Free-Riding on Enforcement in the WTO*

Leslie Johns† Krzysztof J. Pelc‡

July 8, 2015

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 more legal merit in order for countries to overcome the collective action problem, meaning that conditional on being filed, cases that challenge concentrated policies will have less *legal merit* 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.

*This research was supported by the Burkle Center for International Relations at UCLA. Many thanks to seminar participants at Cornell University, the Leitner Program in International and Comparative Political Economy at Yale University, and the Southern California International Law Symposium; as well as the annual conference of the International Studies Association. In particular, we thank Alex Debs, Julia Gray, Greg Shaffer, Richard Steinberg, and Rachel Wellhausen.

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

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

1 Introduction

Like many international organizations, the World Trade Organization (WTO) relies upon its members to challenge possible violations. Its dispute settlement system sometimes 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 began, 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 quickly 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? We argue that Canada's decision can be explained in part by how the US trade policies affected *other* countries.

Countries that file WTO disputes often pay a private cost for a public good. Accusing a trade partner of a violation inevitably antagonizes it, putting diplomacy at risk. For example, Japan's foreign affairs ministry has repeatedly prevented 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.[4] Filing a WTO dispute also requires considerable legal capacity, which many countries lack (Busch, Reinhardt and Shaffer, 2009). 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.[5] Even the US Trade

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

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

[3]DS357: *US—Agricultural Subsidies.*

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

[5]Davis (2012*a*). We thank Richard Steinberg for this point. By one estimate, the average complainant's relevant

Representative—one of the most frequent WTO litigants—faces budget constraints that limit its ability to file cases (Brutger, 2014). In contrast, the benefits from a legal challenge 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.[6] Complainants sometimes extract private benefits in settlements, yet the public benefits of litigation are so large that trade scholars routinely refer to WTO dispute settlement as a public good (Bown, 2005). 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 measurable drawbacks.

The *diffuseness* of a trade policy affects the degree to which litigation is a public good. Some trade policies, like the US's 1998 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.[7] 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.

We argue that more diffuse policies take longer to be challenged, and risk not being challenged at all. 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 likely to provide it. This suggests that when a trade policy is more diffuse, affected countries face a more significant collective action problem, reducing the likelihood that an individual country will bear the cost of WTO dispute settlement. We use a formal model to show that when dispute settlement is sufficiently costly, the individual effect matches the aggregate effect: when a trade policy is more

exports fall by about \$56 million in the run-up to a dispute (Bown and Reynolds, 2015), though this conceals much variation among disputes.

[6]WTO Dispute Settlement Understanding, Article 3.5.

[7]Other countries may benefit indirectly by a dispute's spill-over effects on jurisprudence (Pelc, 2014). We discuss this below.

diffuse, it is less likely to be challenged by *any* WTO member.

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.

First, we examine the timing of WTO disputes. Our infinite-horizon model shows that 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 for longer than a more concentrated policy. We measure the *enforcement delay* for each observed WTO dispute—the time between when a policy was implemented and when it was challenged at the WTO—and show that cases involving more diffuse policies experience a longer enforcement delay than cases involving concentrated policies.

Second, we examine the legal merit of WTO cases. While litigation always comes with uncertainty, some cases are better—are more likely to generate a pro-complainant ruling—than others. 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 sufficiently high legal merit. 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 exceptionally high legal merit for an affected country to file it. While we cannot directly test how legal merit affects selection, we can examine the legal merit of cases that are selected. We show that in observable WTO disputes, cases that challenge concentrated policies have less legal merit—are less likely to generate pro-complainant rulings—than cases that challenge diffuse policies.

We test our arguments on all WTO disputes from 1995 to 2013. To measure enforcement delay, we use data from Bown and Reynolds (2015) on the implementation dates for policies that are challenged at the WTO. To measure legal merit, we collected an original dataset of the ruling direction—whether the complainant prevailed on a legal argument—in all 4,484 individual legal claims across all WTO disputes.

One implication of our argument, which we do not test here, 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. In the conclusion, we discuss this and other implications of our argument, as well as a possible alternative explanation for our empirical findings.

2 Theory

2.1 Primitives and Structure

We present an infinite-horizon game with discrete time ($t = 1, 2, \ldots$). 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 $\tau_i > 0$).[8]

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 i privately learns its *type*, α_{it}—this represents the domestic pressure on i to challenge the policy in period t.[9] 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.[10]

Because we care about filing decisions, we model WTO dispute settlement in reduced

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

[9] Assumptions about the distribution of α_{it} are in the Appendix.

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

form.[11] 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 did not file—benefit from such outcomes. We let parameter $r > 0$ represent the *legal merit* of the case, and payoff $r\tau_i$ represent country i's expected per period payoff from WTO dispute settlement. We thus assume that legal merit increases each affected country's expected payoff from dispute settlement.

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 $k > 0$ 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.

2.2 Equilibrium Behavior and Comparative Statics

We first identify a property of the weak perfect Bayesian equilibrium for our game.[12]

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.

[11] The conclusion discusses a model extension with a more detailed dispute settlement system.

[12] This 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.

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. 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.

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

Legal merit, 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 legal merit increases. However, the expected utility functions change at different rates. When a country files, it knows 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 legal merit makes filing the case more attractive relative to not filing.

Proposition 4. *When the legal merit of the case 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 addresses diffusiveness because increasing the trade stake of one country also increases the total impact of the trade policy on all countries, $\tau = \Sigma_i \tau_i$. To examine diffuseness, we must control for the total impact of the trade policy. We therefore impose an additional assumption for the remaining results: we 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.

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. 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.[13]

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.*

2.3 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 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 chal-

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

lenges. 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.*

Second, we showed that diffuseness and legal merit 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 legal merit of the case. So there is a clear selection effect: conditional on being filed, a case that challenges a diffuse policy should have more legal merit, on average, than a case that challenges a concentrated policy.

Proposition 8. *In observable WTO disputes, cases that challenge diffuse policies will, on average, have more legal merit than cases that challenge concentrated policies.*

This argument is illustrated in Figure 1. The x-axis represents legal merit (r) and the y-axis represents a country's expected benefit from filing the case. Proposition 4 showed that as legal merit increases, a country's expected benefit from filing increases. When the expected benefit is positive, a country will file; when the expected benefit is negative, a country will not file. Proposition 5 showed that increasing the number of affected countries makes the trade policy more diffuse, reducing an individual country's incentive to file. The line graphs in the bottom half of Figure 1 show that increasing the number of countries (from 2 to 3 to 4) increases the minimum value of legal merit, $\overline{r}(n)$, needed for a country to file the case. So conditional on being filed, cases with a more diffuse impact should, on average, have more legal merit.

[Insert Figure 1 here.]

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 cases with more legal merit 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 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. 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 enforcement?

3.1 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 considerable 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 our more conservative findings, looking strictly at 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.

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 use 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.[14] If only two countries exported into the defendant country, and if they exported the same amount, 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.

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 US used to coerce other countries to amend their policies, there was no underlying traded product, even as the legislation had widespread consequences.[15] We code such cases as missing for the purpose of our first two diffuseness measures, but 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—Multilateral Policy—attempts to capture this vari-

[14] The HHi measure is:

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

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

[15] DS152: *US—Section 301 Trade Act.*

ation. This dichotomous variable comes from Bown and Reynolds (2015) and indicates whether a trade policy affects many members, or only a small subset. About 49% of the disputes in our sample are coded as concerning multilateral protectionist policies. The challenge of Section 301 is one such a multilateral policy. By comparison, most antidumping disputes are coded as non-multilateral, since they most often target a single country.[16] 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. Even without running our survival model estimates, the descriptive statistics for Multilateral Policy are telling: on average, multilateral policies in our sample take 825 days to be challenged, while non-multilateral cases take 720 days. If we consider the implementation date of policies that began in the GATT era, multilateral cases take over 80% longer to be filed.

Additionally, we include two control variables measuring the 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 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. Finally, we add a variable indicating the year in which a dispute was initiated, to control for possible 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

[16]The measure goes beyond simply looking at the legal clause at issue. For example, most safeguards are applied to imports from all countries, but some safeguards are targeted at only a few countries. Bown and Reynolds include such distinctions in their coding. Bown and Reynolds (2015) code all disputes as "global" or "partial" policies. Our use of the term "multilateral" corresponds to their use of the term "global".

our observations are censored, i.e. most disputes count only one complainant. Yet a handful of violations are challenged by more than one party, sometimes not simultaneously. We account for the common characteristics of such violations by clustering errors on the common violation.

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, Multilateral Policy. We thus begin our analysis with a univariate Cox model that only considers the relation between Multilateral Policy and the likelihood the measure is challenged in any given period, shown in Column 1 of Table 1. This estimation thus exploits our maximal sample.

In Column 2 of Table 1, we then add our two economic measures of diffuseness, Number of Countries Affected and Disputed Trade Flows HHi, as well the two trade flows variables, and the initiation year. In Column 3, 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 them 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 143 days), while others, like Australia, take much longer (2,871 days). The two superpowers fall somewhere around the membership median in this respect (1,654 days for the US and 1,155 days for the EU). 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. Nonetheless, they are enough to suggest the utility of the frailty parameter. Finally, in Columns 4-6 of Table 1, we rerun our estimation from Column 2, including our diffuseness indicators one at a time.

[Insert Table 1 here.]

The results in Table 1 support our theoretical expectations. 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. The type of legal discrimination at issue has a consistent effect: MULTILATERAL POLICY is associated with a significant decreased rate of legal challenge, though it comes short of high significance in Column 2. The effect is highly substantively significant throughout. As an example, in our first model, the rate of legal challenge decreases by 32% when the legal violation at issue is of a multilateral nature. To put this in terms of duration, we can re-estimate this same model using a parametric Weibull survival model (not shown), which suggests that multilateral policies take 64% more time to be challenged. Because of the stringent assumptions underlying parametric duration models, our non-parametric estimates in Table 1 remain more reliable.

Our two economic indicators of diffuseness also 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, 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. We graph the cumulative hazard associated with a concentrated effects policy versus a diffuse effects policy, holding all else equal, in Figure 2.[17] As the figure makes clear, more diffuse policies face a considerably lower rate of challenge than more concentrated policies.

[Insert Figure 2 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 a country has at stake, the greater the hazard rate and the shorter the enforcement delay. In contrast, Proposition

[17]In Figure 2, we assume that concentrated (diffuse) policies are non-multilateral (multilateral) policies with one standard deviation below (above) the average number of countries affected, and one standard deviation above (below) the average trade HHi.

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 consistently negatively related to the hazard rate, meaning that all else equal, violations where the rest of the world has more at stake result in a greater enforcement delay. This relationship is highly robust, and 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—have a lesser rate of legal challenge, and thus generate longer enforcement delays, on average.[18] In sum, enforcement choices appear individually, if not socially, rational.

3.2 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 legal merit of disputes. The simplest way of assessing legal merit is to look at the outcome of rulings. On average, we should expect disputes with greater legal merit to win more claims. 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 consistent that countries are able to observe legal merit, and choose to challenge those violations that have the highest likelihood of success before judges. 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.[19] 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,[20] 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 percentage of claims won by the complainant, which varies

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

[19] A "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.

[20] For 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.

from 0 to 1; its mean is 0.74 for panel rulings, and 0.69 for appellate rulings.

Indeed, 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 the WTO's jurisprudence than panel rulings (Bhala, 1998-1999). As a result, when assessing the legal merit of a dispute by looking at the direction of rulings, we are interested in rulings "net of appeal". The resulting dependent variable, Claims Won Net of Appeal, considers the panel ruling, as modified (or not) by the AB in the case of an appeal. In the absence of an appeal, this measure simply considers the percentage of claims won at the panel stage. Of course, not all claims are of equal importance, yet it would be a difficult task to account for this in a systematic fashion. One way of addressing this problem is to reduce the win to a binary outcome, where importance of individual claims is less likely to matter: keeping with existing work, we thus also code our measure of legal success in a binary fashion, 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.

Our explanatory variables of interest remain as described above in Section 3.1: two economic measures of diffuseness, Number of Countries Affected and Disputed Trade Flows HHi, and one legal measure, Multilateral Policy, indicating the nature of the policy. A simple descriptive statistic provides early support for our expectations: multilateral cases result in a pro-complainant ruling nearly twice as often as non-multilateral cases (0.58 vs. 0.32), and this difference is highly statistically significant in a t-test. Of course, this relationship does not account for selection in any way. We include 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 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. Finally, we include a Dispute Year variable 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 (Busch and Reinhardt, 2006; Kucik and Pelc, forthcoming; Porges, 2003). Yet the number of third parties in the room should not, by itself, have a direct effect on the direction of the ruling. Indeed, the bivariate correlation between our outcome variable, Claims Won Net of Appeal, and Number of Third Parties is -0.004, suggesting that there is little to no relationship between the two. The other variables in the first stage selection equation are Complainant GDP (logged) and Defendant GDP (logged), and Total Trade Stake. 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. As always, we cluster our robust standard errors on the common dispute, to allow for the fact that some disputes are filed over the same violation.

[Insert Table 2 here.]

The results are shown in Table 2. The dependent variable in Column 1 is the binary indicator of pro-complainant rulings, while in Column 2 we use the continuous measure, coded as the percent of claims won. Our beliefs about the relationship between the concentration of benefits and legal merit find broad support. 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 worst the prospects of the case. And disputes over multilateral cases fare better on average, though this effect loses statistical significance in Column 2. Yet even in that model, the three concentration variables, taken together, remain highly jointly significant.

These effects are also substantively important. Looking at our first model, disputes that fall under Multilateral Policy are 70% more likely to result in a pro-complainant ruling than non-multilateral policies. More starkly still, policies that rank as diffuse on all three indicators are four times more likely to result in a pro-complainant ruling than those that are considered concentrated.[21] In short, disputes where the benefits of enforcement are more concentrated appear

[21] See fn. 17, supra.

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, though this effect becomes insignificant with the continuous dependent variable. Although the defendant's market size is negatively associated with the complainant's success rate, while the complainant's market size is positively associated to it, both of these indicators fall short of significance, and they appear to have no effect on the selection stage. 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.

4 Conclusion

Protectionist trade policies vary in how diffuse their effects are. Legal challenges of highly diffuse policies are largely public goods because individual complainants pay a private cost to provide enforcement that 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 (Pelc, 2014), and the enforcement of trade rules today may deter future violations (Bown, 2004; Davis, 2012*b*). 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. Put simply, more diffuse protectionism generates a more severe collective action problem.

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 multilateral legal issues, that concern a greater potential number of WTO members, are challenged at a 32% 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 level of legal merit 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.

Might our empirical results be explained by a different theoretical mechanism?[22] One alternative explanation for our legal merit results is that WTO panels might issue biased rulings. Suppose that panels are more likely to rule in favor of complainants that are challenging trade policies that harm more WTO members.[23] Then more diffuse policies should be more likely to generate pro-complainant rulings. 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.[24] Yet even if we assume that panels make biased rulings for non-legal reasons, this alternative explanation contradicts our empirical findings on enforcement delay. If diffuseness increases the likelihood that WTO panelists support the complainant, then diffuseness also increases the expected utility of filing the case. This implies that diffuseness should increase the rate of challenges, and reduce enforcement delay. Thus, while a panel bias argument explains our second set of empirical results, it is inconsistent with our first set of empirical results.

One striking implication of our findings, which we leave for further research, is that if our theoretical argument is correct, countries should spread the pain of trade protectionism as much as possible when they provide import relief to domestic industries. Rather than using trade policies

[22]See the Supplemental Appendix for this extension.
[23]For example, see Carrubba, Gabel and Hankla (2008).
[24]In fact, a complainant can challenge a policy that has not harmed it.

with a concentrated impact, such as countervailing duties and antidumping duties, countries have incentive to use more diffuse policies when possible to avoid legal challenges. 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. However, 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. Additionally, we show that when a country challenges a concentrated policy, it often does so at the expense of legal merit. 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. The temptation to free-ride on enforcement means that litigation by a few cannot serve the interests of all.

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 $0 \leq \alpha_L < \frac{k}{\tau_i} < \alpha_H$ 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:

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

Equilibrium Behavior

Proof of Proposition 1. Country i will file iff:

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

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

$$\rho_i = \Pr\left(\alpha_{it} < \overline{\alpha}_i\right) = F\left(\overline{\alpha}_i\right) \quad \text{and} \quad \rho = \prod_k \rho_k = \prod_k F\left(\overline{\alpha}_k\right) \quad \text{and} \quad \rho_{-i} = \prod_{j\neq i}\rho_j = \frac{\prod_k F\left(\overline{\alpha}_k\right)}{F\left(\overline{\alpha}_i\right)}$$

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

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

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

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

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

$$\lim_{\delta\to 0}\Psi^i = \overline{\alpha}_i - \frac{k}{\tau_i} = 0 \quad \Leftrightarrow \quad \lim_{\delta\to 0}\overline{\alpha}_i = \frac{k}{\tau_i}$$

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

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

where $\overline{\alpha} = (\overline{\alpha}_1, \overline{\alpha}_2, \ldots \overline{\alpha}_n)$. □

***

Note that Ψ^i is continuously differentiable in its arguments, $\Psi^i_{\overline{\alpha}_i} = 1 - \delta\rho > 0$, and:

$$\Psi^i_{\overline{\alpha}_j} = \delta f(\overline{\alpha}_j) \frac{\rho}{\rho_i \rho_j} \left[\frac{k}{\tau_i} - \frac{\delta}{1-\delta} b + r + \int_{\alpha_L}^{\overline{\alpha}_i} \alpha f(\alpha) \, d\alpha - \overline{\alpha}_i \rho_i \right]$$

Therefore, the Jacobian matrix is nonsingular for small δ:

$$\mathbf{J} = \begin{bmatrix} \Psi^1_{\overline{\alpha}_1} & \ldots & \Psi^1_{\overline{\alpha}_n} \\ \ldots & \ldots & \ldots \\ \Psi^n_{\overline{\alpha}_1} & \ldots & \Psi^n_{\overline{\alpha}_n} \end{bmatrix} \quad \Rightarrow \quad \lim_{\delta \to 0} \det(\mathbf{J}) = \det(\mathbf{I}) = 1 > 0$$

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 $\frac{\partial \overline{\alpha}_1}{\partial \tau_1}$. By the implicit function theorem, $\frac{\partial \overline{\alpha}_1}{\partial \tau_1} = \frac{-\det(\mathbf{B})}{\det(\mathbf{J})}$ where:

$$\mathbf{B} = \begin{bmatrix} \Psi^1_{\tau_1} & \Psi^1_{\overline{\alpha}_2} & \ldots & \Psi^1_{\overline{\alpha}_n} \\ \Psi^2_{\tau_1} & \Psi^2_{\overline{\alpha}_2} & \ldots & \Psi^2_{\overline{\alpha}_n} \\ \ldots & \ldots & \ldots & \ldots \\ \Psi^n_{\tau_1} & \Psi^n_{\overline{\alpha}_2} & \ldots & \Psi^n_{\overline{\alpha}_n} \end{bmatrix}$$

The Supplemental Appendix shows: $\lim_{\delta \to 0} \det(\mathbf{B}) > 0$. This means $\frac{\partial \overline{\alpha}_1}{\partial \tau_1} < 0$ for small δ. □

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

$$\mathbf{C} = \begin{bmatrix} \Psi^1_{\tau_n} & \Psi^1_{\overline{\alpha}_2} & \ldots & \Psi^1_{\overline{\alpha}_n} \\ \Psi^2_{\tau_n} & \Psi^2_{\overline{\alpha}_2} & \ldots & \Psi^2_{\overline{\alpha}_n} \\ \ldots & \ldots & \ldots & \ldots \\ \Psi^n_{\tau_n} & \Psi^n_{\overline{\alpha}_2} & \ldots & \Psi^n_{\overline{\alpha}_n} \end{bmatrix}$$

The Supplemental Appendix shows: $\lim_{\delta \to 0} \det(\mathbf{C}) < 0$. This means $\frac{\partial \overline{\alpha}_1}{\partial \tau_n} > 0$ for small δ. □

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

$$\mathbf{G} = \begin{bmatrix} \Psi_r^1 & \Psi_{\overline{\alpha}_2}^1 & \dots & \Psi_{\overline{\alpha}_n}^1 \\ \Psi_r^2 & \Psi_{\overline{\alpha}_2}^2 & \dots & \Psi_{\overline{\alpha}_n}^2 \\ \dots & \dots & \dots & \dots \\ \Psi_r^n & \Psi_{\overline{\alpha}_2}^n & \dots & \Psi_{\overline{\alpha}_n}^n \end{bmatrix}$$

The Supplemental Appendix shows: $\lim_{\delta \to 0} \det(\mathbf{G}) > 0$. This means $\frac{\partial \overline{\alpha}_1}{\partial r} < 0$ 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 $\overline{\alpha}_i \in (\alpha_L, \alpha_H)$. When $\tau_i = \frac{\tau}{n}$, the system of $\Psi^i(\overline{\alpha})$-equations simplifies to one equation with one endogenous variable, $\overline{\alpha}_n$. 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, $\overline{\alpha}_n$, is defined by:

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

So: $\lim_{\delta \to 0} \overline{\alpha}_n = \frac{kn}{\tau}$. Similarly, the common cutpoint for the $(n+1)$-country game, $\overline{\alpha}_{n+1}$, is defined by:

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

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

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: $1 - (\rho_n)^n = 1 - F(\overline{\alpha}_n)^n$. For small $\delta > 0$, this probability is decreasing in n iff:

$$F(\overline{\alpha}_n)^n < F(\overline{\alpha}_{n+1})^{n+1} \quad \Leftrightarrow \quad \left(\frac{kn}{A\tau}\right)^n < \left[\frac{k(n+1)}{A\tau}\right]^{n+1} \quad \Leftrightarrow \quad \frac{A\tau n^n}{(n+1)^{n+1}} < k$$

$\square$

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: $E[X|n] = \sum_{t=1}^{\infty} t \left[(\rho_n)^n\right]^{t-1} \left[1 - (\rho_n)^n\right] = \frac{1}{1-(\rho_n)^n}$. Therefore:

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

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:

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

Then: $\Psi_r^n(\alpha_i) = \delta (\rho_n)^{n-1} > 0$, $\lim_{r\to\infty} \Psi^n(\alpha_i) = \infty > 0$, and $\lim_{\delta\to 0} \Psi^n(\alpha_i | r = 0) = \alpha_i - \frac{kn}{\tau} < 0 \Leftrightarrow \alpha_i < \frac{kn}{\tau}$. 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 $\overline{r}(\alpha_i, n) > 0$ such that $\Psi^n(\alpha_i | \overline{r}(\alpha_i, n)) = 0$. Type α_i files iff $\overline{r}(\alpha_i, n) \leq r$. Also:

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

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

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

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

References

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 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.

Davis, Christina. 2012*a*. *Why Adjudicate? Enforcing Trade Rules*. Princeton University Press.

Davis, Christina L. 2012*b*. "WTO Adjudication as a Tool for Conflict Management." Princeton University Working Paper.

Kim, Moonhawk. 2012. "Disguised Protectionism and Linkages to the GATT/WTO." *World Politics* 64(3):426–475.

Kono, Daniel Y. 2006. "Optimal Obfuscation: Democracy and Trade Policy Transparency." *American Political Science Review* 100(3):369–384.

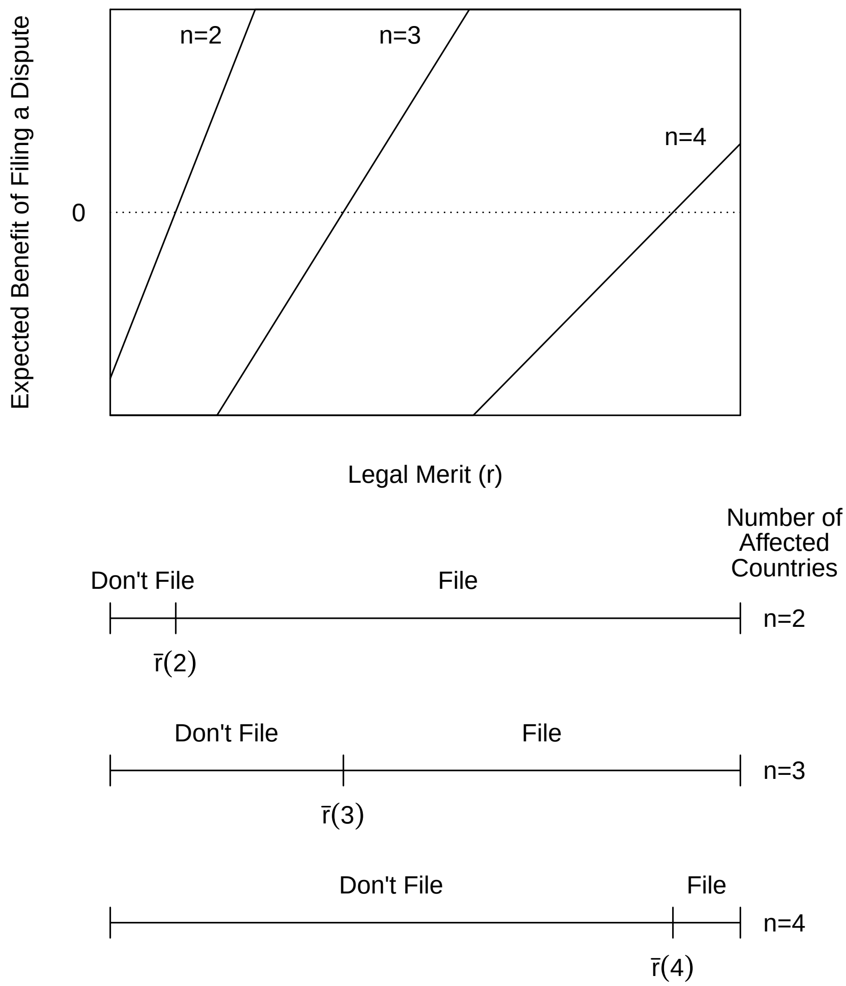
Kucik, Jeffrey and Krzysztof J. Pelc. forthcoming. "Measuring the Cost of Privacy: A Look at the Distributional Effects of Private Bargaining." *British Journal of Political Science* xx:xx–xx.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard 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.

Figure 1: Diffuseness Increases Legal Merit of Observed Cases



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

Figure 2: Cox Proportional Hazards Regression

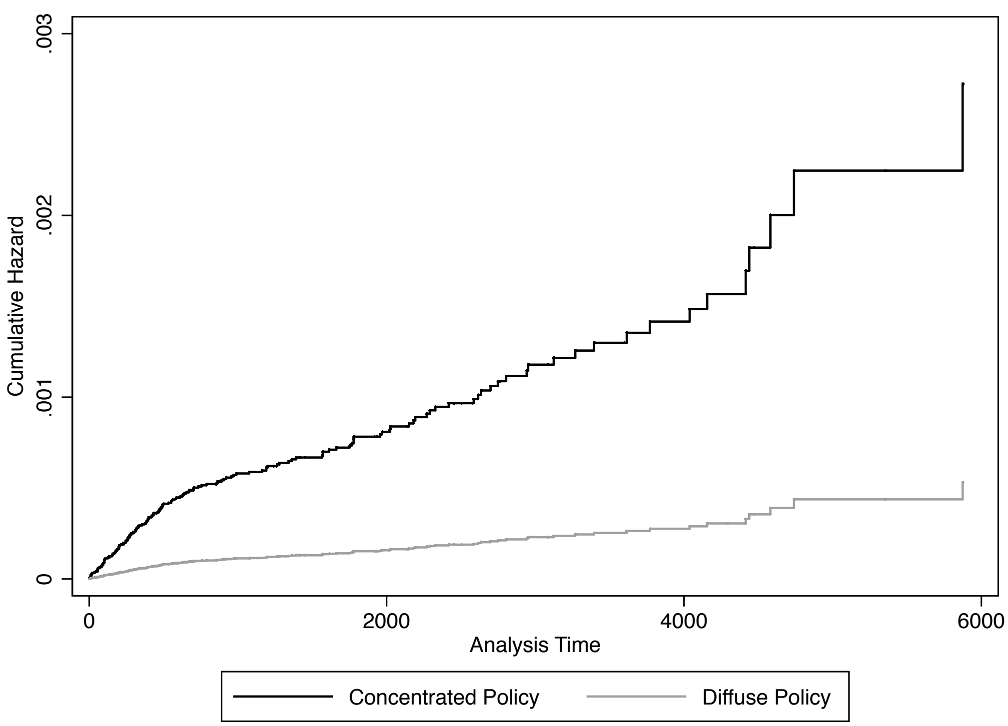


Table 1: Diffuseness of Protectionist Policies and Enforcement Delay

	(1)	(2)	(3)	(4)	(5)	(6)
Number of Countries Affected (logged)		-0.76***	-0.65***	-0.88***		
		(0.12)	(0.13)	(0.11)		
Disputed Trade Flows HHi		0.80*	1.25***		1.52***	
		(0.42)	(0.43)		(0.45)	
Multilateral Policy	-0.38**	-0.21	-0.36**			-0.36**
	(0.17)	(0.16)	(0.17)			(0.18)
Own Trade Stake (logged)		0.47***	0.48***	0.48***	0.44***	0.50***
		(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
ROW Trade Stake (logged)		-0.29***	-0.29***	-0.29***	-0.35***	-0.41***
		(0.04)	(0.04)	(0.04)	(0.03)	(0.04)
Initiation Year		-0.09***	-0.10***	-0.09***	-0.07***	-0.07***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N	26501	26165	26165	26165	26254	26254
Number of Disputes	360	267	267	267	267	267

Duration measured as the number of days elapsed between the implementation of a protectionist policy and its eventual challenge. Columns (1)–(2) and (4)–(6) show Cox regression estimates with robust standard errors in parenthesis clustered on common dispute. Column (3) shows Cox Frailty Model estimates adjusted for shared frailty of a given respondent. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2: Diffuseness of Policies and Pro-complainant Rulings

	(1)	(2)
CLAIMS WON NET OF APPEAL (2nd Stage eq.)	Binary	Percent
Number of Countries Affected (logged)	0.09**	0.06***
	(0.04)	(0.02)
Disputed Trade Flows HHi	-0.70**	-0.24*
	(0.31)	(0.13)
Multilateral Policy	0.38**	0.04
	(0.18)	(0.07)
Total Trade at Stake (logged)	0.04***	0.00
	(0.01)	(0.01)
Defendant GDP (logged)	-0.04	-0.01
	(0.06)	(0.02)
Complainant GDP (logged)	0.03	0.01
	(0.05)	(0.01)
Dispute Year	-0.01	-0.00
	(0.02)	(0.01)
DISPUTE GOES TO RULING (1st Stage eq.)		
Number of Third Parties	0.22***	0.24***
	(0.03)	(0.05)
Total Trade at Stake (logged)	0.04***	0.04**
	(0.00)	(0.02)
Defendant GDP (logged)	0.05	0.06
	(0.04)	(0.05)
Complainant GDP (logged)	0.05	0.05
	(0.04)	(0.05)
N	329	329

Heckman probit selection model (Column 1) and Heckman selection model (Column 2) with maximum likelihood (ML) estimates. First stage estimates likelihood of a ruling. Second stage estimates likelihood of a pro-complainant ruling (Column 1) and percentage of claims ruled in favor of complainant (Column 2). Robust standard errors clustered on the common dispute. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$