
Graduate Institute of International and Development Studies

Center for Trade and Economic Integration

Working Paper Series

Working Paper CTEI-2017-04

Free-Riding on Enforcement in the WTO

Leslie Johns

Department of Political Science, University of California, ljohns@polisci.ucla.edu

Krzysztof J. Pelc

Department of Political Science, McGill University, kj.pelc@mcgill.ca

Rue de Lausanne 132
P.O. Box 136
CH - 1211 Geneva 21
Geneva, Switzerland

Free-Riding on Enforcement in the WTO

Abstract

Countries can challenge potential trade violations using the WTO's dispute settlement system, yet many policies that appear to violate WTO rules remain unchallenged, even when they have a significant economic impact. Why is this? We argue that the likelihood that a country challenges a protectionist policy is linked to how concentrated or diffuse the policy is. When a policy is *concentrated*—because it affects only one country—litigation is a private good, meaning that a country that pays the cost of litigation receives the full benefit of litigation. But when a policy is *diffuse*—because it affects many countries—litigation is a public good and countries face a collective action problem: many countries can benefit from litigation, but each country wants to free-ride by having another country pay the cost. The resulting selection effect has two consequences. First, the free-rider problem reduces the likelihood that a diffuse policy will be challenged in any given period, generating a longer *enforcement delay* for diffuse trade violations. Second, cases must have higher odds of success in order for countries to overcome the collective action problem, meaning that conditional on being filed, cases that challenge concentrated policies will be less likely to succeed in litigation than cases that challenge diffuse policies. We leverage selection effects to test our argument using data on the timing and outcomes of trade disputes. The evidence, which considers all WTO disputes from 1995 to 2013, bears out these beliefs.

1 Introduction

Like many international organizations, the World Trade Organization (WTO) relies upon its members to challenge possible violations. At times, its dispute settlement system appears highly effective, with members challenging protectionist trade policies as soon as they appear. For example, the United States blocked entry to Canadian trucks carrying cattle and swine in 1998, citing health concerns. Canada believed that this policy, which affected only Canadian trucks, violated WTO rules. Canada responded quickly: 15 days after the policy was implemented, Canada filed a dispute and requested expedited consultations with the US.¹

Yet the WTO's dispute settlement system sometimes appears ineffective, with trade violations going unchallenged for years.² For example, the US Agriculture Improvement and Reform Act of 1996 (FAIR Act) violated WTO rules by subsidizing US corn exports. Because subsidies depress world prices and corn is a widely produced commodity, the FAIR Act harmed a great number of countries. However, Canada was the only country that ultimately paid the cost of challenging the FAIR Act, and it did so 4,025 days—over 11 years—after the law's implementation.³ Why did Canada swiftly challenge the US in 1998 over a relatively limited trade restriction on cattle and swine, while ignoring large US corn subsidies for over a decade? More generally, why are there greater enforcement delays for some policies than others? Our claim is that Canada's decision can be explained in part by how the US trade policies affected *other* countries. In this article, we argue that *who else* is affected by trade measures explains which disputes a country chooses to file, and how likely those challenges are to succeed in litigation.

Our argument is not unique to the WTO. Enforcement is one of the key theoretical concepts in the study of international cooperation. Early scholars of international cooperation emphasized

¹DS144: *US—Certain Measures Affecting the Import of Cattle, Swine and Grain from Canada*.

²Chaudoin (2014) argues that such delays in the US are related to electoral cycles.

³DS357: *US—Agricultural Subsidies*.

the role of hegemons and reciprocity in enforcing agreements (Keohane, 1986; Kindleberger, 1973). More recent scholarship emphasizes the role of law and dispute settlement procedures, like those at the WTO (Abbott et al., 2000; Rosendorff, 2005). Regardless of its form, enforcement often serves as a public good: it benefits all members of the cooperative regime. This paper pinpoints the central obstacle in providing such public goods—the individual incentive to free-ride on the efforts of others—in the WTO, one of the key institutions in global governance. It also demonstrates that these incentives shape the kinds of violations that states challenge (and those they do not).

Countries that file WTO disputes pay a private cost for what is often a public good. The costs of filing WTO disputes include sizable legal costs—even the US finds itself constrained as to the number of disputes it can file simultaneously—but also, and perhaps especially, the inevitable political cost that comes from antagonizing a trade partner by challenging their policies in a public forum.⁴ Trade policies that are challenged in the WTO vary in two important ways. First, policies vary in their aggregate economic impact, as measured by the amount of trade at stake. Second, policies vary in the distribution of their impact across states. Some protectionist policies have relatively small but highly concentrated effects, like the US measure on Canadian trucks, where no country other than Canada would have also directly gained from enforcement. Others have far larger but more diffuse effects, like the US corn subsidies, which affect world prices, and thus all corn-producing countries. Regardless of the aggregate economic impact of trade policies, variation in their diffuseness determines the extent to which challenges of these policies represent public goods.

The public goods aspect of WTO enforcement is no secret. Yet until now, scholars have mostly portrayed it as a strength of the WTO’s institutional design. Guzman and Simmons (2005, 591) thus argue that “the wealthiest governments may be supporting the public good of enforcing trade liberalization”. Bown and Hoekman (2005, 862) make a similar claim, stating that “[a]ctive participation in dispute settlement activity by WTO member countries can also have positive externalities if one country’s litigation efforts contribute to the removal of a trade barrier that adversely affected the market access rights of other WTO members.” Davis (2012a) and Bechtel and Sattler (2015) also build on the notion of enforcement as a public good in their own work, representing it as a desirable aspect of dispute settlement.

The promise of positive externalities is in fact woven into the very design of the WTO’s legal rules. The fundamental principle of nondiscrimination, which underlies the trade regime, also extends to dispute settlement. Under Article 3.5 of the Dispute Settlement Understanding (DSU), any solution to a legal challenge must itself be compatible with WTO rules. This design is intended to foster positive externalities. Just as concessions reached in bilateral negotiations during multilateral trade rounds are then extended to all other countries under the most-favored

⁴Some disputes are filed by multiple complainants, but these represent a minority of cases. We examine this in greater detail below.

nation principle (MFN), so too must all concessions offered by a defendant to the complainant during dispute settlement be extended to the entire membership, regardless of who paid the costs of enforcement.⁵

While the challenge of collective action is anything but a novel concept in the study of courts, the rich literature studying the WTO has thus far failed to measure the full consequences of the public goods aspect of enforcement. The main reason for this is methodological. The empirical analysis of WTO disputes is plagued by the dual problems of observability and endogeneity. As in most domestic and international legal systems, we cannot observe all potential disputes—we only observe those disputes that countries actually file. Additionally, filing decisions are endogenous—countries strategically choose which cases to file. This dual problem generates selection effects that hinder valid statistical inference. Rather than being constrained by these selection effects, we use them to our advantage by leveraging the selection process inherent in our formal model’s equilibrium behavior. Namely, we test the implications of our causal argument for observable disputes, given that we know selection is occurring.

We argue that diffuseness of WTO disputes has two overlooked effects, which we derive using a formal model. First, we examine the timing of WTO disputes. Our model shows that holding the aggregate economic impact of a trade policy constant, a country should wait longer, on average, to file a case against a more diffuse policy. So conditional on a policy being challenged at the WTO, a more diffuse policy will have been in place longer than a more concentrated policy. To put it in terms of survival analysis, a diffuse policy should be less likely of being challenged in any given period, controlling for its total trade at stake. Second, we examine the predicted legal success of WTO cases. While litigation always comes with uncertainty, some potential cases are more likely to generate a pro-complainant ruling, because the underlying facts and applicable law are more favorable to the complainant. Because litigation is costly, a country must carefully weigh a potential case’s legal merit against its cost. For a concentrated policy, a country will only pay the cost of dispute settlement if the case has a sufficiently high chance of success. For a diffuse policy, a country has an even higher standard: given the temptation to free-ride on the efforts of others, a potential case must have an especially high chance of success for an affected country to file it. We do not directly test which cases are selected for litigation; instead, we derive predictions about the likely success of cases that *are* selected, and test this against the data. We show that in observed WTO disputes, cases that challenge concentrated policies succeed less often than cases that challenge diffuse policies. In other words, we expect that the incentive to free-ride on enforcement at the WTO will not only produce enforcement delays, but it will also result in the *wrong cases* being filed.

⁵A growing literature looks at the extent to which countries abide by this non-discrimination rule. The private nature of settlements reached during dispute settlement, indeed, fosters the possibility of bilateral opportunism (Bagwell and Staiger, 2004). Yet the very fact that this is being investigated is a testament to the existence and prevalence of the rule.

We test our arguments on 360 WTO disputes from 1995 to 2013. To measure enforcement delay, we rely on data from Bown and Reynolds (2015) on the precise implementation dates for all policies that were eventually challenged at the WTO. To test our predictions about success in litigation, we construct an original dataset of the ruling direction—whether the complainant prevailed on a legal argument—for all 4,484 individual claims brought before the WTO during the period of interest.

One implication of our argument pertains to *how* countries choose to violate their trade commitments. If violations that affect a greater number of trading partners are less likely to be swiftly challenged, then a government that seeks to violate its trade obligations to protect domestic constituents should avoid trade policies with concentrated effects. A government is best off when it can spread the pain around, as it were. Our argument also demonstrates the potential benefit from coordinating enforcement efforts among states: we suggest that private transnational actors may play a facilitating role in this respect. In the conclusion, we return to these and other issues, and discuss the broader generalizability of our argument.

2 Free-Riding on Enforcement

2.1 Private Costs and Public Goods

Many areas of international cooperation involve collaboration problems, in which states can benefit from joint cooperation but each state is tempted to not cooperate (Snidal 1985). Overcoming these temptations is one of the key challenges of global governance. Early scholars of international cooperation emphasized the role of hegemony in providing public goods, like the enforcement of international agreements (Kindleberger, 1973). Others demonstrated how repeated interactions, the number of players, patience, and other factors can affect the enforcement of cooperative agreements (Oye, 1986). Contemporary scholars have built upon the legacy of Keohane (1984) by focusing on role of formal and informal institutions.

For example, the crown jewel at the core of the contemporary international trade regime is the WTO's dispute settlement understanding (DSU). This international legal body is built on decentralized enforcement: while it has no centralized prosecutorial function, it allows countries to challenge one another's policies, and in the absence of a mutually agreeable settlement between the two parties, it adjudicates the matter and offers recommendations for compliance. These are binding in nature. In its scope, its rate of compliance, its rich jurisprudence, the WTO's legal body is arguably the most successful court in the international system.

Nonetheless, there are limits on the effectiveness of international institutions (Martin, 2013). Namely, an institution like the WTO can only adjudicate disputes that states bring before it. As in all legal systems, not all violations that occur in the trade regime are challenged. While

the exact proportion is difficult to estimate, most trade lawyers would agree that most violations are never challenged. In some cases, the stakes are too trivial; in others, the matter may be overly sensitive. Indeed, the trade regime remains in many respects true to its diplomatic origins, and countries try hard to manage one another’s political sensibilities. The result is a process of selection: governments must decide where to allocate their (scarce) enforcement resources to yield the greatest benefit to their exporters, at the least financial and political cost. States can be expected to challenge those policies that cause them the most harm, which are likely to be those that affect the products that they trade in the most. Yet the decision to file is not made in a vacuum, since the same violation often affects more than one country. Governments are aware that others are also choosing whether to challenge such measures. As we argue, the resulting choices of all governments thus become strategically interdependent.

Enforcement is costly for the state that provides it. For example, filing a dispute at the WTO requires considerable legal capacity that many countries lack (Busch, Reinhardt and Shaffer, 2009). The average WTO dispute is estimated to cost in the low millions of dollars in legal costs. While trade disputes often affect tens of millions of US dollars in traded goods, the government agencies that oversee WTO disputes lack the resources to challenge all possible violations.⁶ Even the US and the EU—the two most frequent WTO litigants—are severely constrained by underfunding of the bureaucratic agencies that litigate trade disputes. For example, average annual spending by the Office of the US Trade Representative in 2002–2015 was approximately \$45.4 million.⁷ Only a portion of these resources were actually spent on WTO litigation because the USTR is responsible for both negotiating new trade agreements, like the Trans-Pacific Partnership, and enforcing all existing trade agreements.⁸ When the USTR was threatened with budget cuts in 2013, its top litigator warned that his office “may no longer have the funding to initiate new legal disputes, which would result in reduced enforcement of trade agreements.”⁹ Experts on European trade politics suggest that EC trade litigators face even more severe budgetary constraints (Shaffer, 2003; Greenwood, 2000).

Perhaps more importantly, WTO disputes come with high political costs. Accusing a trade partner of a violation inevitably antagonizes it, putting diplomacy at risk. Japan’s foreign affairs ministry, for instance, regularly prevents the Japanese trade ministry from filing WTO complaints against China because diplomats fear that a WTO case would exacerbate Japan’s ongoing foreign policy conflicts with China.¹⁰ Small countries may be especially prone to such fears. Elsig and

⁶See Davis (2012*a*). By one estimate, the average complainant’s relevant exports fall by about \$56 million in the run-up to a dispute, though this conceals much variation among disputes (Bown and Reynolds, 2015).

⁷Data from 2002–2007 were provided by Fred Ames, the Assistant U.S. Trade Representative for Administration, and data from 2008–2015 came from the annual Congressional Budget Report.

⁸Unfortunately, the USTR does not track expenditures based on its different activities. However, average annual spending by the USTR’s office in Geneva (where WTO litigation actually occurs) was only an average of \$6.4 million during 2008–2015.

⁹“Reif: Sequestration Could Hinder Litigation, Negotiating Efforts At USTR” *Inside US Trade*, March 1, 2013.

¹⁰Author interviews with officials from Japan’s Ministry of Economy, Trade and Industry in Tokyo on November 19, 2013. Notes on file with authors.

Stucki (2011) show how African cotton-producers were unwilling to join as third parties against the United States in *US—Upland Cotton*, let alone as co-complainants, for fear of losing foreign aid. But Shaffer (2003, 137) argues that such concerns also affect US and EU filing decisions since both WTO members must balance the benefits of trade litigation against the desire to maintain “friendlier relations” with their trading partners. The fear of upsetting longstanding relations is routinely brought up to argue against WTO challenges.¹¹

In contrast, the benefits of enforcement are often widespread. WTO rules require each member to provide Most-Favored Nation (MFN) treatment to all other members, meaning that any concessions obtained through dispute settlement must be extended to all members. Article 3.5 of the Dispute Settlement Understanding explicitly requires that “All solutions to matters formally raised under the consultation and dispute settlement provisions of the covered agreements [...] shall be consistent with those agreements”. At the margin, complainants may extract private benefits in settlements, yet the public benefits of litigation are large enough that trade scholars routinely refer to WTO dispute settlement as a public good (Bown, 2005). Countries that file WTO disputes thus find themselves often paying a private cost for a public good. This is usually presented as a positive aspect of the regime: litigation by a few (mostly powerful) countries benefits everyone, because it lifts protectionist barriers that may affect the membership as a whole. Yet we show that this feature of the trade regime also has overlooked consequences.

2.2 Diffuse Policies and Collective Action

The *diffuseness* of a trade policy affects the degree to which litigation is a public good. Regardless of their total economic effect, some trade policies, like the 1998 US restrictions on Canadian trucks, affect only one country, meaning that the policy’s impact is highly concentrated. For such concentrated policies, dispute settlement is largely a private good because a complainant internalizes most of the benefits of challenging a possible violation.¹² Other trade policies, like US corn subsidies, affect a great many countries, meaning that these policies have a highly diffuse impact. For such diffuse policies, dispute settlement more closely resembles a public good. While no WTO dispute is ever a purely private or public good, each dispute lies on the continuum between these two ideal-types. There is considerable variation in the diffuseness of policies that are challenged.

As is well-known, public goods generate collective action problems: when individual effort is needed to create a good that benefits all, each individual has incentive to free-ride on the effort of others (Olson, 1965). All else equal, as more actors benefit from a good, each individual is less

¹¹For instance, in a recent Canadian green energy case, the US was said to be wavering about joining the dispute because of “fear that a stronger stance against the Canadian program could upset the broader U.S.-Canada trade relationship.” See “USTR Cautious on Japan’s WTO Challenge Of Canadian Green Energy Law”, *Inside US Trade*, June 24, 2011.

¹²Other countries may benefit indirectly by a dispute’s spillover effects on jurisprudence (Pelc, 2014). We discuss this below.

likely to provide it. This suggests that enforcement may be most challenging in large, multilateral institutions. Many scholars have examined the special challenges posed by multilateralism, emphasizing that by including more states, multilateral institutions find it more difficult to agree on the nature and depth of cooperation (Downs, Rocke and Barsoom, 1998; Gilligan, 2004; Martin, 1992). Our analysis demonstrates a less-examined consequence of multilateral cooperation.

To preview the expectations derived from our theoretical model of WTO disputes, we put forth two main claims. First, we argue that controlling for the total trade at stake, more diffuse policies take longer to be challenged, and risk not being challenged at all. Enforcement delays are highly costly to affected countries. Each day of delay is a day during which a distortionary policy remains in place, with the economic consequences it entails. Most often, the WTO litigation process does not yield the amount of alleged harm produced by the disputed policy. Yet some WTO cases go through an arbitration process in which damages *are* calculated.¹³ To offer a rough sense of scale, the estimated average annual harm caused by each disputed policy across all such cases was USD \$622 million, or about USD \$1.7 million a day.¹⁴ The aim of the dispute settlement understanding is to minimize such distortionary effects.

Second, we argue that free-riding affects not only the likelihood of enforcement, but also which specific violations are challenged within a cooperative regime, an issue that has not been deeply explored in the study of international cooperation. When a country chooses whether to challenge a trade policy at the WTO, it must perform a cost-benefit analysis. All else equal, each country would prefer to file a stronger case, which will be more likely to succeed in litigation. If a case has a sufficiently high likelihood of success, a country will be willing to pay the cost of filing it. In the case of highly concentrated policies, the cost-benefit analysis is a simple one: a country should file if the direct benefit of litigation to the country is higher than the cost. But if others also gain from enforcement, the calculation becomes more complex, since the country will hope to free-ride on the filing by another government. Holding the total trade stake constant, diffuse policies thus face a higher barrier to enforcement: given the temptation to free-ride on the efforts of others, a potential case must have higher chance of success, all else equal, for an affected country to file it. The result is a testable implication that capitalizes on the known selection process: conditional on being filed, cases that are more diffuse should be more likely to succeed in litigation, all else equal. If the selection of disputes is not occurring on the basis we describe, then we should not expect to see any difference in legal success associated with diffuseness. Indeed, legal merit in the WTO is in no way connected to the harm a measure causes—in fact, countries have standing to challenge measures that do not harm them. Success in litigation is determined only by whether a measure is in violation of a country's WTO obligations.

¹³ These damages are assessed during Article 22.6 proceedings.

¹⁴Horn and Mavroidis (2013) data. This number is not representative of the average WTO challenge, since it is measured only for those disputes where compliance following a ruling was not forthcoming. The takeaway is simply that delayed enforcement comes at a significant cost.

Both of these claims are based on observable implications of an unobservable process: selection. An ideal research design would identify all possible trade violations and then observe which policies are challenged at the WTO. Some scholars have attempted to approximate this research design by focusing on the behavior of specific countries, and describing a prior stage in the selection of disputes. Davis and Shirato (2007) thus examine which disputes are eventually filed among the potential violations identified in the annual report of Japan’s Ministry of Economy, Trade, and Industry. Similarly, Chaudoin (2014) uses US antidumping and countervailing petitions as a pool of potential violations, and then examines which are more likely to be challenged, and at what point in time. Both of these studies have yielded valuable insights about the domestic politics of trade disputes, yet by their own admission, both push the selection process down one level, rather than accounting for it fully. The question then becomes, what process has brought these measures to the attention of, e.g., the Japanese trade ministry, and whether this process is itself prone to selection bias. Using a prior universe of cases also necessarily limits the researcher to examining individual countries, with the risk that the findings may not be generalizable to the behavior of other WTO members. In view of these methodological pros and cons, here we propose a different approach. Rather than trying to systematically control for selection effects—an inherently infeasible task given the size of the WTO and the scale and diversity of domestic trade policies—we design empirical tests that rely upon selection effects. We can only observe cases that are actually filed, so we test our theoretical argument using observable patterns of behavior that should be present in these observable cases if selection is occurring.

2.3 Why Not Coordinate?

Our argument suggests that the suboptimal provision of enforcement might be solved if states could effectively coordinate their efforts (Johns, 2012, 2015). States can, after all, communicate among themselves about their intentions to litigate. They may thus be able to coordinate in groups to challenge policies that affect them, and in this way share the overall cost of enforcement across time. However, from a theoretical perspective, what we know of collective action problems should make us pessimistic about the ability of countries to collude effectively. While states might collectively benefit from taking turns in providing enforcement or filing as co-complainants and sharing litigation costs, the incentive to free-ride remains. Any attempt at sharing the cost of enforcement is likely to suffer from the usual problems of international coordination: countries will prefer that others invest first, and the value of enforcement may vary across cases and time, complicating attempts at “enforcement-trading”.

In his detailed comparative study of EC and US trade disputes, Shaffer notes that the EC and US “rarely collaborate in WTO litigation” for numerous practical reasons, including built-up antagonism from EC-US trade disputes and differences in their legal strategies, organizational

culture, and domestic political institutions (Shaffer, 2003, 127). He writes that: “Trade officials find it difficult enough to coordinate common positions within U.S. and EC internal interagency . . . processes. The coordination of agency positions within transgovernmental contexts is all the more challenging” (134). More importantly, Shaffer highlights the political costs of litigation and argues that the EC and US recognize their strategic incentives “to ‘free-ride’ on the other’s aggressive actions. The more passive party thereby benefits from enhanced market access while retaining friendlier relations with the foreign country for other purposes” (137). Each actor prefers that the other bear the economic and political cost of enforcement.

Similarly, Steinberg (1999) discusses the challenges of EC-US cooperation on Asian trade policy. He writes: “The European Union and United States face many trade problems with third countries that if resolved on an MFN basis would have the qualities of a public good. When one transatlantic power successfully acts alone to resolve such a problem, and the resolution is on an MFN basis, the other power may simply free ride on the result” (Steinberg, 1999, 217). For example, in 1994, the US requested help from the EU in negotiating intellectual property rules with China. The EU refused to assist, forcing the US to bear the full economic and political cost of prolonged negotiations. However, after the US and China reached a successful agreement in 1995, “EU negotiators rushed to China to confirm that they would be able to free ride off the US action” by receiving the same intellectual property protections (Steinberg, 1999, 217).

There have been a few attempts at cooperation in WTO litigation. In *Japan–Alcoholic Beverages*, the EC and US agreed to jointly challenge Japanese liquor laws, but cooperation quickly collapsed when the EC and the US disagreed about how legal criteria should be applied in the case.¹⁵ In contrast, there are a few cases in which multiple states have successfully coordinated their enforcement efforts, largely through the assistance of multinational firms. For example, alcohol lobbying groups in Canada, the EU, and the US were able to successfully coordinate litigation against Korean alcohol tariffs in the 1990s. Similarly, shoemaker Nike played an important role in pressuring both the EU and the US to sequentially challenge Argentinian measures on footwear and apparel (Shaffer, 2003, 140-141). Previous scholars have emphasized the role of multinational corporations in pressuring countries to write new preferential trade agreements and influencing global regulatory standards (Manger, 2009; Sell, 2003). The examples above suggest an additional way in which these corporation can influence global governance: by coordinating enforcement across states. In this way, and as we discuss in the conclusion, transnational actors may help states overcome the free-rider problem in WTO litigation.

While these few anecdotes suggest that cross-national coordination may sometimes be possible, descriptive statistics show no systematic burden-sharing in enforcement. While some cases involve multiple complainants, these represent a small share of WTO disputes: more than 90% of all

¹⁵See Shaffer (2003, 134).

disputes affect multiple countries, yet fewer than 10% of disputes involve multiple complainants.¹⁶ Disputes that feature multiple complainants tend to challenge policies that are highly politically salient, suggesting that the dynamics of these cases may differ from the economic incentives that drive our argument.¹⁷ Additionally, these complainants tend to make legal arguments that challenge different aspects of trade policies, meaning that the precise incentives of co-complainants often differ.¹⁸

While sustained coordination over enforcement is unlikely, any such dynamics should ultimately bias our empirical tests against the effect that we identify. That is, if countries can successfully coordinate—through either alternation or cost-sharing—then we should expect that more diffuse policies would be more likely to be challenged. Such cases should thus be filed more quickly and have lower odds of legal success than cases with a less diffuse effect. Both of these implications directly contradict our empirical findings. Next, we develop the theory to derive these empirical expectations.

3 Theory

One of the central contributions of our model is that it allows us to explicitly consider the *timing* of strategic choices. When studying strategic interaction, social scientists are often interested in timing, yet the confounding technicality of timing models means that they are rarely used in political science, and even more rarely are they tested empirically.¹⁹ Here we present the basic intuitions about our causal mechanism and observable behavior. Technical readers can consult the Appendix for a more mathematical presentation.

3.1 Primitives and Structure

We present an infinite-horizon game with discrete time ($t = 1, 2, \dots$). The game begins when n countries are harmed by a new policy that disrupts their trade. We let $\tau_i > 0$ denote country i 's *trade stake*—this represents the magnitude of country i 's harm from the new policy. Since we care about decisions to enforce WTO rules, we assume the new policy is exogenous and focus on the behavior of affected countries (countries with $\tau_i > 0$).²⁰

We assume that each affected country faces some domestic pressure to challenge the new policy, but the size of this pressure changes exogenously over time. This change can come from election-timing, government composition, macroeconomic shocks, and other factors that are exogenous to our game. Formally, we assume that conditional on reaching period t , each country

¹⁶Descriptives from the data used in the analysis.

¹⁷See *US—Section 301*, the *Zeroing* disputes, or *US—Steel Safeguards* for typical multi-complainant cases.

¹⁸For example, Brazil joined Canada in its challenge of US corn subsidies six full months after Canada's initial challenge. However, when it did so, it brought legal claims over export credit guarantees against a number of agricultural products, over and above corn, thus broadening the initial challenge. See the requests for consultations in WTO DS357 for the Canadian dispute and DS365 for the Brazilian follow-up filing.

¹⁹Two notable examples of timing models in political science are Fearon (1994, 1998).

²⁰We discuss endogenous trade policy—initial decisions about whether and how to violate—in the conclusion.

i privately learns its *type*, α_{it} —this represents the domestic pressure on i to challenge the policy in period t .²¹ Then all countries must simultaneously decide whether to file a WTO dispute. If country i does not file, it receives the payoff $-\alpha_{it}\tau_i$ for period t . This is the domestic cost of not challenging a harmful policy. If no country files, the game progresses to period $t + 1$. If at least one country files, then the dispute goes to the WTO and our model ends.²²

Because we care about filing decisions, we model WTO dispute settlement in reduced form.²³ When the dispute goes to the WTO, all affected countries benefit from having the case resolved. While the complainant might lose a panel ruling, it might alternatively win a ruling or negotiate a settlement in which the policy is partly or completely removed. As discussed above, the most-favored nation principle ensures that all affected countries—including those countries that did not file—benefit from such outcomes. We let parameter $r > 0$ represent the case quality, and payoff $r\tau_i$ represent country i 's expected per period payoff from WTO dispute settlement. This parameter includes expectations about the likelihood and consequences of successful litigation. Higher quality cases are by definition more likely to yield pro-complainant rulings.

We allow the complainant to receive additional private benefits from dispute settlement. These private benefits might come from discriminatory settlements or any other indirect benefits of litigation that only the complainant receives. We let parameter $b > 0$ represent the expected *private benefits*, and payoff $b\tau_i$ represent country i 's expected per period private benefit if it files the case. Finally, we let parameter $c > 0$ represent the one-period litigation cost, and assume that countries have the discount factor $\delta \in (0, 1)$. The Appendix contains each country's expected utility for the possible infinite streams of these per period payoffs.

We initially solve the model for all possible distributions of trade stakes, generating Propositions 1–4. This approach allows us to consider how changes in positive externalities affect enforcement decisions, an issue that we return to in section 3.4. However, to identify the impact of diffuseness, we must make an assumption about the distribution of trade stakes for Propositions 5–8. Namely, we hold the total impact of the trade policy on all countries, $\tau = \sum_i \tau_i$, constant, and then assume that each country's trade stake is an equal share of the total trade stake, $\tau_i = \frac{\tau}{n}$. When there are few affected countries (small n), the overall impact of the trade policy is concentrated. However, as the number of affected countries increases, the total impact of the trade policy is spread across more countries, making it more diffuse. This approach allows us to hold the total impact of the trade policy fixed when taking comparative statics. Diffuseness reduces each individual's incentive to enforce, but also increases the number of players with some incentive to enforce. We are therefore isolating the impact of diffuseness, independent of the overall aggregate benefit of enforcement.

²¹ Assumptions about the distribution of α_{it} are in the Appendix.

²² This is an infinite-horizon game because it can go on forever if no affected country ever files, but it can also end.

²³ The conclusion discusses a model extension with a more detailed dispute settlement system.

3.2 Equilibrium Behavior and Comparative Statics

We first identify a property of the weak perfect Bayesian equilibrium for our game.²⁴

Proposition 1. *When countries are relatively impatient (δ is small), each country adopts a cut-point strategy: conditional on reaching t , high types file the case and low types don't file.*

Given our model's structure, we must constrain the discount factor to identify a reasonable equilibrium. To understand why, suppose that the countries are extremely patient (δ is large). Then an infinite stream of even small expected private benefits will outweigh the one-period litigation cost, and all countries will immediately file the case. This behavior is substantively implausible because countries rarely file WTO disputes the moment new policies appear. More plausible behavior occurs when countries are relatively impatient (δ is small) because the one-period litigation cost deters some types from filing a case.

When a country has a small trade stake, both the expected benefit of filing the case and the domestic cost of not filing are small, so a country will not want to pay the litigation cost. Not surprisingly, a country is more likely to challenge a trade policy when that policy harms the country more.

Proposition 2. *When its own trade stake (τ_i) increases, country i is more likely to file in any given period.*

However, a country's incentives are different when the trade policy causes more harm for another country. When the trade stake of another country j increases, any enforcement action by country i generates larger positive externalities. The aggregate economic benefit of enforcement increases. Country i does not directly care about the trade stake of another country j . However, country j 's trade stake indirectly affects country i by changing i 's beliefs about how j will behave. Countries want the long-term benefit of going to the WTO, but do not want to pay the short-term litigation cost. So country i is less likely to file when another country j is more likely to file, which occurs when j 's trade stake increases. Greater positive externalities can therefore reduce enforcement at the individual-level if they cause a player to believe that someone else is more likely to enforce.

Proposition 3. *When another country's trade stake (τ_j) increases, country i is less likely to file in any given period.*

Case quality, r , also affects a country's behavior. Since all countries benefit when the case is filed, the expected utility from both filing and not filing increase when case quality increases. However, the expected utility functions change at different rates. When a country files, it knows

²⁴This solution concept requires that strategies are sequentially rational and beliefs are consistent with Bayes' Rule where possible. Since types are independent across time and players choose actions simultaneously, we don't need to specify off-the-equilibrium-path beliefs.

that its payoff is increasing as r increases. However, when a country does not file, an increase in r only benefits the country if someone else files the case. So increasing a case's quality makes filing the case more attractive relative to not filing.

Proposition 4. *When the case quality increases, each country is more likely to file the case in any given period.*

Propositions 2 and 3 suggest that the distribution of harm across affected countries matters. However, neither result isolates the effect of diffusiveness because by increasing the trade stake of one country in the results above, we are also increasing the total impact of the trade policy on all countries, $\tau = \sum_i \tau_i$. We now invoke our additional assumption about the distribution of trade stakes, which is described above. That is, we assume that $\tau_i = \frac{\tau}{n}$. Increasing the number of affected countries therefore decreases each individual player's benefit from enforcement because the total trade impact is spread across more players.

When a trade policy's impact is spread across more countries, each country's individual trade stake decreases, exacerbating the collective action problem. Each country is more tempted to free-ride, and therefore is less likely to file the case itself.

Proposition 5. *When the number of affected countries increases, each country is less likely to file in any given period.*

However, this individual-level effect does not necessarily extend to the collective outcome—whether *someone* files a case. At the individual-level, diffuseness makes each country less likely to file because the total benefit of enforcement is distributed across more individuals. Yet diffuseness also increases the number of countries that want to file the case. Which effect is dominant—the individual versus the collective—depends on the model parameters. Suppose we increase the number of affected countries from n to $n+1$. This spreads the impact of the trade policy across more countries, decreasing the likelihood that one of the original n countries will file. If the litigation cost is relatively small, the new country is likely to file the case, offsetting the decrease in the likelihood that one of the original n countries will file. However, as the litigation cost grows, the new country is less likely to file, and the negative impact of diffuseness on the original n countries outweighs the effect of increasing the number of affected countries.²⁵

Proposition 6. *When the litigation cost is large and the number of affected countries increases, the overall probability that the case is filed by at least one country decreases.*

3.3 Empirical Implications

Under an ideal research design, we could identify all possible trade violations, observe which policies are challenged at the WTO, and examine whether diffuse policies are less likely to be

²⁵In the Appendix, we show this logic holds when $\alpha \sim U[0, A]$ and c is large. In the online appendix, we derive the necessary and sufficient condition on the distribution function for Propositions 6–8.

challenged than concentrated policies. However, the available data has selection effects: we can only observe cases that are actually filed. Nevertheless, our model generates two major empirical implications—concerning enforcement delay and legal outcomes—that can be tested using observed legal challenges. Rather than being constrained by selection effects in our empirical analysis, we leverage selection effects in the model to generate hypotheses about observable disputes.

First, the results above concern the likelihood that a dispute is filed in a given period. But because we have an infinite-horizon game, we can make meaningful and rigorous inferences about duration—how long countries will wait to file a case. We refer to this as “enforcement delay”. Because violations with more diffuse effects are less likely to be challenged in a given period, diffuseness increases enforcement delay in observed disputes.

Proposition 7. *In observable WTO disputes, cases that challenge more diffuse policies will, on average, have more enforcement delay, ceteris paribus.*

This finding yields insight into a collective outcome, namely, how long it takes someone to file a dispute. But it also allows us to indirectly test our arguments regarding individual-level behavior, even in the presence of selection effects. As described below, statistical methods for estimating the duration of an outcome rely upon the use of hazard ratios. In our analysis, these hazard ratios indicate the likelihood that the case is filed in a given unit of time. If we construct our analysis at the dispute-country level, we can gain leverage over decision-making by individual states. This allows us to assess the validity of Propositions 2, 3, and 5. All of these results should hold when we restrict attention to those cases that are eventually filed, and hence enter our sample.

Second, we showed that diffuseness and case quality have opposing effects on filing decisions. While diffuseness reduces the likelihood that a case is filed, legal merit increases this likelihood. When deciding whether to file a case, every country must balance the expected benefit from litigation against its expected cost. Suppose that a given country is indifferent between filing and not filing the dispute. If the number of affected countries increases, then the expected benefit of filing decreases. If we wish to offset this effect to ensure that the country remains indifferent, we must increase the quality of the case. So there is a clear selection effect: conditional on being filed, a case that challenges a diffuse policy should be of higher quality, on average, than a case that challenges a concentrated policy. Accordingly, it should be more likely to yield a pro-complainant ruling.

Proposition 8. *In observable WTO disputes, cases that challenge diffuse policies will, on average, be more likely to end in a pro-complainant ruling than cases that challenge concentrated policies, ceteris paribus.*

This argument is illustrated in Figure 1. In the top half of Figure 1, the x-axis represents case quality (r) and the y-axis represents a country’s expected benefit from filing a dispute. Proposition 4 stated that as case quality increases, a country’s expected benefit from filing increases, as

shown by the upward-sloping lines in Figure 1. When the expected benefit is positive (above the dotted line), a country will file; when the expected benefit is negative (below the dotted line), a country will not file. Proposition 5 stated that increasing the number of affected countries makes the trade policy more diffuse, reducing an individual country’s incentive to file. So the expected benefit of filing when there are four affected countries ($n = 4$) is lower than the expected benefit when there are three countries ($n = 3$), which in turn is lower than the expected benefit when there are only two countries ($n = 2$).

[Insert Figure 1 here.]

The bottom half of Figure 1 contains line graphs that show the minimum case quality needed for a country to want to file the case, $\bar{r}(n)$. When there are only two affected countries ($n = 2$), a player will file anytime that the case quality is at least as large as $\bar{r}(2)$. This is the critical value of r at which the expected utility from filing is equal to zero. Similarly, when there are three countries ($n = 3$), the critical value of case quality is $\bar{r}(3)$. Because the expected utility from filing is lower if there are three players than if there are two, the case must be better for a state to be willing to file ($\bar{r}(2) < \bar{r}(3)$). Finally, $\bar{r}(4)$ is the critical level of case quality when there are four affected countries ($n = 4$).

Because both Propositions 7 and 8 are explicitly driven by selection effects, they can be tested on observable cases. We can thus leverage the selection effects that are inherent in the data-generating process rather than being constrained by selection. Namely, we can test Proposition 7 by examining how long trade policies were in effect before being challenged at the WTO. Similarly, we can test Proposition 8 by examining the outcomes from actual WTO disputes. If the perceptions of countries are correct—that is, if higher quality cases are more likely to generate pro-complainant rulings—then cases that challenge diffuse policies should be more likely to generate legal victories for the complainant than cases that challenge concentrated policies.

3.4 Robustness

How robust are our results to changes in specific assumptions?²⁶ We argue that the diffuseness of a policy is analytically distinct from its total aggregate effect. After all, a policy can have a relatively small impact that is diffuse because it is spread across many countries. Similarly, a policy can have a very large impact that is highly concentrated on only a single country. In Propositions 5–8, we isolate the impact of diffuseness by holding the total aggregate effect of a policy constant, and then spreading this effect over more players. This reduces the trade stake of individual players. Alternatively, we might wonder about how strategic behavior changes if each individual’s trade stake stays constant, but the number of affected players increases. This model extension confounds the impact of diffuseness because it requires that the aggregate economic impact of a policy increase.

²⁶All of the model extensions discussed in this section are included in the online appendix.

Nevertheless, individuals continue to have an incentive to free-ride: each individual is less likely to enforce. However, in this model extension, the overall likelihood of enforcement increases because there is an increase in both the number of players that can enforce and the aggregate economic impact of the policy. Therefore, when we test Propositions 7–8 empirically, we must control for a policy’s total trade stake in our analysis.

The model above assumes that players can vary with respect to their trade stake in a given dispute. Yet we know that WTO members can also vary in their ability to bear the cost of litigation: some states find it easier to litigate than others. While the model above assumes that all players pay the same cost for filing a dispute, we can instead allow players to vary in both their trade stake and their cost of litigation. Not surprisingly, players who face a higher cost are less likely to file. However, all of our results continue to hold. Similarly, we might imagine that the political cost for an individual state from filing a lawsuit changes as the number of affected states changes. If there is power in numbers, then the individual cost of challenging a trade policy may decrease if the policy affects many states.²⁷ Under this alternative assumption, all of our results would continue to hold, provided that these decreases are not too large.²⁸ To the extent that there may be some cases in which these decreases are very large, our empirical tests should be biased against supporting our theory.

Might our empirical expectations also be consistent with another theoretical mechanism? Proposition 8 states that more diffuse cases will have more quality due to selection effects. But perhaps WTO panels issue biased rulings. That is, perhaps panels are more likely to rule in favor of complainants that are challenging trade policies that harm more WTO members.²⁹ Then more diffuse policies should be more likely to generate pro-complainant rulings, independent of any selection effect. This alternative explanation is not consistent with WTO law: the legality of a protectionist policy is not affected by the magnitude or the distribution of the policy’s effect. A panel simply decides whether the policy is in violation of a country’s WTO obligations.³⁰ Yet even if we assume that panels make biased rulings for non-legal reasons, this alternative explanation directly contradicts Proposition 7: if diffuseness increases the likelihood that WTO panelists support the complainant, then diffuseness would also increase the expected utility of filing the case. This implies that diffuseness would increase the rate of challenges, and reduce enforcement delay. Thus, while a panel bias argument can also yield the expectation in Proposition 8 (pro-complainant ruling), it would generate the opposite expectation of Proposition 7 (enforcement delay).

Finally, after a state files a WTO dispute, other states can participate at relatively little cost by becoming a third party. Third party participation has been examined at length in other research (Davey and Porges, 1998; Bown, 2005; Busch and Reinhardt, 2006; Kucik and Pelc, 2013;

²⁷We thank Christina Davis for this point.

²⁸The precise condition on changes in cost is specified in the online appendix.

²⁹For example, see Carrubba, Gabel and Hankla (2008).

³⁰In fact, a complainant can challenge a policy that has not harmed it.

Johns and Pelc, 2014, 2016). We draw on this research by using the number of third parties for a given case to identify some of our empirical tests below. The theoretical model above can be easily adjusted so that a state can receive some private benefit via third party participation after another state files the dispute. All of our results continue to hold, provided that the complainant gains a higher private benefit than third parties.

4 Empirics

Testing our two hypotheses requires two distinct datasets. Our first dataset is built at the dispute-country level, and includes data on the trade stake of every WTO member in every WTO dispute since 1995.³¹ Our second dataset is constructed at the dispute level, and considers the proportion of claims the complainant won in each WTO dispute. Next, we describe the data and the estimations that we use to test our two main claims, which together address the question: how does the diffuseness of protectionist trade policies impact their enforcement?

4.1 Free-riding and Enforcement Delays

To test our first hypothesis, we need a measure of ENFORCEMENT DELAY—the time it took for a complainant to file each WTO dispute. The recently coded data from Bown and Reynolds (2015) include the precise implementation date for the trade policy underlying every WTO dispute. This is the date on which the measure being challenged—be it an antidumping duty, a tariff increase, new labeling standards, or an embargo—first went into effect in the country at issue. We compare this implementation date to the date on which each WTO dispute was filed. This allows us to measure, in days, exactly how much time elapsed between the start of a policy and its challenge.

Ideally, our delay measure would “start its clock” only once a trade policy becomes prone to legal challenge. For this reason, we start the delay clock at the WTO’s date of inception even if a given measure was implemented earlier, because of the change in countries’ obligations between the GATT and the WTO period. In an extreme example, in DS162/DS1136, the EU and Japan challenged aspects of the US Antidumping Act of 1916. While the policies at issue had technically been in place for 82 years at the time of the challenge, the grounds for this dispute lay in the WTO’s new Agreement on Antidumping. For this reason, we start the clock on this and similar disputes at the point of the WTO’s inception. Yet our results for the effects of diffuseness are not only robust to, but statistically and substantively stronger, if we consider the GATT implementation dates instead, likely owing to the greater variance in delays. What we show below are thus more conservative findings, considering the WTO period. For similar reasons, we exclude from our sample any observations of countries that were not WTO members at the time of a dispute’s initiation.

³¹This amounts to a dyadic level of analysis, since a WTO dispute can only target one defendant.

We test our theoretical argument using three measures of diffuseness. The first two of these—NUMBER OF COUNTRIES AFFECTED and DISPUTED TRADE FLOWS HHI—capture economic diffuseness by examining trade flows. The third variable—MULTILATERAL POLICY—captures the legal aspect of diffuseness, by distinguishing cases where the discrimination at issue concerns only a few trading partners from those where it affects the entire membership.

Our first measure of diffuseness—NUMBER OF COUNTRIES AFFECTED—is the most straightforward. For each dispute, we identify which products are affected by the trade policy that is being challenged. We then count the number of countries with any trade at stake in the year that the complainant initiated the dispute, or the two preceding years. While some challenged policies concern products that are traded by a great number of countries, like corn, others are exported by a handful of countries, like commercial ships. In our sample, the count variable ranges from 1 to 128 countries. Given that this count is so widely distributed, we take its log.

Our first measure captures the number of countries with trade at stake, but it tells us little about the distribution of trade across these. It could be, for instance, that a dispute with 128 interested parties only has one country that exports non-trivial amounts of the disputed commodity. In such a case, as per Olson (1965), the collective action problem is diminished: we would expect the one country that has disproportionately more at stake to front the costs of enforcement for everyone. In sum, beyond the number of countries with something at stake, the distribution of trade among them matters. Our second measure of diffuseness—DISPUTED TRADE FLOWS HHI—is designed to capture this. We construct a Herfindahl-Hirschman index (HHI) measure of trade flows into the defendant country. The HHI measure is traditionally used to summarize the market structure of an industry in a single figure. It varies from 0 to 1, with 0 indicating a perfectly competitive market with a large number of small actors, and 1 indicating a monopoly with a single actor.³² If only two countries exported the same amount into the defendant country, the HHI measure would be 0.5. In our usable sample, the HHI measure varies from 0.04 to 1. The greater (smaller) this variable, the more concentrated (diffuse) trade in the disputed product is. Because these two measures are derived from the same trade data, however, they are highly negatively correlated, with a bivariate correlation of just over -0.43. To account for this, most of our estimations feature only one of these two variables at a time.

Both of these economic diffuseness variables rely on bilateral trade flow data from the World Integrated Trade Service, which is maintained by the World Bank. This inevitably leaves out non-merchandise disputes for which we cannot quantify the amount of trade. For example, when a group of WTO members successfully challenged Section 301, a piece of legislation that the

³²The HHI measure is:

$$HHi = \sum_{i=1}^n f_i^2$$

where f_i is the trade share from country i to the defendant in the product at issue, and n is the number of countries with non-zero trade to the defendant in the product at issue.

US used to coerce other countries to amend their policies, there was no underlying traded product, even as the legislation had widespread consequences on world trade.³³ We code such cases as missing for the purpose of our first two diffuseness measures—yet our third diffuseness measure allows us to analyze both merchandise and non-merchandise disputes.

At their core, trade disputes involve some form of discrimination. Broadly speaking, the WTO requires that its members abide by two standards of nondiscrimination: national treatment and most-favored nation (MFN) treatment. National treatment requires that each WTO member treat foreign imports no less favorably than the comparable domestic good. When a policy violates the national treatment standard, the entire membership is concerned: all foreign exporters are hurt, and thus all stand to benefit from a legal challenge of the policy. In contrast, most-favored nation treatment keeps WTO members from discriminating among different partners and favoring some over others. By definition, trade policies that violate the MFN standard harm only a subset of WTO members. Trade policies that violate the MFN standard therefore have a relatively concentrated impact, while policies that violate the national treatment standard have a more diffuse impact.

Our third measure of diffuseness—GLOBAL POLICY—attempts to capture this variation. This dichotomous variable comes from Bown and Reynolds (2015) and indicates whether a trade policy is “global”, meaning that it affects the entire membership, or “partial”, meaning that it affects only a subset of members. The WTO caseload is evenly split, with about 49% of the cases in our sample coded as concerning global protectionist policies, and 51% coded as concerning partial policies. Bown and Reynolds (2015) follow a simple coding rule: whenever a disputed government measure excludes any subset of countries, not because of countries’ production profiles, but because of the measure itself, it is coded as partial. Partial policies can thus concern many countries, but they do not concern the entire membership. Conversely, if all foreign countries stand to be affected by the measure if they exported the underlying good, it is coded as a global policy. US Section 301 is thus a global policy. By comparison, most antidumping disputes are coded as partial, since they usually, but not always, target a single country.³⁴ The resulting variable remains necessarily imperfect: the number of countries affected by trade measures is a continuous concept, and reducing it to a dichotomous measure conceals some of this variation. Yet it remains a useful way of thinking about the scope of a policy’s impact. In combination with our other two measures of diffuseness, it provides a fuller sense of whether the trade policy that is being challenged has a diffuse or concentrated impact. While its binary nature may conceal some variation, it also allows for straightforward comparisons. Even without running our survival model estimates, the descriptive statistics for GLOBAL POLICY are telling: on average, partial policies in our sample take 779 days to be challenged, while cases over global policies take 1,034 days—a statistically

³³DS152: *US—Section 301 Trade Act*.

³⁴The measure goes beyond simply looking at the legal clause at issue. For example, most safeguards are applied to imports from all countries, and are thus global, but some safeguards are targeted at a subset of countries, by excluding e.g. free trade agreement partners, in which case they are recorded as partial. Bown and Reynolds include such distinctions in their coding.

significant difference. If we consider the implementation date of policies stretching back into the GATT era, global policy cases take over 80% longer to be filed. The distribution of delay across these two categories of cases is also shown graphically in Figure 2, which illustrates the greater clustering of swift challenges in partial cases.

[Insert Figure 2 here.]

The relationship between our economic and legal measures of diffuseness merits discussion. In cases such as our opening example of the US measure blocking Canadian trucks, a variable indicating concentration of trade flows in the disputed product (cattle and swine) lacks relevance. The legal nature of the policy is logically prior. Yet disputes like the Canadian trucking case are rare: few trade measures truly concern a single country, even among those coded as partial. Most partial disputes affect a significant subset of the membership, in which case economic diffuseness variables remain useful indicators. We thus test our two economic diffuseness variables both on a restricted sample that includes only those disputes that challenged global policies, and on the entire caseload.

Throughout our estimations we include two control variables measuring the exact amount of trade at stake in each dispute. The first variable, OWN TRADE STAKE, captures exports of the disputed product from the country under observation into the defendant market. The second variable, ROW TRADE STAKE, measures the rest of the world's (ROW) exports of the disputed product into the defendant's market—that is, everyone but the country under observation. It follows that the sum of these two variables corresponds to the total world exports of the disputed product into the defendant market. Both trade measures are logged.

We also include a measure of legal capacity. A growing body of work shows the extent to which filing trade disputes is constrained by capacity (Busch, Reinhardt and Shaffer, 2009). Such capacity is not reducible to wealth: countries appear able to learn by doing. Specifically, the more countries participate in the dispute settlement process, the more likely they appear to challenge policies that affect them (Davis and Bermeo, 2009). The single best proxy for legal capacity is thus past experience. We code COUNTRY LEGAL CAPACITY as a running count of the number of prior cases a country has filed.

In our main estimations, we also include a battery of market size variables on both the respondent and the WTO member country level. We add a measure of logged GDP, GDP per capita, and trade dependence for both countries in each case. These are intended to capture any effect of retaliatory power or weakness on either side, or a measure of legal capacity as measured by wealth, though as above, existing work suggests that experience may be the more reliable proxy for capacity (Davis and Bermeo, 2009). Finally, we add a variable indicating the year in which a dispute was initiated, to control for any trends in time.

Using these variables, we estimate a Cox proportional hazards model, shown in Table 1.

Recall that the data are at the dispute-country level of observation: they include information not only about the country that eventually filed, but also about all the countries that did not. We thus right-censor any countries that did not challenge the violation, to account for how, had the violation not been challenged by the eventual complainant, these countries may still have done so. Note that this makes our dataset different from most survival data, since a supermajority of our observations are censored, i.e. most disputes count only one complainant. A handful of violations are challenged by more than one party, sometimes not simultaneously. Given the unusual structure of the dispute-country level data, however, later we also perform the analysis again at the dispute level.

Recall that our assessment of economic diffuseness is limited to those disputes where we can observe trade flows in the disputed product. This leaves out disputes that concern non-merchandise issues. Yet these still vary along our legal diffuseness variable, `GLOBAL POLICY`. We thus begin our analysis with a univariate Cox proportional hazards model that considers only the relationship between `GLOBAL POLICY` and the likelihood the measure is challenged in any given period, controlling for `COUNTRY LEGAL CAPACITY`, which does not require trade flows data. The estimation, shown in model 1 of Table 1, thus exploits our maximal sample. Model 2 adds our set of trade and market control variables, which restricts our analysis to non-merchandise disputes. Models 3 and 4 of Table 1 estimate the effect of our two economic measures of diffuseness, `NUMBER OF COUNTRIES AFFECTED` and `DISPUTED TRADE FLOWS HHI`, in succession. For the reasons mentioned above, the sample for both these estimations is restricted to disputes over global policies. Finally, model 5 includes all three diffuseness measures, despite what we know to be the high (negative) bivariate correlation of our two trade diffuseness variables.

[Insert Table 1 here.]

Throughout Table 1, we adjust the Cox estimation for shared frailty of the defendant country. Shared-frailty survival models are used to model within-group correlation. Accounting for shared frailty is akin to accounting for the panel aspect of time-series data. It might be that some countries have some unobserved attribute that makes their policies more likely to be swiftly challenged, apart from the amount of trade at stake, which we already account for. The salience of a country within the trade regime, its perceived reputation, or other latent qualities affecting the likelihood that a given country is challenged would thus be captured by adding this respondent-specific frailty parameter to the equation. Looking at descriptive statistics suggests that there is considerable variation in the time it takes different countries to be challenged. Some, like Japan, are swiftly taken to task (in an average of 507 days), while others, like the US, take twice as long (1,079 days). The European Union falls somewhere in between, at 826 days. It is worth noting that such descriptives are prone to selection—they do not tell us about the violations that have yet to be challenged, and that have likely been around for longest—and should thus be interpreted

with care. But they are enough to suggest the utility of the frailty parameter.

Because we begin by examining behavior at the dispute-country level, Table 1 tests Propositions 2, 3, and 5.³⁵ We convert hazard ratios into coefficients to make the results easier to read. A negative coefficient represents a decrease in the hazard function, meaning that a challenge is less likely in any given period, resulting in a longer enforcement delay.

What does Table 1 tell us? The type of legal discrimination at issue has a consistent effect: GLOBAL POLICY is associated with a significant decreased rate of legal challenge. The effect is highly substantively significant throughout. As an example, in model 1, the rate of legal challenge decreases by 39% when the legal violation at issue is a global policy. Our two economic indicators of diffuseness follow expectations. The greater the number of countries with a stake in challenging the policy, the longer such challenges take on average. Controlling for the number of countries, and the trade at stake for both the country under observation and the rest of the world, the concentration of trade flows pertaining to the challenged policy also has the expected effect: the positive coefficient on DISPUTED TRADE FLOWS HHI shows that the higher the HHI, meaning the more concentrated trade flows pertaining to the protectionist policy are across members, the shorter the expected delay before a policy is challenged. Both indicators thus support the same belief about the concentration of benefits: more diffuse policies generate a longer enforcement delay. The relationship holds whether they are included together or separately, and whether they are estimated only on global policy disputes (in models 3 and 4) or on the full sample (in model 5)—though the high correlation between our economic diffuseness measures suggests that the estimations in models 3 and 4, where they are included successively, should be viewed as more reliable. That is also where their substantive effect is highest. We graph the cumulative hazard associated with a concentrated effects policy versus a diffuse effects policy on the full sample, holding all else equal, in Figure 3.³⁶ As the figure makes clear, more diffuse policies face a considerably lower rate of challenge than more concentrated policies, yielding support for Proposition 5.

[Insert Figure 3 here.]

Our two trade stakes variables prove equally interesting. Recall that under Proposition 2, we expect that increasing a country's OWN TRADE STAKE makes the country more likely to file all else equal, reducing enforcement delay. As expected, Table 1 shows that the more trade a given country has at stake, the greater the hazard rate and the shorter the enforcement delay. In contrast, Proposition 3 suggests that increasing the trade stake of other countries will decrease a country's willingness to file, thereby increasing delay. As per Table 1, ROW TRADE STAKE is

³⁵The proof strategy for Proposition 7 can be easily applied to convert Propositions 2, 3, and 5 into statements about delay.

³⁶In Figure 3, we assume that concentrated (diffuse) policies are partial (global) policies with one standard deviation below (above) the average number of countries affected, and one standard deviation above (below) the average trade HHI.

consistently and significantly negatively related to the hazard rate (except in model 3, where it falls just short of significance), meaning that all else equal, violations where the rest of the world has more at stake result in a greater enforcement delay. This relationship also holds in a univariate estimation. This finding conveys one striking implication of the free-rider problem: more serious violations—as measured by trade flows in the rest of the world, from the point of view of each country—have a lesser rate of challenge, and thus generate longer enforcement delays, on average.³⁷ In sum, enforcement choices appear individually, if not socially, rational.

Prior legal experience increases the likelihood of a challenge, but none of the market size variables in Table 1 appear to have consistent effects. Larger economies do appear to challenge at a somewhat higher rate, while the effect of GDP per capita, once we control for legal experience, actually appears to have the opposite effect, though inconsistently so. But to be sure, both of these market variables show positive coefficients (meaning a higher rate of challenge) in univariate estimations.

We seek to ensure that the relationship we have identified between diffuseness and enforcement delay is a robust one. In Table 2, we first add fixed effects on the legal area being challenged. The disputes in our data raise 19 different legal issues, including export restrictions, antidumping and countervailing duties and intellectual property (the TRIPs Agreement), health and safety standards (Sanitary and Phytosanitary Barriers), safeguards, services, subsidies, and national treatment, among others. Figure 4, which charts the most salient legal issues in the caseload, suggests that there does seem to be some limited variation in the average enforcement delay across these. As some legal areas may be inherently more or less prone to challenge—following Kono (2006), some measures may be more “opaque”, for instance—we verify that the results are robust to the inclusion of fixed effects on the legal issue. Table 2 suggests they are. In fact, the results on our diffuseness proxies appear stronger on average than in our baseline results in Table 1. Models (1) through (4) of Table 2 include fixed effects on the legal issue, and include our diffuseness proxies simultaneously, and separately. When including NUMBER OF COUNTRIES AFFECTED and DISPUTED TRADE FLOWS HH_i by themselves, in models (3) and (4), we limit the sample to global disputes, as above.

[Insert Table 2 and Figure 4 here.]

In models (5) through (7) of Table 2, we push this analysis further still, by re-estimating our baseline model on a single legal issue, to see whether diffuseness still exerts an observable effect. We pick the most frequent legal issue across the WTO caseload, antidumping and countervailing duties. We rerun our standard estimation, successively including our three diffuseness proxies. It is worth noting that this category of cases is made up for the most part of challenges to partial policies, since antidumping and countervailing duties are targeted. Yet there remains variation on

³⁷If we construct a total trade at stake variable, it too is negatively related to the odds of filing.

our GLOBAL POLICY variable, since an important set of disputes challenges not specific remedies, but the procedures surrounding them.³⁸ Strikingly, even looking only across antidumping and countervailing disputes, we find support for our diffuseness hypothesis: as measured by any of our three proxies, more diffuse cases are challenged at a slower rate, and result in greater enforcement delays.

Accounting for each country's stake in each violation is an integral part of testing our theoretical model. Since our diffuseness measures are at the dispute level, however, it is worth testing our results at the dispute level as well. This also serves to ensure that the results are not driven by the unusual structure of the data, where a majority of observations are censored. A dispute-level test also constitutes a precise test of Proposition 7. Collapsing our data to the dispute level does away with our ROW and own trade stake measures, which we replace by the log of the TOTAL TRADE AT STAKE in the dispute. Similarly, we no longer have a variable for each country's legal capacity. We replace this by a variable that measures the legal capacity of the country with the highest experience, among those countries with trade at stake in the dispute. This is a means of conveying whether there is a country affected by the dispute that has the demonstrated capacity to challenge it.

We present the dispute-level results in Table 3. We alternate the inclusion of our three diffuseness proxies in our first three models, once again adjusting the estimates for shared frailty on the respondent country. Then, in the last three models, we replace this shared frailty parameter with fixed effects on the legal issue being challenged, and robust standard errors clustered on the common dispute. The results are remarkably consistent, even with data collapsed at the dispute level. The legal capacity variable shows little effect one way or another, likely because of its lower variation at the dispute level, compared to its country-level equivalent. Similarly, the trade stake variable, while for the most part positive, is mostly insignificant. This may be because at this level of aggregation, the trade stakes variable is more correlated with our diffuseness measures than in our dispute-country level tests. Indeed, if we run the same TOTAL TRADE AT STAKE variable in a univariate Cox model, the effect appears positive and significant at the 0.01 level. In sum, our main theoretical expectations hold well even at the dispute level of analysis, providing support for Proposition 7.

[Insert Table 3 here.]

We run a number of additional robustness checks.³⁹ One concern might be that our estimations consider the entire membership, while we know that only a subset of countries actively files legal challenges. We have included the COUNTRY LEGAL CAPACITY to account for such variation in the ability to file, but we also go a step further in the online appendix by reestimating

³⁸These include, e.g. the *Zeroing* disputes, and challenges to US statutes like *Continued Dumping and Subsidy Offset Act of 2000*, or the *AD Act of 1916*.

³⁹All these results are available in the online appendix.

our models examining only those countries that have filed a WTO complaint of their own in the past. Our results remain unchanged. We then rerun all our estimations using enforcement delays that stretch back into the GATT period, rather than starting our enforcement delay clock with the WTO's inception. We also modify this clock for agricultural disputes. Because countries were given additional time to implement their new agricultural obligations under the WTO, we start the clock for agricultural disputes at January 31, 2000, the final deadline for all countries. This is an approximate adjustment, since agricultural disputes were also being fought in the first years of the WTO. Our results remain after all these modifications. Diffuse cases result in longer enforcement delays, and the more other countries care about a violation, the less likely a given country is to challenge it swiftly.

4.2 Does the Diffuseness of Benefits from Litigation Affect Dispute Outcomes?

Having found support for our beliefs about enforcement delay, we test the implications of this selection process on the likelihood of pro-complainant rulings, as per Proposition 8. It is well known that WTO disputes display a pro-complainant bias: most rulings find some violation at play. Indeed, this fact is consistent with our theory: international trade rules are sufficiently clear that countries are able to gauge odds of legal success, and choose which cases to challenge accordingly. Yet this pro-complainant bias conceals quite a bit of variation, as complainants file a number of claims in a given dispute, allowing us to measure exactly how much of a dispute was ruled in favor of the complainant.

To construct our variable, we code the direction of every claim in every WTO dispute from 1995 to 2013.⁴⁰ This is a considerable coding exercise: complainants have brought 4,484 such claims over the WTO's history. Of these, a minority are actually ruled on,⁴¹ though some claims receive more than one finding. All told, panels have delivered 1,429 findings on 820 individual claims. We first collapse these findings at the claim level, and then collapse claims at the dispute level, to obtain the number of claims won by the complainant. We divide this number by the total number of claims filed, to obtain the proportion of claims won by the complainant.

Most rulings are appealed, and the Appellate Body (AB) frequently overturns panel rulings. Unlike panels, which rely on *ad hoc* judges, the AB is a standing body and its rulings are thus thought to have greater authority and be more attentive to jurisprudential effect than panel rulings (Bhala, 1998-1999). As a result, when assessing a dispute's legal success, we are interested in rulings "net of appeal". The resulting dependent variable, `RULING WON NET OF APPEAL`, considers the panel ruling, as modified (or not) by the AB in the case of an appeal. Keeping with

⁴⁰A "claim" is an alleged violation of a given Article or sub-article of the WTO texts. These are taken directly from the complainants' request for consultations. There is thus no room for ambiguity in the coding.

⁴¹For instance, all cases alleging national treatment discrimination make a claim under GATT III, yet the panel rarely rules on GATT III, and instead rules on a specific agreement, such as the Agreement on Sanitary and Phytosanitary Measures.

existing work, we code this measure of legal success as a binary indicator, where pro-complainant disputes are those where 90% or more of the claims were ruled pro-complainant. In our usable sample, 44% of disputes fall into this category.

We estimate the success of complainants using a Heckman selection model. The purpose is to account for the selection of cases that reach the ruling stage: about half of all disputes never make it to a ruling, and this risks biasing our results. We thus begin by estimating the odds of a ruling in a first-stage equation, and use those estimates in our second-stage outcome equation. While we argue that a selection model is called for in this case, we go on to show that the results stand when we use a simpler estimation that does not correct for selection.

Our explanatory variables of interest remain as described above in Section 4.1: two economic measures of diffuseness, NUMBER OF COUNTRIES AFFECTED and DISPUTED TRADE FLOWS HHI, and one legal measure, GLOBAL POLICY, indicating the nature of the policy. A simple descriptive statistic provides early support for our expectations: global cases result in a pro-complainant ruling nearly twice as often as partial cases (0.58 vs. 0.32), and this difference is highly statistically significant in a t-test. Of course, this relationship neither controls for confounding factors, nor accounts for selection in any way. In our model, we thus control for market size indicators—COMPLAINANT GDP (LOGGED) and DEFENDANT GDP (LOGGED)—which may exert an impact at both stages of the estimation. Along these lines, we also control for TOTAL TRADE AT STAKE, which is the log of the defendant’s total imports of the disputed product. Because this dataset is at the dispute level, here we make no distinction between own and rest-of-world trade stakes. But we do control for the LEGAL CAPACITY of the complainant, measured as above in Tables 1 and 2. Finally, we control for INITIATION YEAR to account for any change in rulings that may result from cumulative jurisprudence.

To identify our model, we use the NUMBER OF THIRD PARTIES in the room. Many studies have shown that the presence of third parties decreases the odds of settlement, and increases the odds of litigation, in part because litigants have an incentive to posture for the sake of third countries (Busch and Reinhardt, 2006; Porges, 2003; Davey and Porges, 1998; Kucik and Pelc, 2013). Yet the number of third parties in the room should not, by itself, have a direct effect on the direction of the ruling. Recall that the expectation of legal success in the WTO does not increase with the amount of harm a measure may cause. The other variables we include in the first stage selection equation are TOTAL TRADE STAKE, LEGAL CAPACITY, COMPLAINANT GDP, and DEFENDANT GDP. Anecdotal evidence suggests that in very large stakes, defendants cannot allow themselves to concede for domestic reasons without the “political cover” of an unfavorable ruling. We thus expect TOTAL TRADE STAKE to be positively related to the odds of a ruling. We cluster robust standard errors on the common dispute.

[Insert Table 4 here.]

The results are shown in Table 4. The dependent variable is a binary indicator of a pro-complainant ruling. We find broad support for Proposition 8, concerning the relationship between the diffuseness of cases and legal success. Disputes over global policies fare better on average. Additionally, the greater the number of countries with exports at stake in the policy being challenged, the more successful the case, on average. Conversely, the more concentrated trade across those countries, the worse the prospects of the case, though this negative effect falls short of significance in the final model 4, where all three indicators are included simultaneously. Yet even in that model, the three concentration variables, taken together, remain highly jointly significant.

These effects are substantively important. Looking at our first model, disputes over global policies are 79% more likely to result in a pro-complainant ruling than disputes over partial policies.⁴² More starkly still, policies that rank as diffuse on all three indicators are more than twice as likely to result in a pro-complainant ruling than the average dispute, and nearly four times as likely as a concentrated dispute.⁴³ In short, disputes where the benefits of enforcement are more concentrated appear to be worse cases, on average.

The total amount of trade at stake has a positive impact on the odds of litigation, in accordance with intuition. It also has a positive impact on the success rate of cases. Legal capacity and the litigants' market sizes appear to have little consistent effect. Most importantly, the variable that identifies the model, NUMBER OF THIRD PARTIES, bears the expected strong positive relationship with the odds of litigation. Accounting for the selection of disputes into litigation, it appears that disputes over more diffuse policies are more successful cases.

Finally, since selection models are sometimes judged to be precarious, it is worth noting that we obtain equivalent results in a model that omits selection altogether. That is, if we perform our estimation of the proportion of claims won on a single-stage model with all the variables from both stages on the right-hand side, we obtain qualitatively identical estimates. But the selection model does appear warranted. The correlation between the errors in the two equations confirms that these are not independent: a Wald test allows us to reject the null. It remains useful to verify how our expectations fare in the simplest of models, without any correction for selection. In sum, across various types of models, diffuseness appears to bear a consistent relationship with legal success: challenges of diffuse violations are simply better cases.

5 Conclusion

The enforcement of international agreements is often a public good. Yet in institutions that rely on decentralized enforcement, individuals must bear the private cost of enforcement, regardless of

⁴²Specifically, the predicted odds of a pro-complainant ruling go from 13.2% to 23.7%, keeping all other variables at their sample means.

⁴³See fn. 36, *supra*.

its aggregate benefits. In the WTO, legal challenges of highly diffuse policies approximate a public good because litigation benefits many countries. In contrast, a concentrated protectionist policy affects few countries, making enforcement a largely private good. Of course, WTO litigation is never a purely private good—even challenges of highly concentrated policies can yield some broader public benefits. Legal precedent in one case can affect the ability of all countries to successfully challenge future trade violations, and the enforcement of trade rules today may deter future violations (Bown, 2004; Davis, 2012*b*; Johns and Pelc, 2016). Nevertheless, litigation of a concentrated policy remains a largely private good, since the complainant fully internalizes most of the direct benefits of litigation. We argue that as the diffuseness of a trade policy increases, so does the incentive for an affected country to free-ride on enforcement by others, *ceteris paribus*. Put simply, more diffuse protectionism generates a more severe collective action problem that affects not only the likelihood of enforcement, but also which kinds of violations are challenged.

The evidence supports our argument’s two empirical implications. First, our theory suggests that diffuse policies should experience more enforcement delay. We measure diffuseness in three ways, through (i) the number of countries with trade at stake, (ii) the distribution of trade flows across those countries, and (iii) the legal nature of the violation at issue. On all three of these indicators, we find that more diffuse policies are associated with a longer delay between a protectionist policy’s implementation and its eventual challenge. In this way, cases over “global” legal issues, that concern a greater potential number of WTO members, are challenged at a 39% slower rate, representing considerable enforcement delay. We also find evidence supporting our expectation that while a country’s own stake in the dispute should increase its willingness to swiftly challenge a protectionist policy, when *others’* stake in the dispute rises, that country becomes *less* likely to challenge the policy in any given period. Second, our theory suggests that diffuseness increases the minimum expectation of legal success necessary for a country to want to file a dispute. We provide statistical evidence that supports this implication: cases that challenge more diffuse policies are considerably more likely to generate a pro-complainant ruling. In fact, challenges of economically diffuse policies are associated with four times greater odds of a pro-complainant ruling than challenges of concentrated policies, all else equal.

Our argument is not unique to the WTO. Most international dispute settlement bodies—including human rights bodies, investment tribunals, and regional trade agreements—rely upon decentralized enforcement. Only a handful of institutions have centralized enforcement, in which the institution itself has authority to identify and challenge possible legal violations. One avenue for future research is to look for evidence of free-riding in institutions besides the WTO. Consider the European Union (EU), which features a hybrid enforcement system, in which individuals, firms, member-states, and the EU itself all have the ability to challenge possible violations of EU law. The conventional wisdom among EU scholars is that EU law has been enforced and developed primarily through lawsuits filed by individuals and firms, rather than by states and EU

bodies (Alter, 2003; Kelemen, 2011). Yet one area of EU law is enforced almost exclusively by the EU itself: environmental regulation. Roughly 7.5% of the lawsuits brought by the Commission against member-states in 1954–2009 involved environmental policies, while only 1.8% of the lawsuits brought by individuals and firms involved the environment.⁴⁴ Kelemen’s explanation for this pattern matches our own: “most environmental regulation concerns matters of diffuse public interest . . . private parties often lack the individual incentive to commence legal action to secure enforcement” (Kelemen, 2004, 49). Johns (2016) suggests that similar patterns are apparent in other EU issue-areas.

One striking implication of our findings is that if our theoretical argument is correct, countries should want to spread the pain of treaty violations as much as possible. In the realm of international trade, countries have an incentive to use diffuse policies when possible to avoid legal challenges, rather than using trade policies with a concentrated impact, such as countervailing duties and antidumping duties. This incentive might help us understand changes in the nature of trade violations. The growing frequency and importance of WTO litigation since 1995 has coincided with an increase in the use of standards as a tool for trade protectionism (Kono, 2006; Kim, 2012). These policies—such as labeling requirements, health and safety standards, and environment regulation—have highly diffuse effects, affecting all trading partners alike. Many factors naturally go into the design of import relief, but our analysis suggests that one benefit of protectionist standards may be that these are less likely to be challenged at the WTO than policies with a more concentrated effect.

A second implication of our findings pertains to potential cases that we do not observe. The collective action problem ensures that some protectionist policies may never be challenged, or be challenged only after an overly long delay, precisely because they affect many countries. Conditional on a case being filed, the affected countries as a whole would be better off if the case were filed more quickly, thereby reducing the aggregate harm caused by trade violations (Johns, 2016). Additionally, we show that when a country challenges a concentrated policy, it often does so at the expense of legal success. This suggests that limited resources can be spent on the “wrong cases”. All else equal, the affected countries as a whole would be better off if the resources that are spent on relatively weak cases with a concentrated impact were instead spent on stronger cases with a more diffuse impact.

One way in which states may overcome such collective action problems is through coordination by transnational actors, like multinational corporations (MNCs). In international trade, most public-private partnerships occur within the state: firms and industry groups pressure government agencies to enforce trade rules, and provide legal and financial assistance to do so (Brutger, 2014). However, the growth of MNCs suggests that private transnational actors may also be effective at coordinating enforcement by multiple governments, as in the aforementioned *Korea-Alcoholic*

⁴⁴Data is from the EUROPA database, and is available from the authors.

Beverages and Argentinean apparel disputes.

These cases also demonstrate how private transnational efforts at coordination are *themselves* subject to free-riding problems. In the case of Argentinean barriers on footwear, Nike, rather than a coalition of clothing or shoe manufacturers, ultimately bore the cost of enforcement coordination. When Argentina restricted footwear imports in 1997 through a range of measures, Nike was the global behemoth of shoe exports.⁴⁵ When it bore the cost of coordinating WTO litigation, other shoe manufacturers, like Reebok and Adidas, were able to free-ride on the benefits provided by Nike. Trade associations may be able to help individual firms to coordinate, but these organizations are themselves focused on lobbying for domestic regulation and hence organized at the national level. As Shaffer (2003) documents, EC and US exporters created a Transatlantic Business Dialogue (TABD) in the late 1990s to try to better coordinate cross-national litigation at the WTO. These efforts ultimately proved unsuccessful.

This suggests that while transnational actors may be able to play a role in coordinating the enforcement actions of states, they themselves face a free-rider problem. Much as with our country-level argument, enforcement is more likely when the impact of a trade rule is concentrated on one firm—or one group of firms linked by supply chain transactions—than when a trade policy has a more diffuse effect across many firms (Johns and Wellhausen, 2016). In the case of both countries and firms, collective action affects not only delays in enforcement, but also which legal violations are challenged, and which are not. Overall, the temptation to free-ride on enforcement means that litigation by a few cannot serve the interests of all.

⁴⁵In that year, Nike controlled over 35% of the global market share in athletic footwear. Its closest competitors were Reebok and Adidas, which controlled 14.5% and 10.3% of the global market, respectively (Locke, 2002).

Appendix

We assume types α_{it} are independently and identically distributed across countries and time using distribution function F , which has full support on $[\alpha_L, \alpha_H]$, where $0 \leq \alpha_L < \frac{c}{\tau_i} < \alpha_H$ for every i . Let ρ_{-i} denote i 's belief about the probability that no other country will file. Let V_i denote i 's continuation value. Then conditional on reaching t :

$$\begin{aligned} EU_{it}(\text{file}|\alpha_{it}, \tau_i) &= \frac{\delta}{1-\delta} (r+b) \tau_i - c \\ EU_{it}(\text{don't file}|\alpha_{it}, \tau_i) &= -\alpha_{it} \tau_i + (1-\rho_{-i}) \frac{\delta}{1-\delta} r \tau_i + \rho_{-i} \delta V_i \end{aligned}$$

Equilibrium Behavior

Proof of Proposition 1. Country i will file iff:

$$\begin{aligned} \frac{\delta}{1-\delta} (r+b) \tau_i - c &\geq -\alpha_{it} \tau_i + (1-\rho_{-i}) \frac{\delta}{1-\delta} r \tau_i + \rho_{-i} \delta V_i \\ \Leftrightarrow \alpha_{it} &\geq \frac{c}{\tau_i} - \frac{\delta}{1-\delta} b - \rho_{-i} \frac{\delta}{1-\delta} r + \frac{\delta \rho_{-i}}{\tau_i} V_i \equiv \bar{\alpha}_i \end{aligned} \quad (1)$$

Equilibrium behavior is therefore monotonic and i 's best response function is characterized by $\bar{\alpha}_i$ in (1). So:

$$\rho_i = \Pr(\alpha_{it} < \bar{\alpha}_i) = F(\bar{\alpha}_i) \quad \text{and} \quad \rho = \prod_{k=1}^n \rho_k = \prod_{k=1}^n F(\bar{\alpha}_k) \quad \text{and} \quad \rho_{-i} = \prod_{j \neq i} \rho_j = \frac{\prod_{k=1}^n F(\bar{\alpha}_k)}{F(\bar{\alpha}_i)}$$

In an interior equilibrium $\bar{\alpha}_i \in (\alpha_L, \alpha_H)$ for every i , and:

$$\begin{aligned} V_i &= \int_{\alpha_L}^{\bar{\alpha}_i} \left[-\alpha \tau_i + (1-\rho_{-i}) \frac{\delta}{1-\delta} r \tau_i + \rho_{-i} \delta V_i \right] f(\alpha) d\alpha + \int_{\bar{\alpha}_i}^{\alpha_H} \left[\frac{\delta}{1-\delta} (r+b) \tau_i - c \right] f(\alpha) d\alpha \\ &= \frac{1}{1-\delta \rho} \left[(1-\rho) \frac{\delta}{1-\delta} r \tau_i - (1-\rho_i) \left(c - \frac{\delta}{1-\delta} b \tau_i \right) - \tau_i \int_{\alpha_L}^{\bar{\alpha}_i} \alpha f(\alpha) d\alpha \right] \end{aligned} \quad (2)$$

Substituting (2) into (1) shows that $\bar{\alpha}_i$ is defined by:

$$\Psi^i \equiv \bar{\alpha}_i (1-\delta \rho) - (1-\delta \rho_{-i}) \left(\frac{c}{\tau_i} - \frac{\delta}{1-\delta} b \right) + \delta \rho_{-i} \left(r + \int_{\alpha_L}^{\bar{\alpha}_i} \alpha f(\alpha) d\alpha \right) = 0$$

To see that this best response function can generate an interior equilibrium, note that $\Psi_{\bar{\alpha}_i}^i = 1 - \delta \rho > 0$. Because Ψ^i is strictly increasing in $\bar{\alpha}_i$, if there exists an $\bar{\alpha}_i$ that satisfies $\Psi^i(\bar{\alpha}_i) = 0$, this value is unique. Also:

$$\lim_{\delta \rightarrow 0} \Psi^i = \bar{\alpha}_i - \frac{c}{\tau_i} = 0 \quad \Leftrightarrow \quad \lim_{\delta \rightarrow 0} \bar{\alpha}_i = \frac{c}{\tau_i}$$

Recall that by assumption, $\frac{c}{\tau_i} \in (\alpha_L, \alpha_H)$ and α has full support over $[\alpha_L, \alpha_H]$. So i has a unique interior cutpoint, $\bar{\alpha}_i \in (\alpha_L, \alpha_H)$, for small $\delta > 0$. Since this holds for any i , there exists a Bayesian Nash equilibrium in which strategies are defined by the system:

$$\begin{aligned} \Psi^1(\bar{\alpha}) &= 0 \\ \Psi^2(\bar{\alpha}) &= 0 \\ &\dots \dots \dots \\ \Psi^n(\bar{\alpha}) &= 0 \end{aligned}$$

where $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \dots, \bar{\alpha}_n)$. □

Note that Ψ^i is continuously differentiable in its arguments, $\Psi_{\alpha_i}^i = 1 - \delta\rho > 0$, and:

$$\Psi_{\alpha_j}^i = \delta f(\bar{\alpha}_j) \frac{\rho}{\rho_i \rho_j} \left[\frac{c}{\tau_i} - \frac{\delta}{1-\delta} b + r + \int_{\alpha_L}^{\bar{\alpha}_i} \alpha f(\alpha) d\alpha - \bar{\alpha}_i \rho_i \right]$$

Therefore, the Jacobian matrix is nonsingular for small δ :

$$\mathbf{J} = \begin{bmatrix} \Psi_{\alpha_1}^1 & \cdots & \Psi_{\alpha_n}^1 \\ \cdots & \cdots & \cdots \\ \Psi_{\alpha_1}^n & \cdots & \Psi_{\alpha_n}^n \end{bmatrix} \Rightarrow \lim_{\delta \rightarrow 0} \det(\mathbf{J}) = \det(\mathbf{I}) = 1 > 0$$

where \mathbf{I} is the identity matrix. We can therefore use the implicit function theorem.

Proof of Proposition 2. Because the indexing is arbitrary, we solve for $\frac{\partial \bar{\alpha}_1}{\partial \tau_1}$. By the implicit function theorem, $\frac{\partial \bar{\alpha}_1}{\partial \tau_1} = \frac{-\det(\mathbf{B})}{\det(\mathbf{J})}$ where:

$$\mathbf{B} = \begin{bmatrix} \Psi_{\tau_1}^1 & \Psi_{\alpha_2}^1 & \cdots & \Psi_{\alpha_n}^1 \\ \Psi_{\tau_1}^2 & \Psi_{\alpha_2}^2 & \cdots & \Psi_{\alpha_n}^2 \\ \cdots & \cdots & \cdots & \cdots \\ \Psi_{\tau_1}^n & \Psi_{\alpha_2}^n & \cdots & \Psi_{\alpha_n}^n \end{bmatrix}$$

The online appendix shows: $\lim_{\delta \rightarrow 0} \det(\mathbf{B}) > 0$. This means $\frac{\partial \bar{\alpha}_1}{\partial \tau_1} < 0$ for small δ . \square

Proof of Proposition 3. Because the indexing is arbitrary, we solve for $\frac{\partial \bar{\alpha}_1}{\partial \tau_n}$. By the implicit function theorem, $\frac{\partial \bar{\alpha}_1}{\partial \tau_n} = \frac{-\det(\mathbf{C})}{\det(\mathbf{J})}$ where:

$$\mathbf{C} = \begin{bmatrix} \Psi_{\tau_n}^1 & \Psi_{\alpha_2}^1 & \cdots & \Psi_{\alpha_n}^1 \\ \Psi_{\tau_n}^2 & \Psi_{\alpha_2}^2 & \cdots & \Psi_{\alpha_n}^2 \\ \cdots & \cdots & \cdots & \cdots \\ \Psi_{\tau_n}^n & \Psi_{\alpha_2}^n & \cdots & \Psi_{\alpha_n}^n \end{bmatrix}$$

The online appendix shows: $\lim_{\delta \rightarrow 0} \det(\mathbf{C}) < 0$. This means $\frac{\partial \bar{\alpha}_1}{\partial \tau_n} > 0$ for small δ . \square

Proof of Proposition 4. Because the indexing is arbitrary, we solve for $\frac{\partial \bar{\alpha}_1}{\partial r}$. By the implicit function theorem, $\frac{\partial \bar{\alpha}_1}{\partial r} = \frac{-\det(\mathbf{G})}{\det(\mathbf{J})}$ where:

$$\mathbf{G} = \begin{bmatrix} \Psi_r^1 & \Psi_{\alpha_2}^1 & \cdots & \Psi_{\alpha_n}^1 \\ \Psi_r^2 & \Psi_{\alpha_2}^2 & \cdots & \Psi_{\alpha_n}^2 \\ \cdots & \cdots & \cdots & \cdots \\ \Psi_r^n & \Psi_{\alpha_2}^n & \cdots & \Psi_{\alpha_n}^n \end{bmatrix}$$

The online appendix shows: $\lim_{\delta \rightarrow 0} \det(\mathbf{G}) > 0$. This means $\frac{\partial \bar{\alpha}_1}{\partial r} < 0$ for small δ . \square

Proof of Proposition 5. Suppose trade stakes are $\tau_i = \frac{\tau}{n}$. By Proposition 1, there exists a unique Bayesian Nash equilibrium for small $\delta > 0$ in which each country chooses an interior cutpoint $\bar{\alpha}_i \in (\alpha_L, \alpha_H)$. When $\tau_i = \frac{\tau}{n}$, the system of $\Psi^i(\bar{\alpha})$ -equations simplifies to one equation with one endogenous variable, $\bar{\alpha}_n$. Let ρ_n denote the *ex ante* probability that an arbitrary country in the n -country game does not file in a given period. In equilibrium, the common cutpoint, $\bar{\alpha}_n$, is defined by:

$$\Psi^n = \bar{\alpha}_n [1 - \delta(\rho_n)^n] - [1 - \delta(\rho_n)^{n-1}] \left(\frac{cn}{\tau} - \frac{\delta}{1-\delta} b \right) + \delta(\rho_n)^{n-1} \left(r + \int_{\alpha_L}^{\bar{\alpha}_n} \alpha f(\alpha) d\alpha \right) = 0$$

So: $\lim_{\delta \rightarrow 0} \bar{\alpha}_n = \frac{cn}{\tau}$. Similarly, the common cutpoint for the $(n+1)$ -country game, $\bar{\alpha}_{n+1}$, is

defined by:

$$\begin{aligned}\Psi^{n+1} &= \bar{\alpha}_{n+1} \left[1 - \delta (\rho_{n+1})^{n+1} \right] - [1 - \delta (\rho_{n+1})^n] \left[\frac{c(n+1)}{\tau} - \frac{\delta}{1-\delta} b \right] \\ &\quad + \delta (\rho_{n+1})^n \left(r + \int_{\alpha_L}^{\bar{\alpha}_{n+1}} \alpha f(\alpha) d\alpha \right) = 0\end{aligned}$$

So: $\lim_{\delta \rightarrow 0} \bar{\alpha}_{n+1} = \frac{c(n+1)}{\tau}$. Then: $\lim_{\delta \rightarrow 0} \bar{\alpha}_n < \lim_{\delta \rightarrow 0} \bar{\alpha}_{n+1}$. \square

Proof of Proposition 6. Suppose trade stakes are $\tau_i = \frac{\tau}{n}$ and $\alpha \sim U[0, A]$. Conditional on reaching t , the probability at least one country files in period t when there are n countries is: $1 - (\rho_n)^n = 1 - F(\bar{\alpha}_n)^n$. For small $\delta > 0$, this probability is decreasing in n iff:

$$F(\bar{\alpha}_n)^n < F(\bar{\alpha}_{n+1})^{n+1} \Leftrightarrow \left(\frac{cn}{A\tau} \right)^n < \left[\frac{c(n+1)}{A\tau} \right]^{n+1} \Leftrightarrow \frac{A\tau n^n}{(n+1)^{n+1}} < c$$

\square

Proof of Proposition 7. Suppose trade stakes are $\tau_i = \frac{\tau}{n}$. For period t , suppose there is a “failure” if no one files, and a “success” if at least one country files. Let X denote the number of time periods until the first success. Then X is a geometric random variable and: $E[X|n] = \sum_{t=1}^{\infty} t [(\rho_n)^n]^{t-1} [1 - (\rho_n)^n] = \frac{1}{1 - (\rho_n)^n}$. Therefore:

$$E[X|n] < E[X|n+1] \Leftrightarrow \frac{1}{1 - (\rho_n)^n} < \frac{1}{1 - (\rho_{n+1})^{n+1}} \Leftrightarrow (\rho_n)^n < (\rho_{n+1})^{n+1}$$

This holds whenever Proposition 6 holds. \square

Proof of Proposition 8. Suppose trade stakes are $\tau_i = \frac{\tau}{n}$. By the Proof of Proposition 1, the marginal benefit for type α_i from filing when there are n countries is:

$$\Psi^n(\alpha_i) = \alpha_i [1 - \delta (\rho_n)^n] - [1 - \delta (\rho_n)^{n-1}] \left(\frac{cn}{\tau} - \frac{\delta}{1-\delta} b \right) + \delta (\rho_n)^{n-1} \left(r + \int_{\alpha_L}^{\alpha_i} x f(x) dx \right)$$

Then: $\Psi_r^n(\alpha_i) = \delta (\rho_n)^{n-1} > 0$, $\lim_{r \rightarrow \infty} \Psi^n(\alpha_i) = \infty > 0$, and $\lim_{\delta \rightarrow 0} \Psi^n(\alpha_i|r=0) = \alpha_i - \frac{cn}{\tau} < 0 \Leftrightarrow \alpha_i < \frac{cn}{\tau}$. So for small δ and high α_i -values, i always files, regardless of r . But for small δ and low α_i -values, the intermediate value theorem ensures that there exists a unique critical value $\bar{r}(\alpha_i, n) > 0$ such that $\Psi^n(\alpha_i|\bar{r}(\alpha_i, n)) = 0$. Type α_i files iff $\bar{r}(\alpha_i, n) \leq r$. Also:

$$\begin{aligned}\Psi^{n+1}(\alpha_i) &= \alpha_i [1 - \delta (\rho_{n+1})^{n+1}] - [1 - \delta (\rho_{n+1})^n] \left(\frac{c(n+1)}{\tau} - \frac{\delta}{1-\delta} b \right) \\ &\quad + \delta (\rho_{n+1})^n \left(r + \int_{\alpha_L}^{\alpha_i} x f(x) dx \right)\end{aligned}$$

So $\lim_{\delta \rightarrow 0} [\Psi^n(\alpha_i) - \Psi^{n+1}(\alpha_i)] = \frac{c}{\tau} > 0$ and:

$$\Psi^n(\alpha_i|\bar{r}(\alpha_i, n)) = 0 = \Psi^{n+1}(\alpha_i|\bar{r}(\alpha_i, n+1)) < \Psi^n(\alpha_i|\bar{r}(\alpha_i, n+1))$$

which implies $\bar{r}(\alpha_i, n) < \bar{r}(\alpha_i, n+1)$. \square

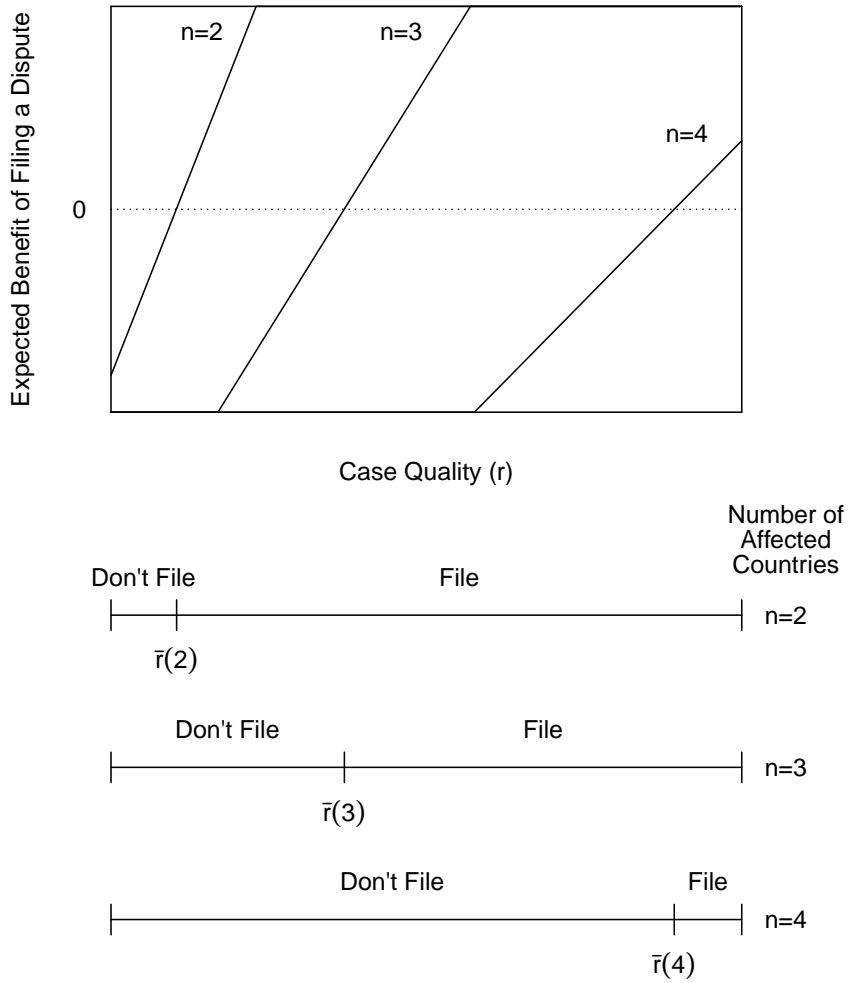
References

- Abbott, Kenneth W., Robert O. Keohane, Andrew Moravcsik, Anne-Marie Slaughter and Duncan Snidal. 2000. "The Concept of Legalization." *International Organization* 54:401–419.
- Alter, Karen J. 2003. "Resolving or Exacerbating Disputes? The WTO's New Dispute Resolution System." *International Affairs* 79:783–800.
- 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.
- Bechtel, Michael M and Thomas Sattler. 2015. "What is Litigation in the World Trade Organization Worth?" *International Organization* 69(02):375–403.
- Bhala, Raj. 1998-1999. "The Myth about Stare Decisis and International Trade Law (Part One of a Trilogy)." *American University International Law Review* 14:845–956.
- Bown, Chad P. 2004. "On the Economic Success of GATT/WTO Dispute Settlement." *The Review of Economics and Statistics* 86(3):811–823.
- Bown, Chad P. 2005. "Participation in WTO Dispute Settlement: Complainants Interested Parties, and Free Riders." *World Bank Economic Review* 19:287–310.
- Bown, Chad P. and Bernard M. Hoekman. 2005. "WTO Dispute Settlement and the Missing Developing Country Cases: Engaging the Private Sector." *Journal of International Economic Law* 8:861–890.
- Bown, Chad P. and Kara M. Reynolds. 2015. "Trade Flows and Trade Disputes." *The Review of International Organizations* 10(2):145–177.
- Brutger, Ryan. 2014. "Screening for Success: The Effect of Firm Signaling on WTO Case Selection." Princeton University Working Paper.
- Busch, Marc L. and Eric Reinhardt. 2006. "Three's a Crowd: Third Parties and WTO Dispute Settlement." *World Politics* 58:446–477.
- Busch, Marc L., Eric Reinhardt and Gregory Shaffer. 2009. "Does Legal Capacity Matter? Explaining Dispute Initiation and Antidumping Actions in the WTO." *World Trade Review* 8:559–577.
- Carrubba, Clifford J., Matthew Gabel and Charles Hankla. 2008. "Judicial Behavior under Political Constraints: Evidence from the European Court of Justice." *American Political Science Review* 102:435–452.
- Chaudoin, Stephen. 2014. "Audience Features and the Strategic Timing of Trade Disputes." *International Organization* 68(4):877–911.
- Davey, William J. and Amelia Porges. 1998. "Comments on Performance of the System I: Consultations and Deterrence." *International Lawyer* 32:695–707.
- Davis, Christina. 2012a. *Why Adjudicate? Enforcing Trade Rules*. Princeton University Press.
- Davis, Christina L. 2012b. "WTO Adjudication as a Tool for Conflict Management." Princeton University Working Paper.
- Davis, Christina L. and Sarah Blodgett Bermeo. 2009. "Who Files? Developing Country Participation in GATT/WTO Adjudication." *The Journal of Politics* 71:1033–1049.
- Davis, Christina L and Yuki Shirato. 2007. "Firms, Governments, and WTO Adjudication: Japan's Selection of WTO Disputes." *World Politics* 59(02):274–313.
- Downs, George W., David M. Rocke and Peter N. Barsoom. 1998. "Managing the Evolution of Multilateralism." *International Organization* 52:397–419.
- Elsig, Manfred and Philipp Stucki. 2011. "Low-Income Developing Countries and WTO Litigation: Why Wake Up the Sleeping Dog?" *Review of International Political Economy* 19:292–316.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88:577–592.

- Fearon, James D. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52:269–305.
- Gilligan, Michael J. 2004. "Is There a Broader-Deeper Trade-off in International Multilateral Agreements?" *International Organization* 58:459–84.
- Greenwood, Justin. 2000. Organized Business and the European Union. In *Organized Business and the New Global Order*, ed. Justin Greenwood and Henry Jacek. St. Martin's Press pp. 77–98.
- Guzman, Andrew T and Beth A Simmons. 2005. "Power Plays and Capacity Constraints: the Selection of Defendants in World Trade Organization Disputes." *The Journal of Legal Studies* 34(2):557–598.
- Johns, Leslie. 2012. "Courts as Coordinators: Endogenous Enforcement and Jurisdiction in Adjudication." *Journal of Conflict Resolution* 56:257–289.
- Johns, Leslie. 2015. *Strengthening International Courts: The Hidden Costs of Legalization*. University of Michigan Press.
- Johns, Leslie. 2016. "The Design of Enforcement: Collective Action and the Enforcement of International Law." UCLA Working Paper.
- Johns, Leslie and Krzysztof J. Pelc. 2014. "Who Gets to Be in the Room? Manipulating Participation in WTO Disputes." *International Organization* 68:663–699.
- Johns, Leslie and Krzysztof J. Pelc. 2016. "Fear of Crowds in WTO Disputes: Why Don't More Countries Participate?" *Journal of Politics* 78:88–104.
- Johns, Leslie and Rachel Wellhausen. 2016. "Under One Roof: Supply Chains and the Protection of Foreign Investment." *American Political Science Review* 110:31–51.
- Kelemen, R. Daniel. 2004. *The Rules of Federalism: Institutions and Regulatory Politics in the EU and Beyond*. Harvard University Press.
- Kelemen, R. Daniel. 2011. *Eurolegalism: The Transformation of Law and Regulation in the European Union*. Harvard University Press.
- Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press.
- Keohane, Robert O. 1986. "Reciprocity in International Relations." *International Organization* 40:1–27.
- Kim, Moonhawk. 2012. "Disguised Protectionism and Linkages to the GATT/WTO." *World Politics* 64(3):426–475.
- Kindleberger, Charles P. 1973. *The World in Depression, 1929-1939*. University of California Press.
- Kono, Daniel Y. 2006. "Optimal Obfuscation: Democracy and Trade Policy Transparency." *American Political Science Review* 100(3):369–384.
- Kucik, Jeffrey and Krzysztof J Pelc. 2013. "Measuring The Cost Of Privacy: A Look At The Distributional Effects Of Private Bargaining." *British Journal of Political Science* pp. 1–29.
- Locke, Richard M. 2002. "The Promise and Perils of Globalization: The Case of Nike." MIT Industrial Performance Center Working Paper 02-008.
- Manger, Mark S. 2009. *Investing in Protection: The Politics of Preferential Trade Agreements between North and South*. Cambridge University Press.
- Martin, Lisa L. 1992. "Interests, Power, and Multilateralism." *International Organization* 46:765–792.
- Martin, Lisa L. 2013. Against Compliance. In *Interdisciplinary Perspectives on International Law and International Relations*, ed. Jeffrey L. Dunoff and Mark A. Pollack. Cambridge University Press pp. 591–610.

- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.
- Oye, Kenneth A. 1986. *Cooperation Under Anarchy*. Princeton University Press.
- Pelc, Krzysztof. 2014. "The Politics of Precedent in International Law: A Social Network Application." *American Political Science Review* 108:547–564.
- Porges, Amelia. 2003. "Settling WTO Disputes: What Do Litigation Models Tell Us?" *Ohio State Journal on Dispute Resolution* 19:141–184.
- Rosendorff, B. Peter. 2005. "Stability and Rigidity: Politics and the Design of the WTO's Dispute Resolution Procedure." *American Political Science Review* 99:389–400.
- Sell, Susan K. 2003. *Private Power, Public Law: The Globalization of Intellectual Property Rights*. Cambridge University Press.
- Shaffer, Gregory C. 2003. *Defending Interests: Public-Private Partnerships in WTO Litigation*. Brookings Institution Press.
- Steinberg, Richard H. 1999. The Prospects for Partnership: Overcoming Obstacles to Transatlantic Trade Policy Cooperation in Asia. In *Partners or Competitors? The Prospects for U.S.–European Cooperation on Asian Trade*, ed. Richard H. Steinberg and Bruce Stokes. Rowman & Littlefield Publishers pp. 213–250.

Figure 1: Diffuseness Increases the Quality of Observed Cases



Note: Figure created from simulations of equilibrium behavior in R.

Figure 2: Distribution of Enforcement Delay

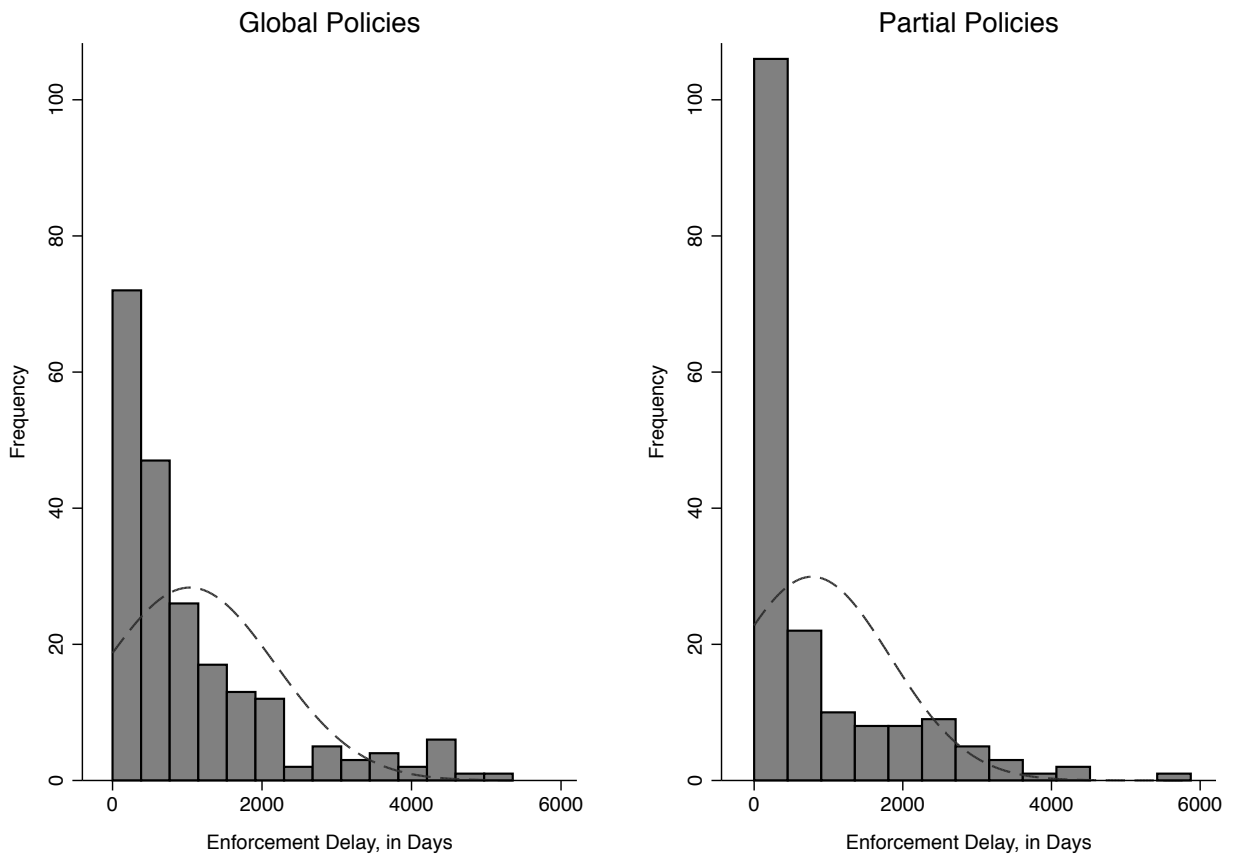


Figure 3: Cox Proportional Hazards Regression

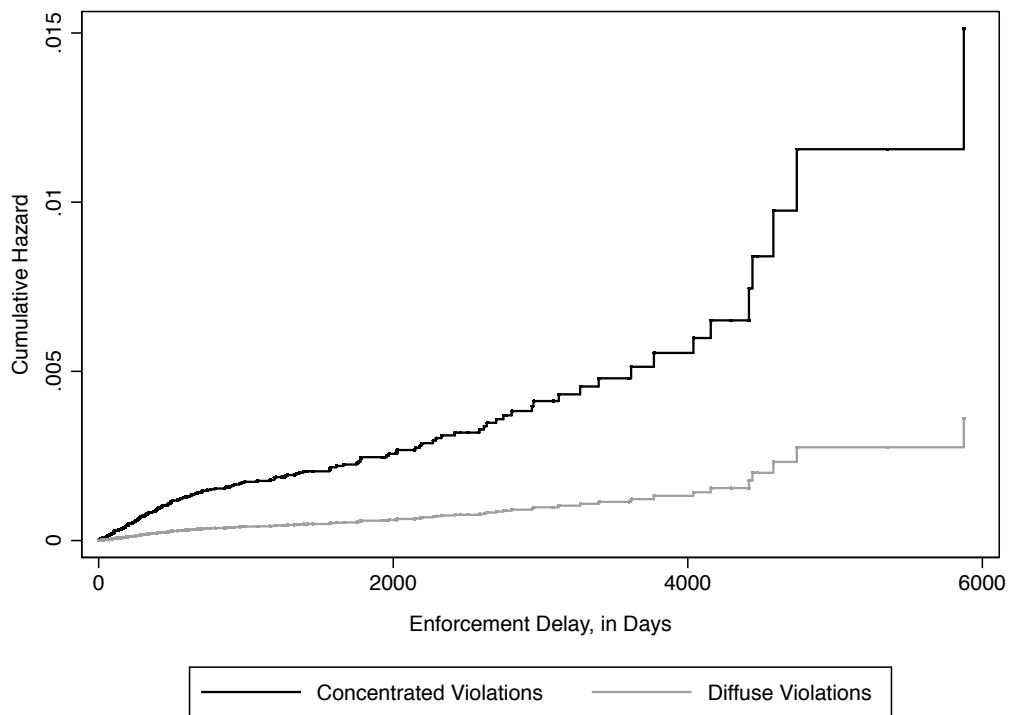


Figure 4: Enforcement Delay by Legal Issue

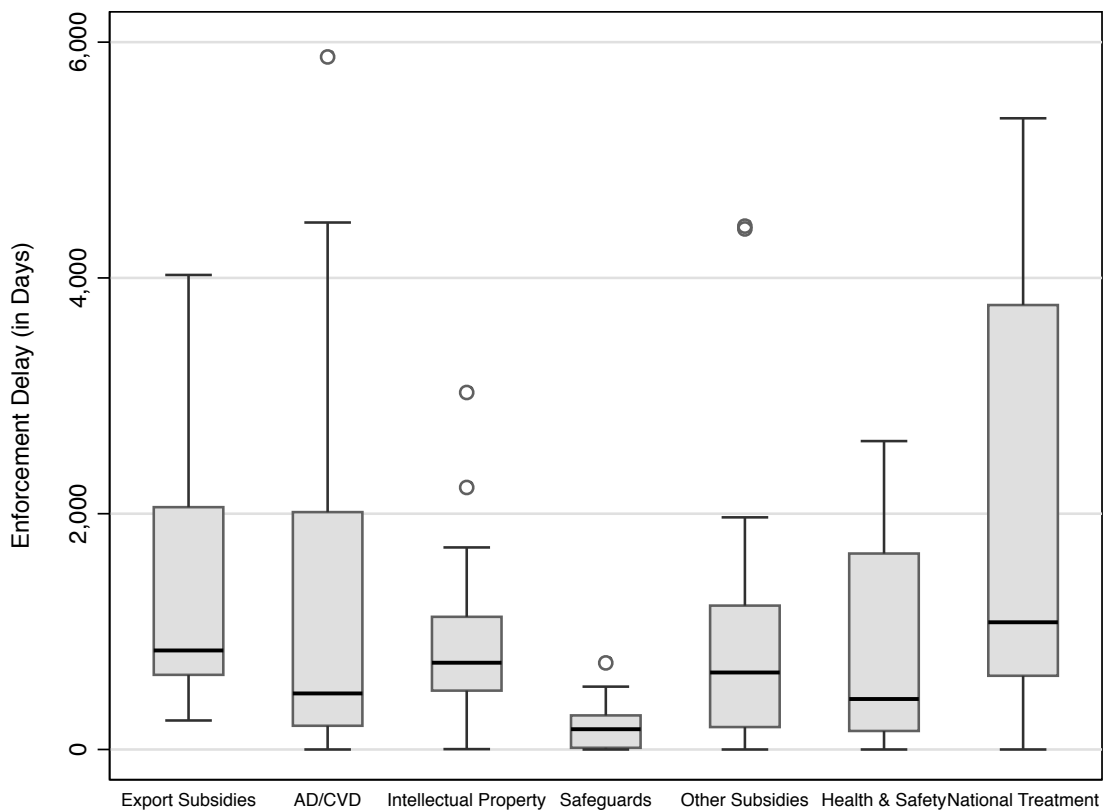


Table 1: Diffuseness of Violations and the Rate of Legal Challenge

	(1)	(2)	(3)	(4)	(5)
Global Policy	-0.49*** (0.15)	-0.65*** (0.18)			-0.39** (0.18)
Number of Countries Affected (log)			-1.29*** (0.18)		-0.66*** (0.14)
Disputed Trade Flows HHi				1.93** (0.92)	0.81* (0.46)
Own Stake (log)		0.43*** (0.03)	0.32*** (0.04)	0.29*** (0.04)	0.41*** (0.03)
ROW Trade Stake (log)		-0.32*** (0.03)	-0.04 (0.06)	-0.18*** (0.05)	-0.23*** (0.04)
Complainant Legal Capacity	0.04*** (0.00)	0.01** (0.00)	0.01* (0.01)	0.01** (0.01)	0.01** (0.00)
Country GDP (log)		0.15** (0.06)	0.16 (0.10)	0.18* (0.10)	0.13** (0.06)
Country GDP/cap (log)		-0.20*** (0.07)	-0.07 (0.12)	-0.03 (0.12)	-0.22*** (0.07)
Country Trade Dependence		-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Respondent GDP (log)		-0.24** (0.10)	0.08 (0.12)	-0.07 (0.14)	-0.09 (0.09)
Respondent GDP/cap (log)		0.08 (0.16)	-0.21 (0.14)	-0.16 (0.20)	-0.03 (0.14)
Respondent Trade Dependence		0.00 (0.00)	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)
Initiation Year	-0.14*** (0.02)	-0.10*** (0.02)	-0.21*** (0.04)	-0.19*** (0.04)	-0.11*** (0.02)
Disputes	All	All	Global	Global	All
N	15,747	14,944	7,072	7,072	14,894

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Cox Proportional Hazards estimates with shared frailty parameter on respondent country. Sample in models (3) and (4) limited to GLOBAL disputes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2: Diffuseness of Violations and the Rate of Challenge: Alternative Estimations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Global Policy	-0.93*** (0.24)	-0.91*** (0.24)			-2.71*** (0.79)		
Number of Countries Affected (log)	-0.72*** (0.14)		-1.20*** (0.21)			-0.55*** (0.20)	
Disputed Trade Flows HHI	0.98** (0.48)			1.92** (0.94)			1.59*** (0.52)
Own Trade Stake (log)	0.43*** (0.03)	0.45*** (0.03)	0.35*** (0.04)	0.33*** (0.04)	0.72*** (0.08)	0.69*** (0.08)	0.63*** (0.07)
ROW Trade Stake (log)	-0.23*** (0.04)	-0.34*** (0.03)	-0.04 (0.06)	-0.20*** (0.05)	-0.65*** (0.08)	-0.59*** (0.08)	-0.58*** (0.07)
Country Legal Capacity	0.01*** (0.00)	0.01*** (0.00)	0.02*** (0.00)	0.02*** (0.00)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
Initiation Year	-0.10*** (0.02)	-0.10*** (0.02)	-0.20*** (0.05)	-0.15*** (0.04)	-0.07** (0.03)	-0.07** (0.03)	-0.06 (0.03)
Disputes	All	All	Global	Global	AD/CVD	AD/CVD	AD/CVD
N	15,622	15,674	7,448	7,448	5,705	5,705	5,705

Cox Proportional Hazards estimates. Models (1) through (4) include shared frailty parameter on respondent country, and fixed effects on legal issue. Models (5) through (7) include robust clustered standard errors on common dispute, with the sample limited to antidumping or countervailing duties cases. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3: Diffuseness of Violations and the Rate of Challenge: Dispute-Level

	(1)	(2)	(3)	(4)	(5)	(6)
Global Policy	-0.65*** (0.12)			-0.51** (0.21)		
Number of Countries Affected (log)		-0.25*** (0.09)			-0.23** (0.09)	
Disputed Trade Flows HHi			0.65** (0.30)			0.91*** (0.29)
Total Trade at Stake (log)		0.04** (0.02)	0.02 (0.02)	-0.00 (0.01)	0.03 (0.02)	0.01 (0.01)
Legal Capacity Of Affected Members		0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)
Initiation Year	-0.01 (0.01)	-0.02 (0.02)	-0.01 (0.03)	0.01 (0.03)	-0.00 (0.03)	0.01 (0.02)
N	360	265	265	265	265	265

Cox Proportional Hazards estimates at the dispute level. Shared frailty parameter on respondent country in models (1) through (3). Robust standard errors clustered on common dispute in Columns (4) through (6), with fixed effects on legal issue. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4: Concentration of Benefits and Legal Success

	(1)	(2)	(3)	(4)
RULING WON NET OF APPEAL (2nd Stage eq.)				
Global Policy	0.18*			0.17*
	(0.10)			(0.09)
Number of Interested Countries (log)			0.07**	0.07**
			(0.04)	(0.03)
Disputed Trade Flows HHi		-0.81***		-0.25
		(0.26)		(0.20)
Total Trade at Stake (log)	0.02**	0.03***	0.02**	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
Legal Capacity (log)	0.04	0.00	0.03	0.05
	(0.06)	(0.11)	(0.10)	(0.06)
Complainant GDP (log)	-0.00	0.01	-0.00	-0.00
	(0.03)	(0.06)	(0.05)	(0.03)
Defendant GDP (log)	-0.02	-0.11***	-0.07**	-0.03
	(0.03)	(0.04)	(0.03)	(0.03)
Dispute Year	-0.01	-0.00	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Constant	14.07	4.50	22.57	20.42
	(20.92)	(25.58)	(26.96)	(20.43)
DISPUTE GOES TO RULING (1st Stage eq.)				
Number of Third Parties	0.20***	0.17***	0.18***	0.20***
	(0.03)	(0.05)	(0.04)	(0.03)
Total Trade at Stake (log)	0.06***	0.05***	0.05***	0.06***
	(0.01)	(0.01)	(0.02)	(0.01)
Legal Capacity (log)	0.01	-0.07	-0.05	0.01
	(0.10)	(0.16)	(0.16)	(0.10)
Complainant GDP (log)	0.04	0.10	0.10	0.04
	(0.06)	(0.09)	(0.09)	(0.06)
Defendant GDP (log)	0.04	0.09	0.07	0.04
	(0.04)	(0.06)	(0.06)	(0.04)
Constant	-4.10**	-6.50**	-5.86**	-4.09**
	(1.88)	(2.58)	(2.60)	(1.84)
N	369	158	158	369

Heckman selection model with maximum likelihood (ML) estimates. First stage estimates likelihood of a ruling. Second stage estimates likelihood of a pro-complainant ruling. Models (2) and (3) limited to GLOBAL disputes. Robust standard errors clustered on the common dispute. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$