 0.16 | 0.25 |
| Robust standard errors clustered as indicated. | | | | | |
| \*\* p<0.01, \* p<0.05 |  |  |  |  |  |

**Table A3: Logistic regression models of on-time compliance**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | | | | |  |
|  |  | **(A6)** | **(A7)** | **(A8)** | **(A9)** | **(A10)** |
|  |  | Country-  Year FEs |  |  |  |  |
| *Extensions†* |  | -2.41\*\* |  |  | -0.64\* | -0.63\* |
|  |  | (0.81) |  |  | (0.29) | (0.31) |
| *Follows†* |  |  | -0.25 |  |  |  |
|  |  |  | (0.15) |  |  |  |
| *Narrows†* |  |  |  | -0.29 |  |  |
|  |  |  |  | (0.39) |  |  |
| *Third Parties†* |  | -0.52 | -0.60 | -0.61 | -0.49 | -0.57 |
|  |  | (1.01) | (0.36) | (0.36) | (0.36) | (0.35) |
| *Trade Share* |  | -1.31\*\* | -0.11 | -0.12 | -0.13 | -0.15 |
|  |  | (0.49) | (0.09) | (0.09) | (0.10) | (0.10) |
| *GDP Share* |  | -7.39 | -2.11\*\* | -1.98\*\* | -2.02\* | -1.96\* |
|  |  | (4.45) | (0.79) | (0.74) | (0.79) | (0.79) |
| *US Respondent* |  |  | 1.16 | 1.04 | 1.15 | 1.27 |
|  |  |  | (0.63) | (0.60) | (0.66) | (0.67) |
| *AD Dispute* |  | -3.35\* | -1.20\* | -1.15\* | -1.31\* | -1.36\* |
|  |  | (1.55) | (0.56) | (0.56) | (0.58) | (0.59) |
| *Case Law†* |  | -0.09 | 0.56\*\* | 0.52\*\* | 0.57\*\* | 0.59\*\* |
|  |  | (0.63) | (0.20) | (0.19) | (0.21) | (0.20) |
| *Leg. Measure* |  | -2.11 | -0.39 | -0.45 | -0.47 | -0.47 |
|  |  | (1.69) | (0.44) | (0.44) | (0.45) | (0.43) |
| *GATT Dispute* |  |  |  |  | -0.37 |  |
|  |  |  |  |  | (0.64) |  |
| *Subsidy Dispute* |  |  |  |  | 0.13 |  |
|  |  |  |  |  | (0.45) |  |
| *Agric. Dispute* |  |  |  |  | -0.65 |  |
|  |  |  |  |  | (0.65) |  |
| *Firms (Compl.)* |  |  |  |  |  | 0.11 |
|  |  |  |  |  |  | (0.10) |
| *Firms (Resp.)* |  |  |  |  |  | -0.09 |
|  |  |  |  |  |  | (0.12) |
| Constant |  | -8.05 | 2.64\* | 2.56\* | 2.99\* | 2.77\* |
|  |  | (4.60) | (1.18) | (1.17) | (1.23) | (1.21) |
| Observations |  | 158 | 158 | 158 | 158 | 158 |
| Log Likelihood |  | -28.32 | -95.94 | -97.18 | -93.82 | -93.96 |
| Notes: \*\*p<0.01, \*p<0.05, †log units. Robust standard errors clustered by dispute group. | | | | | | |

**Table A4. Additional Controls: Economic Stakes and Time Trend**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **(A11)** | **(A12)** | **(A13)** | **(A14)** |
|  | *On-time* | *On-time* | *On-time* | *On-time* |
| VARIABLES | *Comply* | *Comply* | *Comply* | *Comply* |
|  |  |  |  |  |
| *Extensions†* | -0.13\* | -0.13\* | -0.14\* | -0.13\* |
|  | (0.06) | (0.06) | (0.06) | (0.06) |
| *Third Parties†* | -0.12 | -0.09 | -0.10 | -0.07 |
|  | (0.07) | (0.08) | (0.08) | (0.09) |
| *Trade Share* | -0.03 | -0.02 | -0.03 | -0.02 |
|  | (0.02) | (0.02) | (0.02) | (0.02) |
| *GDP Share* | -0.47\* | -0.43\*\* | -0.44\*\* | -0.39\* |
|  | (0.17) | (0.15) | (0.15) | (0.20) |
| *US Respondent* | 0.25 | 0.28\* | 0.26\* | 0.22 |
|  | (0.13) | (0.12) | (0.13) | (0.15) |
| *AD Dispute* | -0.26\* | -0.24\* | -0.26\* | -0.39\*\* |
|  | (0.11) | (0.11) | (0.11) | (0.12) |
| *Case Law†* | 0.11\*\* | 0.12\*\* | 0.13\* | 0.17\*\* |
|  | (0.04) | (0.03) | (0.06) | (0.06) |
| *Leg. Measure* | -0.08 | -0.01 | -0.08 | -0.17 |
|  | (0.09) | (0.09) | (0.09) | (0.10) |
| *Total Trade†* | 0.02 |  |  |  |
|  | (0.05) |  |  |  |
| *Merchandise Disp.* |  | 0.26\*\* |  |  |
|  |  | (0.10) |  |  |
| *Year Consult. Request* |  |  | -0.01 |  |
|  |  |  | (0.02) |  |
| *Year AB Ruling* |  |  |  | -0.02 |
|  |  |  |  | (0.02) |
| Constant | 0.96\*\* | 0.70\* | 1.06\*\* | 1.05\*\* |
|  | (0.42) | (0.26) | (0.24) | (0.27) |
| Observations | 157 | 158 | 158 | 109 |
| Sample | Full | Full | Full | AB Ruling |
| R-squared | 0.16 | 0.20 | 0.16 | 0.26 |
| Robust standard errors clustered by dispute group. \*\* p<0.01, \* p<0.05 | | | | |

**Table A5. Parallel Analysis with Compliance Delay**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **(A15)** | **(A16)** | **(A17)** | **(A18)** | **(A19)** | **(A20)** |
|  | *Time to* | *Time to* | *Time to* | *Time to* | *Time to* | *Time to* |
| VARIABLES | *Comply* | *Comply* | *Comply* | *Comply* | *Comply* | *Comply* |
|  |  |  |  |  |  |  |
| *Extensions†* | 0.54\*\* |  |  |  | 0.51\*\* | 0.53\* |
|  | (0.20) |  |  |  | (0.16) | (0.20) |
| *Third Parties†* | -0.18 | -0.01 | -0.004 | -0.004 | -0.01 | -0.04 |
|  | (0.27) | (0.26) | (0.25) | (0.25) | (0.22) | (0.25) |
| *Trade Share* | 0.25 | 0.28\* | 0.26 | 0.26 | 0.29\* | 0.31\* |
|  | (0.14) | (0.13) | (0.14) | (0.14) | (0.13) | (0.13) |
| *GDP Share* | 1.08 | 1.19 | 1.20 | 1.19 | 1.23 | 1.20 |
|  | (0.75) | (0.74) | (0.74) | (0.74) | (0.76) | (0.75) |
| *US Respondent* | -0.83 | -1.02 | -0.96 | -0.92 | -0.97 | -1.03\* |
|  | (0.55) | (0.54) | (0.54) | (0.54) | (0.52) | (0.48) |
| *AD Dispute* | 0.56 | 0.66 | 0.63 | 0.61 | 0.71 | 0.59 |
|  | (0.56) | (0.59) | (0.58) | (0.59) | (0.59) | (0.55) |
| *Case Law†* | -0.42\* | -0.31 | -0.32\* | -0.30 | -0.42\*\* | -0.32\* |
|  | (0.19) | (0.16) | (0.16) | (0.16) | (0.14) | (0.15) |
| *Leg. Measure* | 0.95\* | 0.90\* | 0.89\* | 0.96\* | 1.06\*\* | 0.90\* |
|  | (0.37) | (0.37) | (0.37) | (0.39) | (0.38) | (0.36) |
| *Year Consult. Request* | 0.07 |  |  |  |  |  |
|  | (0.07) |  |  |  |  |  |
| *De-Meaned Ext.* |  | 0.42\* |  |  |  |  |
|  |  | (0.19) |  |  |  |  |
| *Follows* |  |  | 0.21 |  |  |  |
|  |  |  | (0.11) |  |  |  |
| *Narrows* |  |  |  | 0.40 |  |  |
|  |  |  |  | (0.30) |  |  |
| *GATT* |  |  |  |  | 1.32\*\* |  |
|  |  |  |  |  | (0.50) |  |
| *Subsidies* |  |  |  |  | 0.03 |  |
|  |  |  |  |  | (0.35) |  |
| *Agriculture* |  |  |  |  | 0.78\* |  |
|  |  |  |  |  | (0.39) |  |
| *Firms (Compl.)* |  |  |  |  |  | -0.03 |
|  |  |  |  |  |  | (0.05) |
| *Firms (Resp.)* |  |  |  |  |  | -0.10 |
|  |  |  |  |  |  | (0.10) |
| Constant | 2.37 | 2.11 | 2.18 | 2.23 | 0.79 | 1.93 |
|  | (1.44) | (1.37) | (1.40) | (1.38) | (1.39) | (1.35) |
|  |  |  |  |  |  |  |
| Observations | 111 | 111 | 111 | 111 | 111 | 111 |
| R^2 | 0.24 | 0.22 | 0.22 | 0.21 | 0.32 | 0.24 |
| Robust standard errors clustered by dispute group. \*\* p<0.01, \* p<0.05 | | | | | | |

**Table A6: Duration with Cox Proportional Hazard Models**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **(A21)** | | **(A22)** | |
|  | **(a)** | **(b)** | **(c)** | **(d)** |
| *Accumulation of Prior Extensions*† | -0.39 \*\* |  | -0.33 \*\* |  |
|  | (0.07) |  | (0.09) |  |
| *Accumulation of Prior Precedent*† |  | -0.22 \* |  | -0.22 \* |
|  |  | (0.09) |  | (0.09) |
| *Third Parties*† | -0.50 \*\* | -0.65 \*\* | -0.43 \* | -0.52 \*\* |
|  | (0.16) | (0.14) | (0.17) | (0.15) |
| *Legislative* | -0.37 | -0.42 | -0.27 | -0.33 |
|  | (0.39) | (0.42) | (0.25) | (0.26) |
| *Trade Share* | -0.15 \*\* | -0.12 \*\* | -0.15 \*\* | -0.11 \*\* |
|  | (0.04) | (0.04) | (0.04) | (0.05) |
| *GDP Share* | -1.00 \*\* | -0.93 \*\* | -1.16 \*\* | -1.06 \*\* |
|  | (0.23) | (0.29) | (0.25) | (0.32) |
| *US Respondent* | 1.56 \*\* | 1.16 \*\* | 1.17 \*\* | 0.82 \*\* |
|  | (0.28) | (0.31) | (0.26) | (0.22) |
| *AD Dispute* | -0.53 | -0.62 | -0.38 | -0.45 |
|  | (0.40) | (0.41) | (0.26) | (0.27) |
| *Case Law*† | 0.50 \*\* | 0.51 \*\* | 0.50 \*\* | 0.51 \*\* |
|  | (0.08) | (0.07) | (0.09) | (0.09) |
| Observations | 680 | 680 | 435 | 435 |
| R^2 | 0.20 | 0.18 | 0.24 | 0.23 |
| Disputes | 142 | 142 | 142 | 142 |
| …with compliance | 116 | 116 | 116 | 116 |
| Start of Period | Consultations | | Panel Ruling | |
| Notes: \*\*p<0.01, \*p<0.05, †log units. Robust standard errors clustered by respondent. | | | | |

**Table A7. Additional Controls**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **(A23)** | **(A24)** | **(A25)** | **(A26)** | **(A27)** |
| VARIABLES | *On-time Comply* | *On-time Comply* | *On-time Comply* | *On-time Comply* | *On-time Comply* |
| *Extensions†* | -0.13 \* | -0.13 \* | -0.15 \* | -0.13 \* | -0.15 \* |
|  | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) |
| *Third Parties†* | -0.09 | -0.12 | -0.10 | -0.05 | -0.11 |
|  | (0.07) | (0.07) | (0.07) | (0.07) | (0.07) |
| *Trade Share* | -0.02 | -0.03 | -0.03 | -0.02 | -0.03 |
|  | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
| *GDP Share* | -0.45 \*\* | -0.43 \*\* | -0.55 \*\* | -0.17 | -0.43 \*\* |
|  | (0.15) | (0.15) | (0.15) | (0.15) | (0.15) |
| *US Respondent* | 0.28 \* | 0.25 \* | 0.34 \*\* | 0.14 | 0.19 |
|  | (0.13) | (0.12) | (0.12) | (0.13) | (0.20) |
| *AD Dispute* | -0.27 \* | -0.32 \*\* | -0.30 \*\* | -0.24 \* | -0.26 \* |
|  | (0.11) | (0.11) | (0.11) | (0.11) | (0.11) |
| *Case Law†* | 0.10 \* | 0.11 \*\* | 0.13 \*\* | 0.08 \* | 0.12 \*\* |
|  | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) |
| *Leg. Measure* | -0.08 | -0.06 | -0.08 | -0.09 | -0.07 |
|  | (0.09) | (0.09) | (0.09) | (0.08) | (0.09) |
| *Great Powers* | -0.18 |  |  |  |  |
|  | (0.12) |  |  |  |  |
| *Remedy Dispute* |  | 0.14 |  |  |  |
|  |  | (0.10) |  |  |  |
| *Compliance Ratio* |  |  |  | 0.76 \*\* |  |
|  |  |  |  | (0.13) |  |
| *Civil Law* |  |  |  |  | 0.09 |
|  |  |  |  |  | (0.17) |
| Constant | 0.95 \*\* | 1.00 \*\* | 1.05 \*\* | 0.32 | 1.04 \*\* |
|  | (0.25) | (0.23) | (0.24) | (0.22) | (0.24) |
| Observations | 158 | 158 | 147 | 158 | 158 |
| Zeroing Omitted? | No | No | **Yes** | No | No |
| R^2 | 0.17 | 0.18 | 0.20 | 0.27 | 0.16 |

Notes: \*\*p<0.01, \*p<0.05, †log units. Robust standard errors clustered by respondent.

**Table D1. Share of all WTO disputes and AB reports by issue area**

*Provision All Disputes Our Sample*

*(%) (%)*

Agriculture 16 17

Anti-dumping 18 17

GATT 1994 79 90

Safeguards 8 14

SPS 6 6

Subsidies 21 25

TBTs 9 9

Textiles 4 5

TRIPs 7 3

*Notes:* This table shows the share of disputes by issue area for all WTO cases and just those that end in appeal. The total percentages exceed 100 for each column because many disputes involve more than one GATT/WTO Agreement.

**Table D2. Share of precedent applications that extend the law by issue area**

*Provision Citations Extensions Share (%)*

Agriculture 438 45 10.3

Anti-dumping 303 34 11.2

GATT 1994 1615 178 11.0

Safeguards 192 32 16.7

SPS 268 17 6.3

Subsidies 403 43 10.7

TBT 271 28 10.3

Textiles 79 5 6.3

TRIPs 57 0 0

Total (all disputes) 5,518 575 10.4

*Notes*: This table shows the total number of total citations (i.e., precedent applications) and the number of extensions by issue area. The mean amount of extensions across the entire caseload is 10.4%. The most controversial areas of the law, such as antidumping, are extended at a rate consistent with the mean.

1. The Trade Law Guide jurisprudence citators identify all reports (both Appellate and Panel) that have cited a past decision as well as paragraphs in which they were cited. It does not show cases that were cited in parties’ submissions. Rather it shows cases used actively or passively by the AB as support, authority, or context in its analysis or ruling. [↑](#footnote-ref-1)
2. We exclusively examine substantive AB decisions. Compliance proceedings under Articles 21 and 22 are excluded from this analysis. Neither are they examined for references to other Appellate Body reports. [↑](#footnote-ref-2)
3. There are two special circumstances. In some instances, the AB applies the rationale from one agreement to disputes over another, a technique that we coded as expansion. For example, AB reports in several disputes about WTO flexibility provisions, including DS295 *Mexico* – *Definitive Anti-dumping Measures on Beef and Rice*, cite the Agreement Establishing the World Trade Organization. In other cases, the AB mentions a prior decision without openly adopting any of the described behaviors, hence we exclude this from the analysis. [↑](#footnote-ref-3)
4. Appellate Body Report, *Chile – Price Band System and Safeguard Measures Relating to Certain Agricultural Products*, WTO Doc. WT/DS207/AB/R (adopted Oct. 23, 2002). [↑](#footnote-ref-4)
5. Appellate Body Report, *Canada – Certain Measures Concerning Periodicals*, WTO Doc. WT/DS31/AB/R (adopted July 30, 1997). [↑](#footnote-ref-5)
6. Appellate Body Report, *United States – Sunset Reviews of Anti-Dumping Measures on Oil Country Tubular Goods from Argentina,* ¶ 188, WTO Doc. WT/DS268/AB/R (adopted Dec. 17, 2004). [↑](#footnote-ref-6)
7. Appellate Body Report, *United States – Final Anti-Dumping Measures on Stainless Steel from Mexico*, ¶ 160, WTO Doc. WT/DS344/AB/R (adopted May 20, 2008) (emphasis added). [↑](#footnote-ref-7)
8. *US – Stainless Steel (Mexico)*, *supra* note 90, ¶ 158 (citing Appellate Body Report, *Japan – Taxes on Alcoholic Beverages*, p. 14, WTO Doc. WT/DS8/AB/R, WT/DS10/AB/R, WT/DS11/AB/R (adopted Nov. 1, 1996) [hereinafter *Japan – Alcoholic Beverages II*]). In that case the E.U. argued “[w]hether as a matter of doctrine or practice, a high value is placed on consistency, certainty and predictability of the jurisprudence, particularly as regards decisions rendered by the highest courts”). [↑](#footnote-ref-8)
9. Office of the United States Trade Representative, Report on the Appellate Body of the World Trade Organization 58 (2020), <https://ustr.gov/sites/default/files/Report_on_the_Appellate_Body_of_the_World_Trade_Organization.pdf.> [↑](#footnote-ref-9)
10. The AB concluded “the panel's obligation to make an 'objective assessment of the matter' under Article 11 of the DSU and, in part, from the obligations imposed by Article 4.2, to the extent that those obligations are part of the claim. Thus, as with any claim under the provisions of a covered agreement, panels are required to examine, in accordance with Article 11 of the DSU7, whether the Member has complied with the obligations imposed by the particular provisions identified in the claim”, See AB Report, *US – Lamb,* WTO Doc.WT/DS177/AB/R para. 105. [↑](#footnote-ref-10)
11. Appellate Body Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WTO Doc.WT/DS285/AB/R (adopted Apr. 20, 2005). [↑](#footnote-ref-11)
12. Appellate Body Report, *Argentina – Measures Relating to Trade in Goods and Services*, ¶ 6.24, ¶ 6.25 WTO Doc. WT/DS453/AB/R (adopted May 9, 2016). [↑](#footnote-ref-12)
13. *Id.* ¶ 6.25. [↑](#footnote-ref-13)
14. *Id.* ¶ 6.25. [↑](#footnote-ref-14)
15. *Id.* ¶ 6.31. [↑](#footnote-ref-15)
16. Appellate Body Report, *Canada* *– Measures Affecting the Importation of Milk and the Exportation of Dairy Products*, WTO Doc. WT/DS103/AB/R, WT/DS113/AB/R (adopted Oct. 27, 1999). [↑](#footnote-ref-16)
17. Appellate Body Report, *United States – Tax Treatment for “Foreign Sales Corporations”*, WTO Doc. WT/DS108/AB/R (adopted Mar. 20, 2000). Notice that the initial interpretation of the word “payments” in *Canada – Dairy*, in the specific context of Article 9.1(c) was an expansive one, imputing into its meaning, “payment in kind” which may be read to include foregone revenue. In *US - FSC*, the AB effectively shrinks the meaning of subsidy for the entire agreement to mean only foregone revenue in a move that all but destroys the young and fragile precedent set in *Canada – Dairy*, and creates expansive implications, with that limited interpretation. [↑](#footnote-ref-17)
18. *Id.* ¶ 138. [↑](#footnote-ref-18)
19. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/33, June 25, 1997. [↑](#footnote-ref-19)
20. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/50, December 14, 1998. [↑](#footnote-ref-20)
21. *Ibid.* [↑](#footnote-ref-21)
22. *Ibid.* [↑](#footnote-ref-22)
23. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/62, July 2, 1999. [↑](#footnote-ref-23)
24. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/103, June 6, 2001. [↑](#footnote-ref-24)
25. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/189, June 17, 2005. [↑](#footnote-ref-25)
26. *Ibid*. [↑](#footnote-ref-26)
27. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/200, January 25, 2006. [↑](#footnote-ref-27)
28. *Ibid*. [↑](#footnote-ref-28)
29. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/241, December 14, 2007. [↑](#footnote-ref-29)
30. Minutes of Meeting, Dispute Settlement Body, WT/DSB/M/250, July 1, 2008. [↑](#footnote-ref-30)