

# How Do Third Parties Affect Compliance in the Trade Regime?

Jeffrey Kucik, University of Arizona

Lauren Peritz, University of California, Davis

A core insight of the literature on dispute settlement at the World Trade Organization (WTO) is that third party countries help enforce the organization's multilateral objectives, including the fundamental principle of nondiscrimination. Little is known, however, about when countries comply with WTO rulings and whether these bystander states play a role. We introduce new data on compliance, measured as whether losing countries make tangible domestic reforms to bring policy in line with WTO rulings. We show that compliance is significantly less likely in disputes with more third parties. Using a variety of estimation techniques, including controlling for nonrandom selection into legal rulings, we demonstrate a robust correlation between third party participation and noncompliance. Our findings highlight a risk of stringent enforcement and suggest that compliance problems threaten to undercut the operation of the multilateral trade regime.

When do states comply with international legal rulings? We examine compliance in the context of the World Trade Organization (WTO), whose dispute system is widely considered the legal backbone of the global trade regime. One way the WTO tries to strengthen enforcement is by allowing third party countries to participate in disputes. Literature shows that third parties help protect the WTO's core principle of nondiscrimination. Third parties are thought to spoil the opportunity for litigants to strike discriminatory settlements (Busch and Reinhardt 2006), and they increase the punishment costs incurred by violators (Maggi 1999).

Do third parties, by playing an enforcement role, promote compliance with WTO rulings? We offer the first test of this question using original data on compliance, which we define as whether governments make *tangible policy changes* in the wake of adverse rulings.

The answer appears to be no. Relying on a variety of estimation strategies, we find that, while bystander states increase the likelihood of a ruling, compliance is actually less likely in the presence of third parties. One possible expla-

nation is that third parties are known to bring their own issues and preferences to bear on the dispute. This may expand the scope of disputes such that compliance becomes more burdensome for the defendant. Our finding lends insight into why bystander governments fear overcrowding in WTO disputes and suggests a hazard of stricter enforcement in international adjudication (Johns 2015; Johns and Pelc 2015; Rosendorff 2005).

## THIRD PARTIES AS ENFORCERS

Third parties are widely believed to play an enforcement role at the WTO. Third parties protect the founding principle of nondiscrimination by spoiling early settlements between litigants. These are bilateral deals that allow litigants to resolve their disputes before litigation. Private settlement creates opportunities for discriminatory deals that fail to extend trade concessions to the full membership (Bagwell and Staiger 2004; Kucik and Pelc 2016).

Busch and Reinhardt (2006) identify three ways bystander states increase the likelihood that a dispute goes to a formal ruling. First, simply having more eyes in the room

---

Jeffrey Kucik (jkucik@email.arizona.edu) is an associate professor in the School of Government and Public Policy and in the James E. Rogers College of Law (by courtesy) at the University of Arizona. Lauren Peritz (lperitz@ucdavis.edu) is an assistant professor in political science at the University of California, Davis.

Financial support for this research was provided by the National Science Foundation, grant 1421326. An online appendix with supplementary material is available at <https://doi.org/10.1086/711058>.

can reduce the bargaining space as litigants adopt harder-line postures (Stasavage 2004). Second, third parties complicate bargaining by bringing their own issues and preferences to bear (Davey 2005; Waincymer 1996). Finally, bystanders may seek clarification of the law when systemic issues are at play (Johns and Pelc 2014; Porges 2003). For these reasons, third parties encourage public rulings. Rulings can be preferable because they eliminate the opportunity for litigants to reach deals by offering side payments that violate WTO norms. Thus, they allow for better enforcement of the WTO's nondiscrimination principle and deliver benefits to the broader membership.

Third parties should also lend greater weight to WTO decisions. In a system with decentralized enforcement, involving more states ought to increase the material and reputational costs of violating the agreement (Maggi 1999). Hence, there is general consensus that third parties are beneficial to the WTO dispute process. One implication is that third parties should promote downstream compliance. After all, enforcement is only beneficial if it leads to compliant behavior.

Unfortunately, it is also possible that third parties reduce compliance rates. For one thing, stricter enforcement is known to backfire in international law (Downs, Rocke, and Barsoom 1996). Moreover, third parties bring their own arguments to bear on the dispute (Busch and Reinhardt 2006). Doing so increases the likelihood of a ruling, but it may decrease compliance if third parties expand the dispute. When third parties introduce new issues or legal arguments, they may effectively increase the burden placed on losing defendants by expanding the ruling. The result would be less compliance. (Our appendix, available online, discusses some disputes in which this has occurred.)

## RESEARCH DESIGN

We examine the effect of third parties using the most complete record of WTO dispute compliance to our knowledge. In doing so, we build on foundational work by Hudec, Kennedy, and Sgarbossa (1993) on General Agreement on Tariffs and Trade (GATT) era trade disputes. Researchers have shied away from analyzing the determinants of compliance because collecting adequate data is a formidable task.<sup>1</sup> Yet we cannot make inferences about the strength of international law to affect state behavior without looking beyond just rulings.

1. Busch and Reinhardt (2003) consider compliance with nine WTO disputes between the United States and Europe. Davey (2005) looks at 58 disputes during the first 10 years of WTO operation and finds that, 83% of the time, defendants reported some form of implementation, but he does not verify whether the measures achieved compliance.

This is the most comprehensive effort to do so in the trade realm. Our data contain one observation per WTO dispute from DS1 to DS415—those initiated between years 1995 and 2012.<sup>2</sup> Our effective sample is the 155 disputes that went to rulings that found the defendant in violation.<sup>3</sup> The analysis is limited to rulings because we cannot observe the terms of private settlement.<sup>4</sup>

## Coding compliance

We define compliance as domestic reforms that bring trade policies in line with WTO rulings. Policy reform is not the only way to comply; governments can accept retaliation. However, reform represents a tougher, more meaningful test, namely, whether a government is willing to dismantle a trade barrier that protects important domestic interests.

Collecting data on compliance requires going beyond the WTO because those institutional sources supply an incomplete picture. First, the WTO relies on voluntary reporting from defendants who routinely declare compliance but do not provide supporting evidence of implementation.<sup>5</sup> Second, the WTO relies on complainants to pursue unresolved violations. Yet countries often drop disputes if additional litigation appears futile.<sup>6</sup> Therefore, the true extent of compliance cannot be inferred through relying on the WTO's information alone.

To code compliance, we sought out evidence from primary sources in seven languages.<sup>7</sup> We reviewed the results of the panel ruling to identify specific WTO-inconsistent measures as well as the implementation deadline. For each of the 155 disputes, we looked for any official statutory and regulatory measures passed by defendants, that is, tangible policy changes made before the deadline, the "reasonable period of time" for implementation. Our coding reflects

2. The sample ends in 2012, to ensure that WTO disputes had sufficient time to resolve.

3. There were 170 rulings, with findings of violation in 160. When the defendant lost and had the ruling overturned on appeal, there is no question of compliance; we do not code those cases.

4. Because settlements through mutually agreed solutions often have trade repercussions indicative of discrimination, they are unlikely to represent compliant behavior. See Gray and Potter (2019) on the diplomatic asymmetries that lead to settlement.

5. A clear example is DS132: "Mexico—High Fructose Corn Syrup." Mexico reported to the WTO that it had implemented the panel rulings, but a compliance investigation demonstrated Mexico remained in violation.

6. We cannot rely on article 21.5 proceedings or suspensions of concessions as a definitive indicator of noncompliance, as it would underrepresent the actual rate.

7. In addition to English, we used Arabic, Chinese, Japanese, Korean, Portuguese, and Spanish language sources. These cover the official languages of almost all defendant governments and 95% of cases.

legislative as well as administrative measures governments implemented to correct their violations.

For disputes over WTO-inconsistent domestic legislation, we located those laws from national parliamentary or congressional websites. Some were repealed, others amended, to bring the defendant into compliance. In other cases, amendments were proposed but failed, suggesting noncompliance. In disputes over administrative measures, like antidumping duties or safeguards, we located policies from trade ministries, executive orders, and other bureaucratic actions and recorded whether they were rescinded or modified. We found several instances of partial compliance (e.g., duties were lowered on some but not all the products implicated in a ruling).

In addition, we surveyed industry publications and news reports to verify whether ostensible compliance measures were contested. If an industry group in the defendant country claimed it lost valuable trade protections, we take this as further evidence of compliance. When stakeholders in the complainant state accused the defendant of insufficient measures, we take this as further indication of noncompliance. The appendix illustrates our coding methodology.

We found evidence of full compliance in 101 disputes, partial compliance in 14, and noncompliance in the remaining 40 disputes. Among disputes that received an adverse ruling, there are 30 different complainants and 19 different defendants. The European Union (28 losses) and the United States (58 losses) are the most frequent litigants. We summarize our data on (non)compliance by country in the appendix.

### Variables

We construct two dependent variables. Any Compliance is a dichotomous coding of whether there was either full or partial compliance (1) as opposed to none (0). Degree of Compliance is an ordinal coding where we disaggregate full (2) from partial (1) compliance and where noncompliance remains 0.<sup>8</sup> Our main explanatory variable is the number of third parties in each dispute. Roughly half of all dispute filings have no third parties, and most of the remaining cases have six or fewer. However, some disputes attracted as many as 24. Since the distribution is skewed, we use the natural log (ln Third Parties).<sup>9</sup>

8. Because the ordinal measure assumes proportionality, we prefer the dichotomous measure. In robustness checks, we distinguished between on-time and delayed compliance, finding consistent support.

9. When there are no third parties, we substitute a zero for the log-transformed count. We also considered a measure counting just those third parties who supported the complainant (Busch and Pelc 2010), with consistent results.

The models control for many confounding factors. These include participants' relative market size (GDP Share), calculated as the defendant's gross domestic product (GDP) divided by the sum of complainant and third party GDPs. They also include the share of a defendant's trade accounted for by the complainants and third parties (Trade Share), to measure both litigant trade ties and the retaliatory threat facing defendants should they fail to comply (Bown 2005). We control for the number of times the disputed issues have been ruled on before (Times Ruled), as a proxy for issue salience. We include a 0/1 indicator of whether the case involves trade Remedies and Clarity cases since compliance rates may vary by issue area.<sup>10</sup> We control for whether a dispute was filed under Article XXII of the agreement, which the complainant can use to facilitate third party participation. Finally, we account for lawsuits against EU/US Respondents. These members have the legal/bureaucratic resources to defend their interests and enjoy market power over the membership.

We address the possibility that third parties simply proxy the dispute's underlying economic and political stakes. We measure the product-specific Disputed Imports and the share of Employment in the affected industries. We address the rationale for and coding of these control variables and other robustness checks in the appendix.

### Nonrandom selection

One potential source of nonrandom selection requires particular attention: whether a case ends in a formal ruling. It is possible that third parties overwhelmingly participate in high-profile, controversial cases: exactly those disputes in which defendants are less likely to back down (either during consultations or after rulings). In that scenario, third parties could exert their influence by spoiling early settlements for highly resolved defendants and thus would exert no independent effect on the probability of compliance.

We model the selection process using several additional variables.<sup>11</sup> We account for whether third party countries entered proceedings citing Systemic Interest since these cases are more likely to be litigated.<sup>12</sup> We identify disputes that

10. We code antidumping, countervailing duty, and safeguard disputes as "remedy" cases. We code for disputes over newer technical aspects of the law in which the WTO is called on to provide clarity in interpretation. Clarity cases are trade-related intellectual property rights (TRIPS), investment measures (TRIMs), services, technical barriers to trade (TBTs), and sanitary and phytosanitary measures (SPS; Kucik 2019).

11. See the appendix for other robustness checks for nonrandom selection.

12. This is an indicator for disputes in which at least one third party joined by citing a systemic interest rather than a substantive trade interest. It is coded by evaluating each country's request to join proceedings (Johns and Pelc 2014).

addressed politically sensitive issues: the WTO agreements on Agriculture, Sanitary and Phytosanitary Measures, or Services. Respondents may be inclined to settle sensitive topics behind closed doors before litigation heightens public scrutiny. We account for the proportion of times defendants lost previous rulings on the same issues implicated in the dispute (Adverse Frequency). If complainants bring a lawsuit containing claims that are often won, they may push for a ruling assuming another victory. Finally, we control for the proportion of disputes that go to a ruling each year in case there are time trends.

## FINDINGS

Table 1 presents our statistical results in which we regress compliance on the logged count of third parties using probit, ordered probit, linear, and Heckman selection models. Our unit of analysis is the dispute, and we cluster standard errors by dispute group.<sup>13</sup>

Models 1 and 2—our dichotomous and ordinal measures, respectively—demonstrate that third parties are strongly, negatively correlated with compliance. The effects in model 1 are substantively large. When no bystander countries intervene, which occurs in roughly 6% of rulings, the predicted probability of compliance with an adverse ruling is 0.96 [0.87, 0.99].<sup>14</sup> When 10 participate, representing the 75th percentile, the probability of compliance drops to just over half (0.62 [0.48, 0.73]). Models 3 and 4 repeat the analysis with linear probability models. Models 5 and 6 show that the economic and political stakes, measured by bilateral disputed merchandise imports (5) and industry employment share (6), are not predictors of compliance. Across these models, third party participation remains strongly correlated with noncompliance. Additional tests of dispute stakes are provided in the appendix.

Next, we model compliance while explicitly accounting for the selection into adverse rulings in models 7 and 8. We find a strong negative correlation between third party participation and noncompliance. Confirming the existing literature, we find that disputes are more likely to go to ruling when there are third party participants or when a larger share of the respondent's trade is with the complainant and bystander countries. To identify the selection model, we rely on a combination of variables. The Systemic importance of a dispute is a powerful predictor of rulings but has little bearing

on compliance with those rulings.<sup>15</sup> Disputes over sensitive topics are far more likely to be settled early. When the case law reveals that the complainant has a strong chance of prevailing in litigation (Adverse Frequency), complainants are less likely to settle and more likely to seek a ruling. Models 7 and 8 reveal only modest residual correlation between the two stages, suggesting selection into rulings is not driving our substantive finding.<sup>16</sup> The magnitude of the third parties' effect on compliance is diminished. This is consistent with previous findings that third parties limit their participation for fear of "overcrowding."

## DISCUSSION

Compliance rates are either high or low depending on one's point of view. Many international organization scholars argue that compliance with international law is relatively common. However, realists may actually be surprised that an organization like the WTO, without stronger enforcement, would see so much compliance—particularly given the money at stake.

Either way, we find that third parties, who are supposed to strengthen enforcement at the WTO, actually drive down compliance rates. These findings are consistent with the idea that transparency can backfire in global governance (Stasavage 2004) and that countries will renege on their commitments when pressured too greatly (Downs et al. 1996). They also resonate with skepticism about implementation of trade commitments (Gray and Kucik 2017). And, they reinforce the argument that third party participation can be self-defeating (Johns and Pelc 2014). From the WTO's point of view, then, there is a balance between allowing bystander participation and compliance with the rules.

What explains the negative association between third parties and compliance? One possibility is that one of the very same mechanisms that spoil settlement—third parties bringing their own issues and preferences to the table—also drives down compliance with rulings. Third parties can effectively expand the scope of the dispute by introducing new arguments. As WTO jurisprudence states, "a panel is not limited by the arguments made by the parties to a dispute. . . . It can certainly consider the arguments made by third parties."<sup>17</sup> This makes early settlement less likely, but it also means that the price tag attached to rulings is higher. Rulings may ask

13. The WTO sometimes combines complaints into a single legal proceeding (e.g., DS8, DS10, and DS11).

14. Brackets report 95% confidence intervals. Predicted probabilities are calculated holding variables at their means.

15. The correlation between systemic cases and rulings is  $\rho = 0.534$ , whereas the correlation between systemic cases and compliance is  $\rho = 0.132$ . See the appendix.

16. There is modest correlation in the error between the selection and outcome stages, and the insignificant Inverse Mills Ratio suggests that the modeled selection process is not driving the outcome.

17. DS376: "Australia—Apples" panel report, paragraph 7.76.

Table 1. Baseline Estimates

	Compliance						Ruling Selection	Compliance Outcome	Ruling Selection	Compliance Outcome
	Probit (1)	Ordered Probit (2)	Linear (3)	Linear (4)	Probit (5)	Probit (6)				
Third Parties	-.65** (.19)	-.56** (.18)	-.18** (.05)	-.34** (.11)	-.76** (.28)	-.78* (.36)	.62** (.14)	-.17* (.07)	.60** (.14)	-.14** (.07)
Trade Share	1.27* (.56)	1.11* (.57)	.40* (.16)	.71* (.33)	1.62* (.69)	2.13* (1.06)	2.17** (.41)	.45* (.21)	2.18** (.39)	.54* (.21)
Times Ruled	.01 (.01)	.01 (.01)	.00 (.00)	.00 (.00)	-.00 (.01)	-.02 (.01)	-.00 (.00)	.00 (.00)	.00 (.01)	.00 (.00)
GDP Share	-.69 (.40)	-.52 (.39)	-.14 (.11)	-.28 (.23)	-1.02* (.43)	-.73 (1.63)	-.20 (.27)	-.11 (.11)	-.06 (.26)	-.12 (.11)
EU/US										
Respondents	.45 (.35)	.38 (.31)	.10 (.10)	.21 (.21)	.99* (.47)		.45* (.22)	.12 (.09)	.50* (.23)	.13 (.09)
Article XXII	.64* (.28)	.53* (.25)	.17* (.08)	.33* (.16)	.47 (.41)	.75 (.69)	-.29 (.20)	.12 (.08)	-.22 (.19)	.11 (.08)
Remedy	-.03 (.31)	.08 (.28)	.01 (.08)	.06 (.18)	.04 (.44)	.10 (.90)	-.19 (.21)	.07 (.07)	-.21 (.19)	.07 (.07)
Clarity	-.79* (.32)	-.55 (.31)	-.22* (.09)	-.36 (.19)	-.78 (.45)	-1.70 (1.04)				
Disputed Imports					.08 (.08)					
Employment						-.14 (.18)				
Systemic							.75** (.23)		.67** (.23)	
Agriculture							-.47 (.29)			
Sanitary and Phytosanitary							-.76 (.46)			
Services							-.93* (.42)			
Rulings by Year									.12 (.95)	
Adverse Frequency									-.70** (.25)	
Constant	.89 (.63)		.74** (.17)	1.35** (.36)	.75 (.65)	.90 (2.10)	-1.61** (.31)	.57* (.29)	-1.70** (.44)	.43 (.29)
N	153	153	153	153	115	61		373		373
$\rho$								.35		.38
Inverse Mills Ratio								.15 (.15)		.14 (.15)

Note. Standard errors in parentheses.

\*  $p < .05$ .

\*\*  $p < .01$ .



more expansive trade concessions in the presence of third parties than they otherwise would have. The result is less compliance. This article is not the final word but rather the necessary first step: using new data to systematically establish a strong relationship between third parties and noncompliance.

## ACKNOWLEDGMENTS

The authors wish to thank James Adams, Stephen Chaudoin, Julia Gray, Leslie Johns, and Daniel Kono. Previous versions of this article were presented at the annual meeting of the International Studies Association (2018); the Bozeman, Montana, workshop on security and cooperation (2019); and the University of California, Berkeley's MIRTH speaker series (2019).

## REFERENCES

- Bagwell, Kyle, and Robert W. Staiger. 2004. "Multilateral Trade Negotiations, Bilateral Opportunism and the Rules of GATT/WTO." *Journal of International Economics* 63 (1): 1–29.
- Bown, Chad P. 2005. "Participation in WTO Dispute Settlement: Complainants, Interested Parties, and Free Riders." *World Bank Economic Review* 19 (2): 287–310.
- Busch, Marc L., and Krzysztof Pelc. 2010. "The Politics of Judicial Economy at the World Trade Organization." *International Organization* 64:257–79.
- Busch, Marc L., and Eric Reinhardt. 2003. "The Evolution of GATT/WTO Dispute Settlement." In John M. Curtis and Dan Ciuriak, eds., *Trade Policy Research, 2003*. Ottawa: Department of Foreign Affairs and International Trade, 143–83.
- Busch, Marc L., and Eric Reinhardt. 2006. "Three's a Crowd: Third Parties and WTO Dispute Settlement." *World Politics* 58 (3): 446–77.
- Davey, William. 2005. "The WTO Dispute Settlement System: The First Ten Years" *Journal of International Economic Law* 8 (1): 17–50.
- Downs, Geroge W., David M. Rocke, and Peter N. Barsoom. 1996. "Is the Good News about Compliance Good News about Cooperation?" *International Organization* 50 (3): 379–406.
- Gray, Julia, and Jeffrey Kucik. 2017. "Leadership Turnover and International Trade Commitments." *Comparative Political Studies* 50 (14): 1941–72.
- Gray, Julia, and Philip Potter. 2019. "Diplomacy and the Settlement of International Disputes." Working paper, University of Pennsylvania.
- Hudec, Robert, Daniel Kennedy, and Mark Sgarbossa. 1993. "A Statistical Profile of GATT Dispute Settlement Cases, 1948–1989." *Minnesota Journal of Global Trade* 2 (1): 1–114.
- Johns, Leslie. 2015. *Strengthening International Courts: The Hidden Costs of Legalization*. Ann Arbor: University of Michigan Press.
- Johns, Leslie, and Krzysztof Pelc. 2014. "Who Gets to Be in the Room? Manipulating Participation in WTO Disputes." *International Organization* 68 (3): 663–99.
- Johns, Leslie, and Krzysztof Pelc. 2015. "Fear of Crowds in WTO Disputes: Why Don't More Countries Participate?" *Journal of Politics* 78 (1): 88–104.
- Kucik, Jeffrey. 2019. "How Do Prior Rulings Affect Future Disputes?" *International Studies Quarterly* 63 (4): 1122–32.
- Kucik, Jeffrey, and Krzysztof Pelc. 2016. "Measuring the Cost of Privacy: A Look at the Distributional Effects of Private Bargaining." *British Journal of Political Science* 46 (4): 861–89.
- Maggi, Giovanni. 1999. "The Role of Multilateral Institutions in International Trade Cooperation." *American Economic Review* 89 (1): 190–214.
- Porges, Amelia. 2003. "Settling WTO Disputes: What Do Litigation Models Tell Us." *Ohio State Journal on Dispute Resolution* 19:141–84.
- Waincymer, Jeffrey. 1996. "Reformulated Gasoline under WTO Dispute Settlement Procedures: Pulling Pandora out of a Chapeau." *Michigan Journal International Law* 18:141–84.
- Rosendorff, B. Peter. 2005. "Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure." *American Political Science Review* 99 (3): 389–400.
- Stasavage, David. 2004. "Open-Door or Closed-Door? Transparency in Domestic and International Bargaining." *International Organization* 58 (4): 667–703.

Copyright of Journal of Politics is the property of The Southern Political Science Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.