

DISCUSSION PAPER SERIES

No. 10571

**TRADE AGREEMENTS AND
ENFORCEMENT: EVIDENCE FROM WTO
DISPUTE SETTLEMENT**

Chad P. Bown and Kara M. Reynolds

***INTERNATIONAL TRADE AND
REGIONAL ECONOMICS***



Centre for Economic Policy Research

TRADE AGREEMENTS AND ENFORCEMENT: EVIDENCE FROM WTO DISPUTE SETTLEMENT

Chad P. Bown and Kara M. Reynolds

Discussion Paper No. 10571

May 2015

Submitted 20 April 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK

Tel: (44 20) 7183 8801

www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL TRADE AND REGIONAL ECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Chad P. Bown and Kara M. Reynolds

TRADE AGREEMENTS AND ENFORCEMENT: EVIDENCE FROM WTO DISPUTE SETTLEMENT[†]

Abstract

This paper examines implications of the terms-of-trade theory for the determinants of outcomes arising under the enforcement provisions of international agreements. Like original trade agreement negotiations, we model formal trade dispute negotiations as potentially addressing the terms-of-trade externality problem that governments implement import protection above the globally efficient level so as to shift some of the policy's costs onto trading partners. We first extend earlier theoretical models from trade agreement accession negotiations to the setting of enforcement negotiations, and the resulting theory guides our empirical assessment. We use instrumental variables to estimate the model on trade volume outcomes from WTO disputes over 1995-2009. Our evidence is consistent with theoretical predictions that larger import volume outcomes are associated with products that have smaller increases to foreign exporter-received prices (terms-of-trade losses to the importer) as a result of the dispute, larger pre-dispute import volumes, larger import demand elasticities, and smaller foreign export supply elasticities. Dispute settlement outcome differences are also explained by variation in institutionally-motivated measures of retaliation capacity and the severity of the free rider problem associated with foreign exporter concentration.

JEL Classification: F13 and F14

Keywords: dispute settlement, terms of trade, trade agreements and WTO

Chad P. Bown cbown@worldbank.org
The World Bank and CEPR

Kara M. Reynolds reynolds@american.edu
American University

[†] For useful comments and discussions, we thank Kyle Bagwell, Robert Staiger, Giovanni Maggi, Kamal Saggi, Russ Hillberry, Hiau Looi Kee, Wolfgang Keller, Kalina Manova, Aaditya Mattoo, Caglar Ozden, Michele Ruta, Roberta Piermartini, Manfred Elsig, Arevik Mkrtyan, and participants at seminars at Stanford, Vanderbilt, Illinois, Wesleyan, the World Bank, the Geneva Trade and Development Workshop, and the PEIO Conference in Berlin. We acknowledge the generous financial support from the World Bank's KCP Trust Fund. Thanks to Aksel Erbahar, Carys Golesworthy, and Semira Ahdiyyih for excellent research assistance and to Meredith Crowley for graciously sharing her estimated trade elasticities. All errors and omissions are our own.

1 Introduction

What determines the outcome of dispute settlement negotiations arising under the relatively new and highly legalistic procedures of international trade agreements? Dispute settlement under agreements like the WTO is most frequently triggered when one country is alleged to change a policy so as to provide import protection above the limit to which it had agreed in prior negotiations. However, the emergence of evidence that the terms-of-trade theory helps to explain policy changes that take place as the outcome of original trade agreement negotiations - whether through WTO accession or the result of multilateral GATT negotiating rounds - naturally raises the question of whether similar incentives shape the negotiated outcome arising under subsequent use of the agreement's enforcement provisions. Put differently, after a government policy deviation disturbs the originally negotiated trade agreement outcome by moving trade volumes away from globally efficient levels and thus triggering a dispute, does the terms-of-trade theory also help explain differences across the negotiated dispute settlement outcomes? Or, does membership in the trade agreement extinguish all terms-of-trade incentives from government policymaking behavior so that subsequent policy outcomes resulting from dispute settlement negotiations are determined by something else?

The purpose of this paper is to empirically estimate determinants of the outcomes of formal enforcement negotiations that take place under WTO dispute settlement. Like original trade agreement negotiations, we provide a formal model of trade dispute negotiations as potentially confronting the externality problem that arises when a government deviates from the originally negotiated outcome by implementing import protection above the globally efficient level so as to shift some of the policy's costs onto trading partners via a change in the terms of trade. Our specific approach to the examination of subsequent dispute resolution process is motivated by the combination of two insights from the existing literature on trade agreements and dispute settlement. Because these insights speak to the design of the institution established to resolve WTO disputes, we are skeptical that terms-of-trade motives have been completely extinguished and thus no longer play an important role in affecting trade policy outcomes arising under dispute settlement.

First, in an influential paper in the terms-of-trade literature, Bagwell and Staiger (1999) provide a theory-based interpretation of the GATT/WTO principle of *reciprocity* and its implications for how original trade agreement negotiations move countries from a prisoner's dilemma to a jointly efficient outcome. They find reciprocity can serve to coordinate two large countries' tariff changes in a way that neutralizes the otherwise negative (own) terms-of-trade impact that would take place if each country were to implement the same policy change unilaterally. Bagwell and Staiger (2011) derive formal implications for econometric estimation and provide product-level evidence from 16 countries consistent with the theory that negotiated tariff levels resulting from accession to the WTO agreement are related to pre-negotiation import volumes and trade elasticities.

Second, Bown (2002) and others have noted that WTO jurists have *interpreted* the agreement's dispute settlement rules for renegotiation (or retaliation) almost identically to how Bagwell and Staiger (1999) model the reciprocity principle that drives GATT/WTO liberalization negotiations. The WTO limits authorized retaliation in dispute settlement negotiations to a level that - when

viewed through the lens of the Bagwell and Staiger modeling framework - should neutralize the terms-of-trade gain of the respondent (importing) country's WTO violating unilateral policy change that is the subject of the dispute.

The empirical question of this paper is thus whether evidence of the terms-of-trade theory arising from the original trade agreement negotiations setting (Bagwell and Staiger, 2011) also extends to the trade agreement's *enforcement* negotiations setting. While we begin with the Bagwell-Staiger theoretical model, we are forced to adapt its empirical implementation in order to address shortcomings in observability of data that arises in the enforcement setting. For whereas Bagwell and Staiger's empirical examination of tariff negotiations under WTO accessions had access to data on precisely-measured "before" and "after" trade agreement tariff policy *levels*, the enforcement setting typically does not allow for direct measurement of the analogous (before and after) levels of the policy subject to the dispute.¹

A simple examination of the caseload clearly reveals that most policy changes that trigger WTO disputes concern allegations of excessive levels of import protection. Nevertheless, governments typically do not impose this protection as a straightforward tariff change, instead implementing the new protection through a nontariff policy. While the timing of such policy changes may be readily observable, the size of the policy movements are notoriously difficult to measure accurately. Our approach is therefore first to reinterpret the theory to account for what we can better observe and measure, which is data on the "before" and "after" *trade volumes* and *prices*.² Put differently, in order to examine whether the predictions from the terms-of-trade theory also extend to determinants of outcomes under trade agreement enforcement negotiations, we do not attempt to assess the impact of these determinants on the changes to the levels of the *policies* themselves, but instead to the changes in the trade volumes that result from these policy choices.

Our specific approach is to derive a formal estimation equation directly from the underlying theory and investigate empirically its relevance for the enforcement of international agreements and the negotiated outcomes that arise under dispute settlement. We use instrumental variables to estimate the model on data from formal WTO disputes that were initiated and legally concluded between 1995 and 2009 that involve alleged violations over policies that affect goods imports. We present evidence that larger trade volume outcomes are associated with products that had smaller increases to foreign exporter-received prices (terms-of-trade losses) at the conclusion of the dispute, larger pre-dispute import volumes, and a higher ratio of import demand to export supply elasticities. The results hold after extending the model to account for variation across new and institutionally-motivated measures of retaliation capacities across the litigating countries

¹In the Bagwell and Staiger (2011) trade agreement negotiations setting, the "before" policy was the tariff the country implemented prior to its WTO accession negotiations, and the "after" policy was the negotiated tariff following the country's WTO accession. In the trade dispute setting that we introduce below, the "before" policy will be the policy change that allegedly violates the WTO rules and which triggers the WTO dispute, whereas the "after" policy is the one imposed by the importing country at the conclusion of the formal process of WTO dispute settlement negotiations and rulings.

²Our approach is to impose sufficient structure on the estimation so as to control for other factors outside of the theoretical model that may also influence trade volumes and prices of disputed products for the period of the dispute.

that the previous literature has suggested are also likely to affect dispute settlement outcomes (Bown, 2004a).³ Finally, we also present evidence that the variation in trade policy liberalization arising under dispute settlement is also partially explained by the degree to which the exporting countries impacted by the policy change are able to overcome free-riding incentives, i.e., a factor that Ludema and Mayda (2009, 2013) have found to affect negotiated MFN tariff reductions among WTO members more generally.

After establishing these baseline results, we estimate the empirical model of determinants of negotiated outcomes on different subsamples of the underlying population of dispute settlement data so as to better understand where the terms-of-trade model is most likely to break down. First, we find the empirical results are particularly strong for disputes that are not settled, but that conclude with formal legal decisions. This evidence is reassuring as this particular subsample of disputes is the most aligned with our theoretical framework, given that the vast majority of disputes that result in a legal decision contain both a ruling that the respondent country has violated the WTO agreement and an accompanying request that the country should change its import-restricting policy so as to bring it into compliance.⁴ Second, there is also strong evidence that the model helps to explain the trade outcomes for disputes involving high-income countries.

Nevertheless, we show explicitly that the model does not perform well in all settings, and we draw inference from the existing literature on trade agreements to help provide potential interpretations for where the results break down. For example, the framework performs less well in explaining the dispute settlement outcomes for cases involving *developing* countries; one interpretation is that this is consistent with the theoretical literature suggesting that governments in developing countries may be more likely to rely on trade agreements for commitment reasons vis-a-vis relationships with their private sectors (Maggi and Rodriguez-Clare, 1998; Limão and Tovar, 2011) rather than for neutralizing the terms-of-trade impact of policy choices. We also explore whether the results vary across types of import-restricting policies subject to dispute, and we can rule out any concern that our results are limited to disputes over antidumping, and in particular, disputes challenging US antidumping use, one of the most frequently disputed topics.⁵

Overall, our results add to an emerging literature on international trade agreements, as a number of recent contributions demonstrate the empirical relevance of the terms-of-trade theory for the conduct and negotiation of trade policy across a variety of settings both inside and outside the GATT/WTO framework. Broda, Limão and Weinstein (2008) examine a set of *pre-trade*

³Bown (2004a) used an earlier data sample of WTO and GATT dispute settlement outcomes to investigate related questions. While that research presented evidence consistent with the terms-of-trade theory, its reduced form estimation framework was not linked to any formal theoretical model. See also Bown (2004b) and Grinols and Perrelli (2006). Horn, Mavroidis and Nordstrom (2005) provide one of the early and important empirical papers on the economic determinants of WTO dispute settlement activity.

⁴As we describe in more detail below, the Maggi and Staiger (2011) model of dispute settlement provides a theoretical prediction of such a biased pattern of policymaking behavior and rulings arising in equilibrium.

⁵This particular empirical question of interest is motivated by the results of Bown and Crowley (2013) that found, in their assessment of the Bagwell and Staiger (1990) theory of self-enforcing trade agreements, that terms-of-trade incentives could be interpreted as contributing to US use of antidumping over the period 1997-2006, a period which overlaps with our sample.

agreement levels of import protection and find evidence consistent with the theory that market power affects unilaterally-imposed (noncooperative) tariffs. Bagwell and Staiger (2011) examine a set of countries that newly acceded to the WTO between 1995 and 2005 and find that the negotiated, post-accession tariff levels that governments take on after joining the agreement are also consistent with the core theoretical predictions of what such negotiations can deliver. Ludema and Mayda (2013) find evidence that heterogeneity of exporter concentration can also be used to explain variation in the most-favored-nation tariff schedules for many of the long-term members of the GATT/WTO system as of the end of the Uruguay Round. Finally, Bown and Crowley (2013) focus on a government’s time-varying resort to import protection through potentially permissible trade agreement “exceptions” such as antidumping and safeguards. They provide evidence that even when applied import tariff policies are constrained by a trade agreement, governments can resort to other non-tariff policy instruments to respond to terms-of-trade motivated incentives to adjust levels of import protection in the face of trade volume shocks (Bagwell and Staiger, 1990).⁶ The totality of this recent evidence that terms-of-trade incentives affect trade policymaking, including the Ludema-Mayda and Bown-Crowley results that such incentives are not extinguished upon entry into trade agreements, also identifies potential challenges and limits to international cooperation over trade policy. Our evidence below suggests that this would include trade policy changes that arise through dispute settlement negotiations.

While the literature on trade agreement enforcement is much more nascent than that concerning the formation of trade agreements more generally, the recent increased interest in dispute settlement is driven at least partially by the tacit acknowledgment that the trading system’s legal procedures may be playing a particularly important role. First, legal scholars have now established a significant body of research assessing WTO jurisprudence and case law; in this work the system’s dispute settlement procedures are referred to as the “crown jewel” achievement of the international trading system, especially vis-a-vis the much less legalistic dispute system made available under the GATT period of 1947-1994.⁷ Second, in their survey of the bilateral trade flows of the products that member countries eventually subject to WTO dispute settlement, Bown and Reynolds (forthcoming) report that governments have requested the WTO legal system to scrutinize policies that affected nearly \$1 trillion in imports over the period 1995-2011. This translates to roughly \$55 billion annually or 0.5 percent of world imports in 2011.⁸ Furthermore, their detailed examination of the policies under dispute also reveals substantial variation in the scope of products and the size

⁶See also Nicita, Olarreaga, and Silva (2013). A more general survey of the economics literature on trade agreements is Maggi (2014).

⁷For an introduction to WTO law see, for example, Jackson (1997) or Palmetier and Mavroidis (2004). Furthermore, beginning with the WTO case law decisions arising in 2001, Horn and Mavroidis (2003) published an annual series from a project teaming legal scholars and economists that jointly assessed the newly arising WTO legal decisions. The series is now more than a decade old (subsequently having been extended by Bown and Mavroidis, 2013), and in total now covers nearly 100 formal legal decisions arising under WTO Panel Reports and from the Appellate Body.

⁸These figures are a lower bound as they are constructed from bilateral trade data for disputes affecting goods imports only and thus do not include disputes affecting exports or services. They also do not attempt to account for the indirect market access implications of WTO dispute settlement arising through jurisprudence or “off equilibrium” impacts of the dispute settlement system. For more on the latter, see Bagwell, Bown and Staiger (2015, Section 7), which surveys the economics literature on WTO dispute settlement.

of market access at stake across the caseload. For example, at the lower end of the distribution, similar shares of the WTO caseload (e.g., between 14 and 20 percent) involve a disputed policy affecting less than \$1 million in annual bilateral trade and only 1 product at a time as the shares of disputes at the upper end of the distribution that assess policies affecting in excess of \$1 billion in annual bilateral trade and more than 50 products at once.

In addition to WTO members trusting the legal system to rule over challenges to very heterogeneous policies that collectively affect hundreds of billions of dollars of international trade, there is also considerable variation in the trade volume outcomes that arise at the conclusion of these disputes. While we estimate below that bilateral import volumes increase by an average of 12 percent between the critical years of a formal WTO dispute, it is also the case that import volumes do not increase above even the pre-dispute starting levels in 38 percent of disputes, and import growth does not keep up with the general level of overall economic growth of the respondent country in 44 percent of disputes. Our econometric approach provides evidence for when the terms-of-trade theory can help to explain this considerable variation in dispute settlement outcomes.

Finally, our evidence also points to potential areas for emphasis for the evolving theoretical literature that explores the role of dispute settlement provisions in trade agreements modeled as incomplete contracts (Horn, Maggi, and Staiger, 2010). In particular, recent contributions by Maggi and Staiger (2011, 2015a, 2015b) and Staiger and Sykes (2013a) model dispute settlement provisions as helping to fill in some of the gaps of the trade agreement’s incomplete contract.⁹ While these papers typically abstract from terms-of-trade considerations, one reason why they are an important step forward for the literature is because they contain some of the first trade agreement models in which disputes can be triggered as an equilibrium outcome. Indeed, an implication of our results is that a potentially important area for theory to continue to explore is the role of enforcement provisions in trade agreements that are both incomplete *and* motivated by the desire to coordinate policy changes in order to address terms-of-trade externalities.

The rest of this paper proceeds as follows. Section 2 reviews the GATT/WTO institutional setting and the negotiating principle of reciprocity, as interpreted by Bagwell and Staiger (1999), and the resulting parallel definition arising in WTO dispute jurisprudence. Section 3 introduces a theoretical model derived from Bagwell and Staiger (2011) that guides the estimation framework. Section 4 describes the data and variable construction used in the empirical analysis, and Section 5 turns to the econometric estimates. Section 6 concludes.

2 GATT/WTO Negotiations and Dispute Settlement

One of Bagwell and Staiger’s (1999) critical theoretical contributions to the trade agreements literature was to establish the terms-of-trade externality as a central problem that such agreements are seen to solve. In order to establish this result, they provide a formal theoretical interpretation of the GATT/WTO principle of reciprocity.

⁹Other recent theoretical contributions include Beshkar (2010, 2014), Park (2011), Limão and Saggi (2008, 2013), and Ludema (2001).

The basic GATT/WTO principle of reciprocity arises in the text in two critical places. First, governments negotiate tariff reductions in GATT rounds under Article XXVIII *bis*, which contains clear language that participation is voluntary. While the Article XXVIII *bis* language indicates a desire for country negotiators to arrange “reciprocal and mutually advantageous” reductions in tariffs, there are no mandatory requirements for reciprocity to take place in the original GATT/WTO trade agreement negotiations. However, a second and formal application of reciprocity is found in the GATT rules for *renegotiation* of tariffs. Under Article XXVIII, a country is permitted to withdraw its previously granted tariff concessions and thus increase its tariffs. Nevertheless, if it and any adversely affected trading partner cannot come to an agreement as to a level of compensation that is due for such a tariff increase, reciprocity is understood as a limit to the tariff withdrawal (the retaliation response of the trading partner) to the amount that would balance “substantially equivalent concessions.”

A key theoretical contribution of the Bagwell and Staiger (1999) model arises when they provide a mathematical interpretation for this concept of reciprocity, which they then use to derive implications for trade agreements. Their interpretation allows them to show how reciprocity helps to coordinate policy-changing behavior between two large countries starting from a prisoner’s dilemma outcome in which both countries are imposing non-cooperative, “best response” tariffs. They interpret reciprocity as coordinating tariff reductions so that the (own) adverse terms-of-trade impact of each country’s import tariff reduction is neutralized by the positive impact it experiences through the trading partner’s simultaneous tariff reduction. The post-tariff reduction outcome in which each country imposes its “politically optimal” tariff maximizes joint (global) welfare because it achieves higher (and globally efficient) trade volumes relative to the volumes that arose under best response policies but without either country experiencing a change in its terms of trade.¹⁰

What are the implications of this approach for the WTO’s *enforcement* provisions? First, the original GATT/WTO enforcement texts also contain no explicit reference to reciprocity. Furthermore, in the case of a trade dispute in which the respondent country fails to comply with WTO rulings and the WTO must establish a limit to how much the complainant country is able to seek compensation through retaliation, the WTO’s Dispute Settlement Understanding (DSU) states “[t]he level of the suspension of concessions or other obligations authorized by the [Dispute Settlement Body] shall be equivalent to the level of the nullification or impairment” (GATT, 1994, Article 22:4). Thus, the DSU texts were initially unclear as to what would determine the limit to retaliation, such as whether it would also be limited by the principle of reciprocity.

Nevertheless, Bown (2002) notes that in practice the first two WTO disputes to reach the retaliation-defining stage of the WTO’s dispute settlement process (*EC - Banana Regime* and *EC - Beef Hormones*) established jurisprudence which arguably adopted the Bagwell and Staiger (1999) formulation of reciprocity to define the limit to the tariff increase that a complainant country would be authorized to implement if the respondent did not remove the WTO-inconsistent policy.

¹⁰In the next section below we formally define the “best response” and “politically optimal” levels of import tariffs in the context of an economic model.

And while *stare decisis* and binding precedent are not as robust a feature of WTO law as other legal settings, Bown and Ruta (2010) show that the interpretations of the arbitrators in the *EC - Banana Regime* and *EC - Beef Hormones* disputes that limited retaliation to the level defined by the Bagwell and Staiger interpretation of reciprocity were not one-time events. They provide a detailed examination of the decisions in the 10 formal WTO disputes taking place between 1995 and 2008 that led to the phase in which the WTO arbitrators authorized and articulated retaliation levels and suggest that WTO arbitrators have consistently sought to define limits to authorized retaliation in a manner similar to this interpretation of reciprocity.¹¹ One implication that motivates our approach is that, during this period, respondent importing countries are likely to have had a good understanding of the upper limit of retaliation to which they may have found themselves subject if they refused to comply with WTO rulings.¹²

To summarize, the intuition for reciprocity in the enforcement setting is that, in a dispute, the complainant country would be authorized a tariff retaliation that would allow it to neutralize the terms-of-trade impact of the respondent country's original WTO violation. The simultaneous act under WTO dispute settlement of one country (the respondent) removing its WTO-violating policy in order to comply with a legal ruling and a second country (the complainant) ending its WTO-authorized retaliation can be seen as neutralizing the terms-of-trade impact of policy changes, in just the same manner as two countries liberalizing tariffs simultaneously under original WTO agreement tariff liberalization negotiations.¹³

The subsequent analysis is therefore motivated by insights from the underlying Bagwell and Staiger (1999, 2011) theory on reciprocity and its empirical implications for original trade agreement negotiations combined with recognition that WTO jurisprudence interpreted retaliation limits similarly in formal dispute settlement (Bown, 2002; Bown and Ruta, 2010). Our approach examines whether the empirical evidence of the Bagwell and Staiger (2011) trade agreement setting extends to the empirical setting of trade agreement enforcement. In the next section we more formally develop a theoretical model to guide the empirical examination in the remainder of the paper.

3 Theoretical Model

Bagwell and Staiger (2011) develop a multi-country, partial equilibrium model in which the domestic government can impose an ad valorem tariff τ on imports; domestic prices are thus defined as $p = (1 + \tau)p^w$ where p^w is the world price. The objective function of each government is defined as

¹¹Furthermore, in a number of instances in which arbitrators deviated from the definition, Bown and Ruta (2010) suggest that it was not necessarily due to a conceptual dissatisfaction with the Bagwell and Staiger definition but instead can be motivated by limits to data availability (e.g., services trade), measurement issues, or potentially different rules for limiting retaliation under different areas of WTO law, such as subsidies. Most of the disputes in the data set that we estimate below would not fall into these categories, had they reached the stage under which DSU arbitrators determined retaliation limits.

¹²See also the discussion in Schwartz and Sykes (2002) that interprets such retaliation limits as implying a “liability rule” remedy, and thus the implications for efficient breach of the trade agreement contract.

¹³Put differently, if the respondent refuses to comply with the WTO ruling, reciprocity defines the limit to the complainant's retaliation as the amount that offsets the respondent's original terms-of-trade gain associated with violating the agreement.

the weighted sum of producer surplus (PS), consumer surplus (CS), and tariff revenue, according to the equation:

$$W = \gamma PS(p(\tau, p^w)) + CS(p(\tau, p^w)) + (p(\tau, p^w) - p^w)M(p(\tau, p^w)). \quad (1)$$

In this equation $\gamma \geq 1$ reflects potential political economy pressure on the domestic government through a potential extra weight that the government places on producer surplus in its objective function, and $M(p)$ is the level of imports. Like Bagwell and Staiger (2011), we assume that W is globally concave over non-prohibitive τ . For this condition to be met even when the country is “small” (or $\partial p^w / \partial \tau = 0$), it must be the case that

$$W_{pp} < 0. \quad (A1)$$

This assumption is satisfied as long as demand is not too convex and supply is not too concave.

When the domestic government is unconstrained by trade agreements, we assume it chooses to impose its best response tariff (τ^{BR}) to maximize total domestic welfare:

$$W_p \frac{dp}{d\tau} + W_{p^w} \frac{\partial p^w}{\partial \tau} = 0 \quad (2)$$

Note that the partial derivative W_{p^w} , holding domestic prices constant, is equal to $-M(p)$. Thus this first order condition can be rewritten in the form:

$$-\frac{W_p}{p^{w, BR}} = \frac{\sigma^{BR}}{\omega^{*BR}} \frac{M^{BR}}{p^{BR}} \quad (3)$$

where σ^{BR} and ω^{*BR} are the (absolute value of the) elasticity of domestic import demand and foreign export supply faced by the domestic country, respectively, and the superscript BR denotes the levels of import volumes, world and domestic prices, and their trade elasticities, when evaluated at the best response tariff.

Following Bagwell and Staiger (2011), a second level of import protection worth highlighting is the country’s politically optimal tariff, given by τ^{PO} . This is the level of protection the government would impose if it were not motivated by terms-of-trade considerations, and is thus defined as the tariff that satisfies

$$W_p(p^{PO}, p^{w, PO}) = 0 \quad (4)$$

where the superscript PO indicates the politically optimal level of domestic and world prices.

In a series of research, Bagwell and Staiger (1999, 2002) have shown how the GATT/WTO principle of reciprocity can improve global economic efficiency and deliver relief from the terms-of-trade driven prisoner’s dilemma. The principle can be interpreted as one that allows countries to coordinate policies and thereby move from a noncooperative equilibrium in which governments impose best response tariffs (τ^{BR}) to a cooperative equilibrium in which governments impose their politically optimal tariffs (τ^{PO}). The reciprocity principle delivers this outcome because the coor-

minated movement serves to neutralize what would otherwise be an adverse terms-of-trade impact of a unilateral import tariff reduction. In their empirical application, Bagwell and Staiger (2011) further develop this theoretical model in order to estimate determinants of how countries change their tariffs from their pre-WTO levels (interpreted as τ^{BR}) to their post-WTO accession levels (interpreted as τ^{PO}).

The approach we develop below will ultimately examine the relevance of this theoretical model for trade agreement enforcement negotiations that take place under dispute settlement provisions. This is motivated by our discussion in Section 2 that found stark parallels between the Bagwell and Staiger (1999) theoretical interpretation of reciprocity and its implications for trade agreement negotiations and how WTO jurists have interpreted the limits to permissible retaliation that can take place under WTO enforcement in dispute settlement negotiations. We start the theory from the place that countries have signed onto a trade agreement, but nevertheless the domestic country has violated the agreement and once again implemented its best response tariff policy.¹⁴ This country will then face a dispute and we seek to examine determinants of its policy decision to return to the politically optimal tariff.¹⁵ In most all respects, we follow the Bagwell and Staiger (2011) modeling logic and intuition as they would transfer from the WTO agreements negotiations to the WTO enforcement negotiations. The one critical way in which our approach must differ from Bagwell and Staiger (2011) is that our empirical setting is complicated by the fact that, in most instances, governments do not deviate from the WTO agreement by simply implementing an observable best response tariff but instead some nontariff barrier.¹⁶ Thus we use the remainder of this section to reformulate the Bagwell and Staiger (2011) approach to fit our setting and, in particular, observable data.

Consider the simple linear version of the Bagwell and Staiger (2011) model. Domestic demand and supply are defined, respectively, by the following equations:

$$D(p) = \alpha - \delta p \tag{5}$$

$$S(p) = \lambda + \kappa p \tag{6}$$

¹⁴We describe the theory in terms of a direct violation of the trade agreement. Nevertheless, especially since we are interested in measuring determinants of *trade volume* outcomes in lieu of policies, our approach should also apply to instances in which governments deviate from their trade agreement obligations through non-tariff policies, including domestic policies. It is possible under the WTO to pursue trade disputes in which no explicit WTO obligations were violated but in which market access expectations have nevertheless been frustrated; such disputes are triggered by “nonviolation nullification and impairment” claims under GATT Article XXIII:1. For a discussion and one theoretical approach to nonviolation disputes under the GATT/WTO, see Staiger and Sykes (2013a, b).

¹⁵That is, we will not seek to model why it is that the country has already deviated from the politically optimal policy back to the best response policy. We assume that the deviation has taken place and seek to examine determinants of the negotiations back to the politically optimal policy. We do, however, explore some of the potential implications of this assumption in our discussion of the empirical results below.

¹⁶Put differently, Bagwell and Staiger (2011) are able to empirically examine the relevance of the model for tariff negotiations because there is available data on pre-WTO accession and post-WTO accession tariffs.

where both $\delta, \kappa > 0$. Further note that the W_p is defined by the expression:

$$W_p = (\gamma - 1)S(p) + (p - p^w) \frac{\partial M(p)}{\partial p}. \quad (7)$$

Finally, market clearing requires $M(p) = D(p) - S(p)$, which then yields a general formulation for import tariffs in the linear model as

$$\tau = \frac{[\alpha - \lambda] - M(\cdot)}{p^w(\delta + \kappa)} - 1. \quad (8)$$

In the linear model, Bagwell and Staiger (2011) show that the terms-of-trade theory makes the following prediction for an estimating equation for politically optimal tariffs as a function of pre-negotiation (best response) tariffs, import volumes, and world prices

$$\tau^{PO} = \beta_0 + \beta_1 \tau^{BR} + \beta_2 M^{BR}/p^{w, BR}, \quad (9)$$

where $\beta_0 = [(\gamma - 1)\kappa(r - 1)]/\{r[\delta + \kappa - (\gamma - 1)\kappa]\}$, $\beta_1 = (1/r)$, $\beta_2 = -\theta/\{r[\delta + \kappa - (\gamma - 1)\kappa]\}$, $r \equiv p^{w, PO}/p^{w, BR}$, and $\theta \equiv (-\partial M/\partial p)/(\partial E^*/\partial p^w)$.¹⁷ Furthermore, under the model's assumptions, it is straightforward to show that $\beta_0 \leq 0$ as $r \leq 1$ and $\beta_1 \leq 1$ as $r \geq 1$. Finally, $\beta_2 < 0$ since $\theta > 0$ and using $[\delta + \kappa - (\gamma - 1)\kappa] > 0$ by equation (A1). I.e., controlling for the level of the best response tariff, the negotiated (politically optimal) tariff will be lower the larger is the ratio of pre-negotiation import volumes to world prices, or $M^{BR}/p^{w, BR}$.

If data constraints were not an issue, the same approach could be adopted to model trade agreement *enforcement* negotiations that seek to have governments move from their best response policy (under dispute) back to their politically optimal policy. Unfortunately an equivalent test of equation (9) is not empirically possible in the enforcement setting because the *level* of the best response policy deviation that becomes subject to dispute, τ^{BR} , is typically not observable in the data.

Our approach is to instead use information from equation (8) on how the politically optimal and best response tariffs relate to *observable* import volumes. We then substitute this into equation (9) in order to obtain an estimating equation to take to the enforcement data. Solving for an estimation equation of determinants of the politically optimal level of import volumes yields

$$M^{PO} = \psi_0 M^{BR} + \psi_1 (p^{w, PO} - p^{w, BR}) \quad (10)$$

where $\psi_0 \equiv 1 + [\theta(\delta + \kappa)]/[\delta + \kappa - (\gamma - 1)\kappa] > 1$, again because $\theta > 0$ and using $[\delta + \kappa - (\gamma - 1)\kappa] > 0$ by equation (A1), and $\psi_1 \equiv -(\delta + \kappa)^2/[\delta + \kappa - (\gamma - 1)\kappa] < 0$.

There are two key predictions that can be derived from the linear model and equation (10). First, the post-dispute volume of imports (M^{PO}) should be increasing in the (pre-dispute) best response volume of imports (M^{BR}). Second, the post-dispute volume of imports should be decreasing in the

¹⁷In Bagwell and Staiger (2011), equation (9) is given by equation (12) on p. 1248.

world price increase received by the foreign exporter in the *post-dispute* political optimum relative to the *pre-dispute* best response ($p^{w,PO} - p^{w,BR}$).

In more general and nonlinear models in which θ is not constant, we also follow the logic of Bagwell and Staiger (2011) so as to use available information from import demand and foreign export supply elasticities. That is, we can manipulate equation (10) in order to also consider an equation of the following form

$$M^{PO} = \xi_0 M^{BR} + \xi_1 (p^{w,PO} - p^{w,BR}) + \xi_2 \left[\frac{\sigma^{BR}}{\omega^{*BR}} M^{BR} \right] \quad (11)$$

where $\xi_0 \equiv 1$, $\xi_1 \equiv -(\delta + \kappa)^2 / [\delta + \kappa - (\gamma - 1)\kappa] < 0$, and $\xi_2 \equiv [(\delta + \kappa)] / [\delta + \kappa - (\gamma - 1)\kappa] > 0$ using equation (A1). Then, since the parameter on the best response volume of imports alone is equal to 1, we can re-arrange equation (11) to examine the difference between the volume of imports at the conclusion of the dispute (M^{PO}) relative to when the policy was imposed and the dispute was ultimately triggered (M^{BR}), i.e., through examination of

$$M^{PO} - M^{BR} = \xi_1 (p^{w,PO} - p^{w,BR}) + \xi_2 \left[\frac{\sigma^{BR}}{\omega^{*BR}} M^{BR} \right]. \quad (12)$$

Our interpretation of this more general model and equation (12) is then the following. With respect to ξ_1 , the change in the volume of imports associated with the resolution of the dispute should again be decreasing in the world price increase received by the foreign exporter in the *post-dispute* political optimum relative to the *pre-dispute* best response. The import volume increase is larger, the smaller is the terms-of-trade loss to the respondent (importing) country when changing its policy at the end of the dispute.

Furthermore, the change in the volume of imports should be increasing in the *interaction* between the ratio of import demand to export supply elasticities (σ^{BR}/ω^{*BR}) and the (pre-dispute) best response volume of imports. The intuition behind this prediction has two components. First, for a given best response volume of imports M^{BR} , the post-dispute level of imports is increasing in the elasticity of import demand, so that a given tariff decrease generates a larger increase in the quantity of imports demanded. Products with high import demand elasticities would be associated with large economic distortions associated with imposition of the best response tariff. Second, for a given best response volume of imports M^{BR} , the post-dispute level of imports is *decreasing* in the foreign export supply elasticity. I.e., the more market power that the importing country has vis-a-vis the exporter, or the larger is the inverse of the foreign export supply elasticity, the larger will be the politically optimal import volume. Put differently, note that the small importing country case corresponds to $\omega^{*BR} \rightarrow \infty$, in which the final term in equation (12) goes to zero, regardless of M^{BR} or σ^{BR} . Furthermore, in the small country limiting case, $p^{w,BR} \rightarrow p^{w,PO}$ and so by equation (12), $M^{BR} \rightarrow M^{PO}$. On the other hand, the smaller is ω^{*BR} , the greater is the responsiveness of foreign export supply to any given change from the best response to the politically optimal policy, and thus the larger the resulting post-dispute level of imports, *ceteris paribus*.

Before moving to a discussion of the data and our estimation approach, we clarify two impor-

tant aspects of our theoretical framework that are explicitly designed to focus attention on the potential terms-of-trade implications of government policy changes. First is our assumption that the respondent government sets off the event that leads to the dispute through an exogenously triggered implementation of its best response policy. That is, we do not consider formal motives behind why the government made such a policy change and whether it is due to shocks to political preferences (e.g., γ), trade volumes, technology, domestic demand conditions, etc. Second, we also assume that this dispute-triggering event involves the respondent deviating all the way to its best response policy; i.e., we do not consider the theoretical possibility that the country deviates from the initial politically optimal policy to some alternative policy that falls short of the best response.¹⁸ While our approach is designed to empirically investigate whether the determinants of trade dispute outcomes are consistent with the core relationships deriving from the terms-of-trade theory and in general we find evidence of these relationships, we also report results from subsamples of the data illustrating explicitly where this framework breaks down. We provide below potential explanations behind the limitations to the model’s performance in explaining trade liberalization outcomes (and, by extension, inference on policy changes) and whether they may be motivated by the viability of these particular assumptions.¹⁹

4 Data and Estimation

We rely on theoretically motivated equation (12) to ultimately estimate models of the form

$$\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR}) = \xi_1[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})] + \xi_2\left[\frac{\sigma_{gr}^{BR}}{\omega_{gc}^{*BR}} \ln(M_{grc}^{BR})\right] + \nu_{grc}, \quad (13)$$

where g indexes the disputes (products), r indexes respondent (importing) countries, c indexes complainant (exporting) countries, and ν_{grc} is the error term. The theory suggests our estimates to be $\xi_1 < 0$, and $\xi_2 > 0$.

Next we turn to a discussion of the econometric issues associated with estimating equation (13) and our particular instrumental variables approach that addresses these concerns. Then we introduce the data and construction of variables used in the estimation.

¹⁸While this is clearly a simplifying assumption, in the absence of evidence suggesting otherwise, we motivate it as a reasonable first step. Put differently, one important question for a formal model to address would be why, if the respondent country knew it would ultimately face a formal trade dispute for any policy change above the status quo (politically optimal) level, it would chose to implement a policy increase short of the best response level.

¹⁹To our knowledge, Maggi and Staiger (2015b) is one of the few trade agreement models that has both disputes arising in equilibrium and variation in “outcomes” of the dispute, i.e., whether disputes settle early or proceed fully through the legal process and obtain rulings. Nevertheless, the basic underlying trade agreement framework of that model is not designed to capture salient aspects of the terms-of-trade theory of trade agreements that the empirical evidence (cited above) has found to be relevant across a number of different settings for the formation of non-cooperative and cooperative tariffs (and non-tariff barriers) under the WTO. Other theoretical contributions in which disputes can arise in equilibrium, but in which terms-of-trade motives are also not considered, include Beshkar (2010, 2014), Maggi and Staiger (2011), Staiger and Sykes (2013a) and Park (2011).

4.1 Instrumental variables estimation

Two econometric issues arise when we seek to estimate equation (13) that lead us to use instrumental variables (IV) to address potential concerns over endogeneity.

First, a strict interpretation of the model may omit factors that impact both the quantity of imports directly (through ν_{grc} in equation 13) and indirectly through its impact on changes to exporter prices, $[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})]$. Examples would include non-captured import demand or foreign export supply shocks; if left unaddressed, this would lead to inconsistent estimates of ξ_1 .

We construct instruments for the change in exporter-received prices $[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})]$ resulting from the dispute that is uncorrelated with ν_{grc} . Our instrument for $[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})]$ is based on earlier changes in exporter prices for the disputed product - e.g., those that took place because of the original imposition of the allegedly WTO-violating policy. Specifically, if BR is the year after the respondent first imposes its best response policy (i.e., the first full year that the policy is in effect), our approach is to instrument for $[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})]$ with $[\ln(p_{grc}^{w,BR-2}) - \ln(p_{grc}^{w,BR})]$.²⁰ We also follow Khandelwal (2010) and include as an additional instrument the percent change in crude oil prices between the best response and politically optimal years interacted with the log of distance between the respondent and complainant. In addition to being a potential supply shifter, this change in crude oil prices will control for changes in transportation costs over the period of dispute which are necessarily embedded in our calculations of export prices, as described below.

A second potential concern with estimating equation (13) is the endogeneity of the volume of imports in the best response year, M_{grc}^{BR} . Both the volume of imports in the best response year and the growth of imports between the best response and politically optimal years will be a function of shocks occurring during the best response year. To motivate our empirical strategy to address this concern, consider a decomposition of the error term into two components:

$$\nu_{grc}^{BR} = \epsilon_{grc} + u_{grc}^{BR}.$$

Here ϵ_{grc} captures the time-invariant, unobserved factors that influence the volume of imports between two countries, e.g., distance, common border, historical (colonial) relationships, etc.²¹ The second component, u_{grc}^{BR} , captures any time-varying unobserved components to the bilateral volume of imports in the best response year. In estimating equation (13), the time-differenced dependent variable $\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$ differences out the time-invariant component to the error term, ϵ_{grc} . Nevertheless, failing to address the component of the error term associated with the best response year and thus the endogeneity of M_{grc}^{BR} would result in biased coefficient estimates.

We therefore instrument for M_{grc}^{BR} with the average bilateral import volume between 1988 and

²⁰We fully explain the mapping of the timing of policy changes, years $t = BR$ and $t = PO$, and the trade dispute information in the next section.

²¹Given the limited period of time between the best response and politically optimal year, one could also imagine that near time-invariant variables would also be captured in this component, such as factors measuring the comparative advantage of two countries, relative size, and whether the two countries are members of the same regional trade agreement.

the year prior to the best response year.²² This average import volume will be highly correlated with M_{grc}^{BR} as it is a function of the time-invariant determinants of trade captured in ϵ_{grc} . Unlike M_{grc}^{BR} , however, the instrument will be uncorrelated with any shocks limited to the best response year, u_{grc}^{BR} , that remains in our error structure.

Finally, we report below standard tests for under-identification, weak instruments, and over-identification so as to provide an assessment of the quality of our instruments.

4.2 Construction of WTO dispute sample

We begin with a population of 347 formal bilateral (complainant-respondent) WTO disputes that were initiated and legally concluded between 1995 and 2009.²³ We begin with 1995 as that was the first year that the WTO and its Dispute Settlement Understanding was in effect, and we conclude in 2009 as we require two years of post-dispute trade data with which to observe potential changes in trade volumes resulting from the dispute settlement negotiations. We expand the WTO dispute database of Horn and Mavroidis (2008) by adding disputes from more recent years as well as details on the policies and traded products under dispute, as now made available in the data set accompanying Bown and Reynolds (forthcoming).

Figure 1 presents a timeline of the policy changes associated with the dispute settlement process and the means by which we map the timing of the theory to the dispute settlement data.

The first year of interest is the “best response” year, or the first full year that the respondent country has in place the policy that is ultimately the subject of the WTO dispute. In Figure 1, $t = BR$ is given by period 3, and it occurs one year after the alleged WTO-violating policy was first imposed and is thus the first full year that the policy is in effect. While respondent countries are rarely alleged to have simply raised their applied tariffs, but instead are alleged to have imposed a WTO-violating policy through a nontariff barrier (which implies difficulty in measuring the *size* of the ad valorem equivalent of the policy change), identifying the timing of the *imposition* of the best response policy (period 2) is relatively straightforward. In most instances, the year of imposition of the best response policy is available from either official WTO documentation associated with the dispute or from official government notifications available from other sources. However, it is important to note that 18 percent of the disputes in the sample are initiated *without* such a policy change being relevant. To clarify, it is not that the date of the policy change is unobservable, it is simply that the WTO dispute arises under the allegation that the respondent has *failed* to change its policy so as to bring itself into compliance with basic WTO obligations undertaken either at the end of the Uruguay Round or at the point of accession. In these cases, we define $t = BR$ simply to

²²We chose 1988 because it serves as the beginning of the 6-digit Harmonized System import data series.

²³Overall, members filed 402 WTO disputes between 1995 and 2009. However, 55 disputes were associated with alleged policy violations that were still in force as of 2009. Because our model examines a setting in which enforcement negotiations take place bilaterally, we define our unit of observation as a complainant-respondent pair. Therefore, we further clean the population of redundant disputes (i.e., a complainant filing multiple disputes against the same respondent country over the same issue) and break into bilateral pairings any instances in which multiple complainants jointly file a dispute against a common respondent over the same issue.

be the year of the initiation of the formal WTO dispute.²⁴

The best response year is not only necessary for establishing the baseline level of import volumes in equation (13) from which to assess import growth associated with the dispute, but as described above, it is also used to construct our instruments. One of our instruments will be the change in foreign-exporter received prices associated with the timing of the imposition of allegedly WTO-violating policy between period 1 and period 3, or $[\ln(p_{grc}^{w,BR-2}) - \ln(p_{grc}^{w,BR})]$. In light of the discussion above, one clear motivation for assessing the robustness of our results to inclusion of the 18 percent of the dispute settlement sample that involves an allegation of no policy change by the respondent is that there may be measurement issues affecting the performance of $[\ln(p_{grc}^{w,BR-2}) - \ln(p_{grc}^{w,BR})]$ as an instrument. I.e., for disputes in which no respondent policy change took place (and that the failure of a change in policy is what triggered the dispute), foreign-exporter received prices may not have changed prior to the dispute, and thus 18 percent of the caseload may not have sufficiently useful variation for a strong enough instrument. Below we therefore discuss the robustness of our baseline results to dropping these particular observations.

The second year of interest in Figure 1 is the “politically optimal” year, or the year by which the respondent is supposed to have implemented its policy reform following resolution of the dispute. Our rule for establishing the legal “conclusion” of a dispute (period 5 in Figure 1) is the following. Whenever the timing of the announced change in the policy is directly observed because of dispute rulings, other legal filings (e.g., mutually agreeable solutions), revelations through other official government documents (e.g., removal of temporary trade barriers), or other reports, we take $t = PO$ (period 7) to be two years after this notification. This two year window accounts for the WTO’s institutional constraint that provides a “reasonable period of time” to implement policy changes after such rulings and thus for trade flows to be given time to respond.²⁵ In 83 percent of our sample, the timing of post-dispute policy change implemented by the respondent is directly observable. However, in 17 percent of disputes there is no formal announcement of the policy change or legal conclusion. In these instances, we define the dispute’s legal conclusion (period 5) to be the last legal correspondence between the two main litigants in the WTO dispute, and we take the politically optimal year $t = PO$ to be three years after this period. Since it is unclear what, if any, policy changes took place in these disputes, we also report and discuss below the sensitivity of the evidence to dropping these particular observations.

Turn next to Table 1, which provides greater detail about the process by which the population

²⁴Bown and Reynolds (forthcoming) provide an in depth analysis of the trade flow data associated with products prior to the initiation of the dispute, where the years and sources of the information for the policies challenged in each dispute are reported in the accompanying database. One result of interest from that paper is that estimates of the pre-dispute levels of market access at stake and scope of products under dispute are not statistically different for the disputes triggered by policy changes relative to the disputes triggered by the failure to implement a policy change after the Uruguay Round or WTO accession.

²⁵Article 21.3(c) of the WTO’s Dispute Settlement Understanding specifically makes the following suggestion for defining a reasonable period of time for respondents to implement policy changes after rulings: “a guideline for the arbitrator should be that the reasonable period of time to implement panel or Appellate Body recommendations should not exceed 15 months from the date of adoption of a panel or Appellate Body report. However, that time may be shorter or longer, depending upon the particular circumstances.”

of WTO disputes initiated and legally concluded between 1995-2009 is reduced to the sample that is appropriate and available for our modeling and estimation framework. First, the modeling framework described in Section 3 is not necessarily appropriate for all disputes, especially those involving alleged violations to WTO rules affecting a country’s export policies.²⁶ That eliminates 35 of the 307 disputes from consideration for the analysis. Second, we also drop disputes related to services imports or general policies that affect all imports, i.e., those that cannot be matched to any particular products under dispute. The resulting sample is 249 WTO disputes initiated and concluded between 1995 and 2009 that relate to allegations over import policies that can be matched to specific 6-digit Harmonized System (HS-06) import products.

Our estimation procedure also requires matching available disaggregated trade data for volumes and prices for the years around the policy changes taking place prior to and at the conclusion of the WTO dispute. We describe the matching process in greater detail in the next subsection, but our final sample of data used in the estimation includes 167 respondent/complainant pairs, or roughly 70 percent of the total population of 249 WTO disputes initiated and concluded between 1995 and 2009 that targeted imported products.

In our empirical analysis below, we also estimate the model on subsamples of data split by exogenously-determined characteristics, so as to explore the sensitivity of the results. We investigate whether the model explains certain subsamples of disputes better than others, based on differences such as the potential legal status or legal resolution of the process, the type of policy being challenged, or the countries involved.

First, we explore whether the determinants of trade dispute outcomes resolved through the WTO’s legal system are different from those that are settled or withdrawn after dispute initiation but prior to any formal WTO legal rulings. Table 1 reports that a WTO Panel Report, or the first formal legal ruling of importance in any dispute, was issued in slightly less than half (83 out of 167) of the disputes in our sample.

Second, we also characterize the caseload based on how the respondent’s disputed import policy treated (non-complainant) third country exporters of the disputed product. The first type of dispute involves challenges to policies that the respondent imposed on a “global” basis against all trading partners. Examples would include a WTO-inconsistent internal tax, subsidy or domestic regulation that was nevertheless applied on a relatively MFN-conforming basis so as to affect all exporters. The second type of dispute involves challenges to policies that the respondent imposed on a “partial” basis and thus which excluded certain third country exporters. Examples of partial policies would be WTO-inconsistent application of antidumping or countervailing duties or a trade preference scheme.²⁷ Table 1 indicates that in our final sample of 167 disputes, slightly less than half are

²⁶For example, a three country model would be more appropriate to examine the litigation of WTO-inconsistent export subsidies, in which the complainant and respondent are each modeled as having exporters that compete in a common third market and which countries are assumed to have access to export policy instruments as opposed to the import tariffs assumed here. The examination of export restrictions would similarly require an alternative modeling framework that may include different assumptions on available policy instruments, and in that case the complainant would be the importer and the respondent would be the exporter of the disputed product.

²⁷To clarify, our categorization as to whether the disputed policy is “global” versus “partial” is based on our

associated with challenges to “global” policies, and the rest are associated with “partial” policies.

Third, we explore whether our model better fits the high-income respondent subsample of data relative to the low-income importing countries, perhaps, for example, out of the possibility that richer countries may have more market power. As Table 1 indicates, our sample contains a significant number of disputes involving both high-income and developing countries as both complainants and respondents.²⁸

4.3 Variable construction and data

Estimation of equation (13) requires data on trade quantities and prices associated with the products in each dispute. We take this information from a newly constructed database (Bown and Reynolds, forthcoming) which matches disputes to the c.i.f. value and volume (as measured by the net weight in kilograms) of bilateral import data by HS-06 product code from UN Comtrade.²⁹

For each dispute we measure “world” or exporter-received *prices* for the product under dispute from data based on the ratio of the real value of imports associated with the dispute to the netweight (kilograms) of imports, deflating the nominal import value data using the IMF’s world import price index. While we are interested in the impact of changes to prices received by foreign exporters, our unit values are constructed from importer data, which is more reliable than exporter-reported data but which is compiled on a c.i.f. basis and thus includes the insurance and freight costs that arise in getting from the exporter’s (complainant’s) to the importer’s (respondent’s) border. However, because our analysis focuses on changes in these unit prices, our estimates will be unaffected by differences in levels of these freight and insurance costs across products, provided these costs do not change substantially during the period of the dispute.³⁰

Because weight data are not available for all HS-06 products, our approach is to drop any dispute in which volume data is not available for at least 80 percent of the HS-06 product lines. Of the remaining disputes, import volume data is available for almost all HS-06 product lines. For the

judgment of whether the policy excluded some, or was applied to all, third country (non-complainant) exporters of the disputed product. I.e., it is not based on whether the allegation focused on (or was limited to) legal arguments or submissions under GATT Article III (National Treatment) versus Article I (MFN Treatment). Such a characterization would not be possible because some of the disputes in our sample do not move to the stage in which complainants must fully articulate their allegations of respondent misconduct. Furthermore, the two allegations are not mutually exclusive. For a discussion of some of the key economic aspects of National Treatment in the GATT and WTO, see Horn (2006). For a discussion of the role of MFN Treatment, see Horn and Mavroidis (2001).

²⁸Note, however, that there are no least developed countries in our estimation sample, though this is because least developed countries are mostly absent from involvement in WTO disputes altogether, so this phenomenon is not driven by our particular approach.

²⁹Disputes in our sample may target alleged WTO violations affecting products at a finer or more coarse level of aggregation than the HS-06. To the extent that disputes are over traded products at a finer level of aggregation, our approach may mismeasure the volume of imports. Nevertheless, more than 70 percent of disputes in our sample target products measured at the HS-06 or more coarse level of aggregation.

³⁰While freight and insurance rates are probably not time invariant during our sample, our results should be unaffected provided these changes are not correlated with changes in levels of trade protection. Nevertheless, because the error in the change in the exporter prices may be larger the longer is the duration of the dispute (i.e., the larger the difference between year $t = PO$ and $t = BR$), we include the change in crude oil prices interacted with the distance between the respondent and complainant to address changes in transportation costs over the dispute period. We describe this in more detail above when we discuss construction of our instruments.

handful of disputes without volume data for all HS-06 product lines, we drop those HS-06 products with a positive value of imports but missing quantity of imports. This approach provides the best assurance that our price variable, measured in dollars per kilogram, is calculated accurately. As noted earlier, this conservative approach forces us to drop 79 disputes from the population of WTO cases because import volumes are unobserved.³¹

Exporter-received prices exhibit substantial variation in our sample due to the heterogeneous nature of the disputed products. For example, the exporter price associated with the disputes in our sample ranges from \$70 to \$587.20 per kilogram; across the disputes, the standard deviation of prices across all disputes is \$44.16 per kilogram. Rather than estimating the equation using the change in the price level as suggested by equation (12), we deal with this heterogeneity by estimating the model of equation (13) by use of the *percentage change* in the price.

Finally, the elasticity data used to estimate equation (13) is derived from Crowley and Yu (2013). Crowley and Yu (2013) estimate import demand and export supply elasticities for a sample of 11 countries by HS-06 product lines between the years 1988 and 2012. Their estimates are calculated using the structural estimator originally proposed in Broda and Weinstein (2006) and further developed in Soderbery (forthcoming). Dispute-level estimates of elasticities are calculated using a trade value weighted average of the HS-06 product line elasticities associated with each dispute. The Crowley-Yu elasticities are unavailable for the HS-06 product lines and/or complainant countries for approximately 30 percent of our sample. In these cases we approximate the elasticities using the median elasticity in the product line of all other countries within the complainant country’s World Bank income group.³²

Table 2 provides summary statistics for the variables used in the econometric analysis.

4.4 Characteristics of the data on disputes and trade volume outcomes

To fix ideas, consider Figure 2, which presents the time path of the mean growth of import volumes and (foreign) exporter-received prices for the products under WTO dispute, with time normalized around two critical years for each dispute. The first year is one in which the importing country - i.e., the defendant or “respondent” country alleged to have violated WTO rules by imposing an illegal trade restriction - has its best response ($t = BR$) policy imposed. Relative to two years earlier ($t = BR - 2$), the policy is associated with a sharp reduction in import volumes, on average, and a modest reduction in the prices received by the foreign exporter - i.e., the plaintiff or “complainant” in the dispute. On the other hand, two years after the conclusion of the WTO dispute, and by the time the importing country is supposed to have implemented its politically optimal ($t = PO$) trade policy, import volumes have increased, on average, as has the average price received by foreign exporters of the disputed product. Our approach uses model predictions from the terms-of-trade

³¹Such an approach could result in sample selection bias if, for example, low-income countries are less likely to record import volume data than others. The proportions of low income respondents and complainants in our final sample are virtually identical to those in the population of WTO disputes.

³²Developing country respondents account for over three-quarters of the disputes in which we have to approximate elasticities, thus inducing more measurement error into this sub-sample of countries.

theory to explain the variation in the data underlying Figure 2.

Table 3 presents additional information from the data on import volume outcomes arising under the WTO disputes; first consider the full sample of data ultimately to be used in the estimation. The respondent’s bilateral import growth from the complainant associated with the timing of the changes in the disputed policy - i.e., between years $t = BR$ and $t = PO$ (see again Figure 1) - has a mean of 12 percent and a median of 18 percent. On average, 4.57 years pass between $t = BR$ and $t = PO$; furthermore, the mean and median bilateral import growth per annum is 4 and 5 percent, respectively. Sixty-two percent of all disputes result in positive import growth between $t = BR$ and $t = PO$; but in only 56 percent of the disputes is this bilateral import growth larger than the respondent country’s overall level of economic (real GDP) growth over that same period.

The additional columns of Table 3 suggest other interesting results from subsamples of the dispute settlement caseload, a number of which motivate the need for a more formal econometric approach like the one we adopt below. First, disputes that proceed to a formal legal ruling (through a Panel Report) are more likely to result in positive import growth and to result, on average, in slightly higher levels of import growth, than disputes that settle earlier or are withdrawn before any formal legal ruling. While the simple t-tests of differences in the means are not statistically significant at conventional levels, even this result could be viewed as surprising for any theory predicting that disputes that settle early have “better” economic outcomes.³³

Second, there is also little statistical difference in the trade growth outcomes for disputes that confront “global” (as opposed to “partial”) policies. This may also be seen as surprising if, relative to global disputes, partial disputes have complainants that are less likely to face the free-rider problem in which the externality benefits of pursuing litigation spill over to third countries. Note that we also explore econometrically below an additional channel through which the free-rider problem may arise within the subsample of “global” policy violation disputes, as measured by concentration of foreign exporter interests across countries (Ludema and Mayda 2009, 2013).

Third, Table 3 also does not suggest that any pattern to the successful record of market access restoration (import growth) associated with the WTO dispute settlement system has arisen because high-income countries are using it to extract trade liberalization concessions from lower-income countries. For example, high-income respondent countries liberalize their import markets much more substantially on average over the life of these disputes than the developing (respondent) countries. Furthermore, 46 percent of high-income complainants (exporters) experience negative trade growth. On the other hand, disputes initiated by developing country complainants result in average trade growth of 30 percent over the life of the dispute (7 percent per annum). Overall, these potentially unexpected patterns arising from the raw data on trade dispute outcomes motivates the importance of more clearly modeling and estimating the influence of the potential channels of such determinants through a formal econometric analysis.

Finally, compare features of the data sample to be utilized here with the pre-dispute trade

³³Some of the first research exploring the relationship between legal policy concessions and early settlement in GATT-era disputes is Busch and Reinhardt (2000). However, this research did not explore trade flow outcomes.

flow information from the population of disputes collected and surveyed in Bown and Reynolds (forthcoming). For example, one interesting set of characteristics observed in the larger population of disputes involves the vast heterogeneity of market access and scope of products directly affected by the policies under dispute - e.g., 15 percent of cases involve bilateral (complainant-respondent) trade in disputed products of more than \$1 billion per year, while another 14 percent of disputes involve bilateral trade in disputed products of less than \$1 million per year. Furthermore, to the extent that product counts may affect the ability of exporters to organize collectively and overcome the free rider problem in trade policy determination, 20 percent of disputes consider policies affecting only one HS-06 product, whereas another 20 percent of disputes examine policies affecting 50 or more HS-06 products at once.

While the data constructed for the econometric analysis reported below tends to lose observations from the population of disputes disproportionately at the very low end of the distribution,³⁴ it is worth noting that a number of qualitative similarities are preserved: e.g., 8 percent of disputes in the current sample had (pre-best response year) trade of less than \$5 million per year, and 16 percent of disputes had trade that was more than \$1 billion. Furthermore, 12 percent of disputes in this sample address policies directly affecting only 1 HS-06 product, whereas 22 percent of disputes address more than 50 products.

5 Econometric Results

This section reports econometric regression results from estimating equation (13) on a sample of WTO disputes initiated and legally concluded over the period 1995-2009.

5.1 Disputed product import volume growth, the terms-of-trade, and trade elasticities

Table 4 provides our first estimates from equation (13) of the determinants of the growth of disputed product bilateral import volumes of the respondent from the complainant, $[\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})]$. The pre-dispute best response (*BR*) and post-dispute politically optimal (*PO*) years are tied to the timing of the changes in the disputed policy, as illustrated in Figure 1.

Table 4 begins by providing OLS estimates of the impact of the theoretically motivated variables of interest, also ignoring momentarily the potential influence of other control variables. In column (1), the OLS estimated impact of $[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})]$ is negative and the estimated impact of $[(\sigma_{gr}^{BR}/\omega_{gc}^{*BR})\ln(M_{grc}^{BR})]$ is positive, as suggested by the theory, and both estimates are statistically different from zero.

Beginning in column (2), we iteratively introduce our IV estimation, one variable at a time, so as to understand the source of potential underlying differences in the estimates.³⁵ In column (2), for

³⁴ As described earlier, we lose observations due to missing trade volume data, and these losses are disproportionately in disputes where there are also low values of import market access at stake and few HS-06 product codes.

³⁵ As noted in Table 4, the Cragg-Donald Wald F-statistics are well over 10 for most of our specifications, indicating that the excluded variables are valid instruments for the endogenous variables. Similarly, the p-values associated with

example, we instrument only for $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$; while the magnitude of the coefficient estimate on $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$ is reduced, the qualitative pattern to the results is largely unaffected. In column (3), we instrument only for $[(\sigma_{gr}^{BR}/\omega_{gc}^{*BR})\ln(M_{grc}^{BR})]$, and this also results in only a small impact on the estimated coefficients. Column (4) presents our first specification with our full set of instruments to estimate the model; both estimated coefficients of interest in the IV regression are of the theoretically predicted sign and are statistically different from zero. Finally, column (5) presents an initial check on the sensitivity of our estimates to the definition of the timing of the best response year. In column (5), we substitute the year of the initiation of the WTO dispute for our standard definition of the best response year, i.e., the year of the identified policy change that is being challenged in the dispute. The result is little change to the estimated coefficients.

Column (6) of Table 4 presents our preferred specification to the baseline model. In addition to using the full set of instrumental variables, we also introduce a number of variables to control for other factors expected to affect the growth of bilateral, disputed product import volumes between $t = BR$ and $t = PO$. The next section provides a detailed introduction to each variable and a discussion of its estimated impact. Here, we stress that while inclusion of the controls slightly reduces the magnitude of the IV coefficient estimates on $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$ and $[(\sigma_{gr}^{BR}/\omega_{gc}^{*BR})\ln(M_{grc}^{BR})]$, they remain of the theoretically predicted sign and statistically different from zero.

Consider next the magnitudes of the IV estimates in column (6). The estimate of -0.997 on $[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})]$, indicates that disputed product import volume growth will be 0.997 percentage point less for every 1 percentage point increase in foreign-exporter received prices (terms-of-trade loss) associated with the respondent importing country's policy change resulting from the dispute settlement negotiations. Furthermore, the coefficient estimate on $[(\sigma_{gr}^{BR}/\omega_{gc}^{*BR})\ln(M_{grc}^{BR})]$ of 0.001 implies that a one standard deviation increase in this variable will result in an 11 percentage point increase in disputed product import volume growth; i.e., from the baseline prediction (at the means of the data) of 12 percent to nearly double this rate to 23 percent. While this is economically meaningful, the analogous exercise of considering a one standard deviation improvement in the respondent's terms-of-trade (reduction in foreign exporter-received prices) relative to the mean level is a much more sizeable impact, leading to a nearly 70 percentage point increase above the mean prediction for import volume growth.

5.2 The effect of other determinants on disputed product import growth

In addition to the key determinants of interest arising from the theoretical model discussed in the last section, the baseline specification of Table 4, column (6) includes a set of covariates that the prior literature on trade agreements as well as economic theory would predict impact the trade volume growth outcome of negotiations under WTO dispute settlement. For ease of exposition, we group these potential determinants into two broad categories: (i) other political-economic determi-

the Sargan test of overidentifying restrictions confirm that the excluded variables are uncorrelated with our error term. The two Appendix Tables A and B provide a complete reporting of the first stage coefficient estimates associated with the model specifications presented in Table 4.

nants associated with trade agreements, and (ii) other demand- and supply-side determinants and disputant (country-level) characteristics.

5.2.1 Retaliation capacity, exporter concentration, and free riding

The first set of additional covariates introduced in Table 4, column (6) include other political-economy forces through which the institutional process of WTO dispute resolution is likely to affect the trade policy negotiation outcomes.

The first variable is designed to address the self-enforcing nature of WTO dispute settlement and that particular bilateral trading relationships can exhibit asymmetries that may render difficult (or, in the limit, even meaningless) WTO-sanctioned retaliation that any particular complainant country may be authorized to implement against a respondent (Bown, 2004a).³⁶ To control for variation across disputed product import volume growth explained by differences in bilateral retaliation capacities, we introduce a newly constructed and institutionally-motivated variable defined as the ratio of the value of the complainant’s total ‘actionable’ goods imports from the respondent over which the complainant could potentially retaliate relative to the value of imports at stake in the dispute. Construction of the denominator draws on insights from Bagwell and Staiger (1999) and Bown and Ruta (2010).³⁷ The numerator for this retaliation capacity variable on ‘actionable’ imports focuses on bilateral trade in differentiated products using the measure first proposed by Rauch (1999).³⁸ We expect the growth of disputed product import volumes to be increasing in this ratio, i.e., increasing in the level of actionable imports and decreasing in the value of disputed product trade that is at stake in the case.

The evidence in column (6) is a statistically significant and positive relationship between the retaliation capacity variable and the growth of disputed product import volumes. I.e., the larger

³⁶While not part of our estimation sample because the underlying dispute involves a violation over trade in services, the canonical example in which the lack of retaliation capacity is likely to be an important contributor to the inability of the respondent country to comply with adverse WTO legal rulings involves the tiny islands nation of Antigua and Barbuda dispute over Internet gambling brought against the United States. While there have been proposals that countries implement retaliation by something other than goods (tariff) retaliation, the reality is that as of this writing, there has yet to be a case in which a country actually *imposed* retaliation in some form other than tariff retaliation, partially because there are substantial implementation difficulties in doing so. For an extensive discussion of these issues, see the contributions in Bown and Pauwelyn (2010). For an important theoretical contribution examining the role of bilateral trade asymmetries and enforcement of cooperative low tariffs in a repeated game model of trade agreements, see Maggi (1999).

³⁷Bown and Ruta (2010, p. 159) identify one version of the Bagwell and Staiger (1999) definition of reciprocity as the retaliation being limited to that value which makes the absolute change in the complainant’s import volume from the respondent resulting from the retaliation (evaluated at the initial world price) equal to the absolute change in the respondent’s import volume from the complainant resulting from the WTO violation (evaluated at the initial world price). The measure used here in the denominator (the value of imports at the initial world price, $p_{grc}^{w, BR-2} M_{grc}^{BR-2}$) is a proxy for this relationship that serves as the upper bound measure for the amount of trade over which the WTO might authorize retaliation. I.e., if the imposed best response policy were to be prohibitive, so that $M_{grc}^{BR} = 0$, then $p_{grc}^{w, BR-2} M_{grc}^{BR-2}$ is the level of retaliation defined in Bown and Ruta (2010, p. 159).

³⁸Our measure of actionable imports thus strips out trade in homogeneous goods. See, for example, Evenett (2010). Homogeneous goods may not be targeted because they are lower profile and because losses to the respondent would be smaller due to the possibility of arbitrage via the respondent shifting sales to third markets. For more on the anecdotal evidence of the products that retaliating (complainant) governments place onto their retaliation lists, see again the contributions in Bown and Pauwelyn (2010).

is the ratio of complainant’s imports from the respondent relative to the value at stake in the dispute, the more capacity the complainant has to find and exert some meaningful trade retaliation threat to help facilitate respondent country compliance, policy reform, and trade liberalizing import growth as an outcome of the dispute negotiations. The impact is also economically significant; the coefficient estimate of 0.084 implies that a one standard deviation increase in this variable will result in an 18 percentage point increase in disputed product import volume growth; i.e., more than doubling import growth from the baseline prediction (at the means of the data) of 12 percent to 30 percent. Finally, and as will become clearer from our continued discussion below, we can further rule out that this retaliation capacity variable is instead capturing only a size effect (i.e., that larger countries get more trade liberalization than smaller countries, regardless of retaliation capacity) because we separately control for complainant and respondent market sizes through inclusion of measures of each country’s real GDP.

A second important political-economic variable of interest is motivated by the Ludema and Mayda (2013) empirical result that the collective action problem also affects the formation of trade policy. They find that under GATT negotiating rounds and in instances in which exporting countries have incentive to free ride (Olson, 1965), such countries may fail to negotiate jointly with an importing country and liberalize tariffs. Under our approach, and for the set of disputes in which the policy being challenged was a “global” violation that negatively affected all exporting countries and not just the complainant, we consider the impact of the concentration of the exporting countries in the respondent’s import market of the disputed product.³⁹ However, whereas Ludema and Mayda (2013) use the Herfindahl-Hirschman index (the sum of squared export shares) as a measure of exporter concentration, our exporter concentration measure is defined as simply the share of the respondent’s import market that is comprised of the exporting countries that are legally involved in the dispute - i.e., the complainant and all self-identified “interested third party” countries that participate in the formal dispute.⁴⁰

For the “global” violation subset of disputes, column (6) presents evidence that the higher is the concentration amongst the exporting countries legally involved in the dispute, the larger is the subsequent import volume growth arising from the dispute. Put differently, our results on the outcomes arising under WTO dispute settlement negotiations are consistent with the Ludema

³⁹For this reason in column (6) we interact the exporter concentration variable with an indicator for whether the underlying disputed policy was an alleged “global” violation. Put differently, the relevance of the concentration of the exporters in the subsample of disputes over “partial” policy violations is complicated by the fact that the complainant may have been the only exporting country that was negatively impacted by the imposed policy, and the result is an implicit preference associated with discrimination against the complainant. In standard economic models, such a preference would provide third (exporting) countries with an economic interest in the dispute that is aligned with the import-competing industry in the respondent as opposed to exporters of the disputed product in the complainant; i.e., third countries would not necessarily seek to join in the dispute in order to negotiate collectively *with* the complainant country.

⁴⁰To clarify, Ludema and Mayda (2013) do not examine the trade dispute setting to test their theory; their empirical setting focuses on the MFN tariffs that WTO members negotiated and implemented as of the conclusion of the Uruguay Round negotiations in 1995. Furthermore, and unlike our trade dispute setting in which negotiating exporting countries reveal themselves directly through the legal process, Ludema and Mayda rely on the Herfindahl-Hirschman index because the multilateral tariff round negotiations setting does not reveal information on which exporting countries directly participated in the trade liberalization negotiations.

and Mayda (2013) findings on the outcomes arising under WTO MFN tariff-setting negotiations that product markets with more diffuse exporter interests (across countries) result in less trade liberalization by the importing country. And for comparison purposes with the magnitudes of our other estimates, the coefficient estimate of 0.649 implies that a one standard deviation increase in this variable will result in a 22 percentage point increase in disputed product import volume growth; i.e., from the baseline prediction (at the means of the data) of 12 percent to nearly triple this amount at 34 percent.

5.2.2 Other demand- and supply-side determinants and country characteristics

The next set of covariates that we introduce in the baseline specification of Table 4, column (6) are other general demand- and supply-side determinants of disputed product import volume growth during the period of the dispute. We address the general nature of demand shocks by including the respondent's real GDP growth between $t = BR$ and $t = PO$ and find the expected positive relationship, i.e., importing countries enjoying faster rates of overall growth are also more likely to experience more disputed product import volume growth. On the other hand, controlling for supply side shocks at the aggregate level does not lend similarly positive evidence of a relationship between complainant country economic growth during the period and disputed product trade growth, as the estimated coefficient is negative, though not statistically different from zero. Finally, we also include a dummy variable for disputes in which the politically optimal or best response year associated with the dispute (coincidentally) takes place during the 2008-9 period of the global trade collapse. While the coefficient estimate is negative as expected for the indicator that 2008-9 was the politically optimal year associated with the dispute, it is not statistically different from zero.

The third set of covariates introduced in column (6) are disputant-level economic characteristics designed to examine whether variation in dispute settlement outcomes is affected by country size (real GDP) or levels of economic development (real GDP per capita). The coefficient estimates on these variables suggest that, once we control for the other determinants of disputed product import growth described above, there is little evidence that country size or its level of development has a statistically significant independent effect on the import volume growth arising from the negotiations that take place under WTO dispute settlement.

The final sets of columns (7) and (8) in Table 4 introduce country-level (complainant and respondent) fixed effects and then country-level *and* sector-level fixed effects, respectively. Not surprisingly in column (7), inclusion of the country-level fixed effects soaks up most of the variation used to identify the impact of the control variables introduced in column (6).

Most important is that even inclusion of the country-level fixed effects in (7) does not change the qualitative nature of the IV estimates on the theoretically-motivated variables of interest, $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$ and $[(\sigma_{gr}^{BR}/\omega_{gc}^{*BR})\ln(M_{grc}^{BR})]$. However, while leaving unaffected the estimate on $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$, inclusion of sector-level fixed effects in column (8) does eliminate the statistical significance on the estimated impact of $[(\sigma_{gr}^{BR}/\omega_{gc}^{*BR})\ln(M_{grc}^{BR})]$. This calls into question the robustness of the ratio of the trade elasticities multiplied by the best-response import volume

of the disputed product.⁴¹ Nevertheless, while we are left with some skepticism for the estimated impact of $[(\sigma_{gr}^{BR}/\omega_{gc}^{*BR})\ln(M_{grc}^{BR})]$, the estimated impact of the direct role of the change in exporter-received prices, $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$, is robust. We consistently find evidence consistent with the theory that disputed products with larger terms-of-trade losses can be linked to smaller levels of growth of trade volumes resulting from negotiated dispute settlement outcomes.

To summarize the results thus far, our IV approach provides evidence consistent with the theory that the Bagwell and Staiger (1999, 2011) model of trade agreement negotiations also applies to the trade liberalization negotiations under the WTO’s formal dispute resolution procedures. In our full sample of disputes, model estimates are robust to inclusion of a variety of additional covariates, as well as country- and sector-level fixed effects and provide evidence that the growth in disputed product import volumes is decreasing in the size of the respondent importing country’s terms-of-trade loss, i.e., the price increase received by the foreign exporter of the disputed product associated with liberalization of the market. Second, we also provide some evidence that growth in disputed product import volumes is increasing in the ratio of the product’s import demand elasticity to the foreign export supply elasticity, when interacted with the best response (pre-dispute) import volume. Such evidence would be consistent with the theory that import volume growth is expected for products in which the import demand elasticity is high – so that the alleged WTO-violating trade restriction was imposing significant (own) distortions on the respondent’s importing economy – relative to a low foreign export supply elasticity. Nevertheless, this result is weak and sensitive to variable definition and measurement of the underlying trade elasticities.

Our approach also extends evidence provided in other contexts on the determinants of trade policy formation. We provide new evidence consistent with earlier research that the bilateral retaliation capacity of the complainant affects trade dispute outcomes measured as disputed product import growth (Bown, 2004a). Finally, we also find that trade dispute outcomes are affected by the ability of exporting countries to overcome the free rider problem of collective action in order to put pressure on the respondent to liberalize its import market (Ludema and Mayda, 2013). This result holds for disputes in which the policy being challenged is relatively MFN-conforming in nature, defined as policies likely to negatively impact the exports of all other WTO member countries and not only the complainant.

⁴¹Following the approach suggested by Broda, Limão, and Weinstein (2008) to check for concerns over potential measurement error and outliers, we note that our results for the impact of the estimated elasticities ratio variable are also sensitive to slight changes in its definition - e.g., using logs of the ratio in lieu of levels, indicators for “high” versus “low” ratios in lieu of the actual values for the elasticities, etc. We have also explored - with only mixed results - replacing the current elasticities (that provide the best country and disaggregated product coverage) with the foreign export supply elasticity estimates from Broda, Limão, and Weinstein (2008) and the import demand elasticities from Kee, Nicita and Olarreaga (2008). (The export supply elasticities from Broda, Limão, and Weinstein (2008) are measured at HS-04 level and are for a different set of countries; thus we use the median elasticity in the country’s World Bank income group to approximate the elasticities for roughly two-thirds of the sample. As a result, such specifications use a less precise measure of the elasticity associated with each dispute.) While the results for the elasticity ratio variable are sensitive, nevertheless, the estimate on $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$ is consistently unaffected by these alternative approaches.

5.3 The impact of institutional features on dispute settlement outcomes

We have established that the basic terms-of-trade model and estimating framework can be used to understand changes in trade volumes that result from trade liberalization negotiations taking place under WTO dispute settlement. This section extends the analysis by implementing our preferred model specification on different subsamples of data, where we allow institutional features of the trade agreement and dispute settlement process to establish the relevant subsamples under consideration. Table 5 presents the estimates; and for ease of exposition, column (1) re-establishes the benchmark estimates from the full sample of dispute outcomes data, repeating the results presented earlier in Table 4, column (6). While each of the specifications includes all of the control variables in Table 4, column (6) and are available upon request, we do not report them in Table 5 to conserve space.

Table 5, column (2) estimates the model on only the sample of trade dispute outcome observations for which we observe information about the timing of the change in the post-dispute policy. I.e., recall from Table 1 that in 28 of the 167 WTO disputes we are unable to identify and verify whether any policy change resulted from the WTO dispute. A legitimate question of interest for our estimates is thus whether the results are driven by inclusion of such disputes in the sample, which could be problematic if that particular part of the sample suffers from a different form of measurement error.⁴² The sample in column (2) drops such disputes; the resulting estimates are reassuring in that the qualitative pattern of evidence is mostly unchanged.

Column (3) next considers a subsample that eliminates all WTO disputes in which the complainant has filed a challenge to US use of antidumping, one of the most frequent topics arising under dispute settlement. Bown and Crowley (2013) interpret US antidumping use over 1997-2006 as motivated by terms-of-trade incentives and provide empirical evidence consistent with a related Bagwell and Staiger (1990) theory. But since many of the US antidumping import restrictions empirically explained by the Bown-Crowley estimation framework were subsequently challenged by formal WTO dispute settlement, they are therefore part of the current paper’s sample of observations. Nevertheless, column (3) drops disputes to US antidumping and the estimates allow us to conclude that the current empirical results are not limited to that particular slice of the WTO dispute settlement caseload.

The estimates in columns (4) and (5) assess a more general split of the dispute sample by characterizing the discriminatory nature of the allegedly WTO-violating policy that is under dispute. Column (4) reports model estimates from the WTO challenges to policies that the respondent applied on a “global” basis, and thus which negatively affect all other WTO member countries in addition to the complainant. Column (5) reports model estimates from the WTO challenges to policies applied on a “partial” basis, in which some other WTO member countries are excluded from the policy and thus exporting countries arise with the potential to enjoy an implicit preference

⁴²Because we do not observe any announced policy changes for these 28 disputes, we are unable to distinguish whether these disputes actually resulted in no policy change or whether they resulted in an unreported policy change, and for which we would then potentially mis-assign the post-dispute politically optimal year.

(relative to the complainant country) in the respondent country’s disputed product import market. Given the relatively small remaining subsamples of disputes, our purpose is admittedly not to strictly test for differences in parameter estimates. Our more modest goal is to investigate whether certain characteristics of the policy under dispute helps to explain the extent to which our basic modeling approach to trade liberalization negotiations under dispute settlement can potentially explain patterns in the data. With the exception of a couple of the control variables, the model estimates across these two particular subsamples of data are not particularly different from one another.

Columns (6) and (7) split the sample in two based on whether the institutional resolution to the WTO dispute was the issuance of a formal legal ruling (of at least a Panel Report) versus no legal ruling whatsoever. Disputes that did not have at least a Panel Report were either settled early or dropped by the complainant. Specification (6) reveals that the subsample of disputes that received a legal ruling have estimates that are consistent with theoretical predictions and, for the most part, are statistically significant. Furthermore, for most of the key determinants, the estimates are generally larger and more consistently statistically significant than the estimates from the full sample of disputes, as well as the subsample of disputes without legal decisions. This is both interesting and potentially reassuring for a number of reasons. First, examination of these two different subsamples of data is arguably important, given that some models of the dispute settlement process (e.g., Maggi and Staiger, 2015b) contain different predictions for policymaking in cases that reach a legal ruling versus those that settle early. Second, the subsample of disputes without a legal decision contains more observations with uncertainty as to the *timing* of policy changes, and this is likely to translate into additional measurement error for our variables constructed from the trade data.⁴³ Third, there is also the possibility that disputes without legal decisions were “weaker” disputes and are ones for which our dispute settlement modeling framework for trade liberalization may not necessarily apply.

The last two columns of Table 5 utilize the current modeling approach to explore in more depth the potential impact of the complainant country’s capacity to retaliate to influence the trade volume outcomes arising from dispute negotiations. These columns allow us to focus on some of the issues first introduced in Bown (2004a), which examined changes in imports arising as dispute outcomes during an earlier period (1973-1998); i.e., the latter half of the GATT period and the first three years of the WTO.⁴⁴ While the previous research also examined the retaliation capacity of the complainant vis-a-vis the respondent and relied on a number of similar control variables, the main difference is that the earlier approach ignored the *within*-market implications of the terms-of-trade theory for the products under dispute that is the emphasis here. The within-market implications for liberalization of such product markets arises via our formal application of the Bagwell and Staiger

⁴³Note that in 56 of these disputes we observe the timing of the change of policy due to legal filings and other government documents. In the remaining cases we set $t = PO$ as three years after the last formal correspondence between the two litigating parties relating to the dispute.

⁴⁴In addition to different years for the sample of data, the two approaches are also not strictly comparable given that Bown (2004a) focused on the growth in import values, whereas the current approach examines data on import *volumes* only, so as to examine the particular influence of changes in prices on trade outcomes.

(1999, 2011) theory and empirical approach to modeling trade liberalization negotiations.

To explore the implications for retaliation capacity, we divide the sample based on values of our retaliation capacity variable, described earlier as the ratio of actionable (non-homogeneous) imports available for the complainant's retaliation relative to the value of trade at stake in the dispute. The higher is this ratio, the less constrained is the complainant to identify potential goods for retaliation relative to the maximum amount of trade over which the WTO could potentially authorize the complainant to retaliate.

Column (8) provides estimates from the half of the data sample for which the ratio is above the median and in which the complainant is much less constrained in its capacity to implement a potentially WTO authorized retaliation. The model's estimates are generally aligned with theoretical predictions and are statistically significant. The other half of the sample is defined by cases in which the complainant is more constrained in its capacity to implement a sufficient level of retaliation that the WTO might authorize. In column (9), the model performs less well, though the coefficient estimates for the theoretically-motivated covariates are so imprecisely estimated that we cannot rule out that they are equivalent to the estimates from high retaliation capacity subsample. Finally, it is also instructive to note that the retaliation capacity variable is not significantly different from zero even in the high retaliation capacity subsample of disputes. Combined, these results suggest that there may be a threshold level of retaliation capacity for which the terms-of-trade model is useful to explain trade dispute negotiation outcomes; however, marginal increases to retaliation capacity has little additional effect.

5.4 Dispute settlement outcomes and economic characteristics

Table 6 provides a last set of estimates of our preferred, baseline model specification on alternative subsamples of data defined by their (exogenously determined) economic characteristics. Again, column (1) provides estimates for the baseline specification from the entire sample of disputes as a point of reference.

First, we examine specification on different subsamples of data based on income categories for the countries involved in the disputes. Columns (2) and (3) split the sample in two based on the income status of the respondent country, which, in our framework, is the importing country that has been alleged to have implemented a WTO-inconsistent trade restriction. Columns (4) and (5), on the other hand, split the sample in two based on the income status of the complainant (exporting) country that has initiated the WTO dispute.

Table 5's specifications (2) through (5) indicate that our evidence in support of the terms-of-trade theory is driven by the subsamples in which the complainant and/or the respondent is a high-income WTO member country. In particular, the change in foreign exporter-received prices (the terms of trade loss), the ratio of elasticities interacted with pre-dispute import volumes, and the exporter concentration variable are each of the theoretically predicted sign and are statistically significant in the high-income country specifications (2) and (4). To the extent that high-income countries may be more likely to form trade agreements and initiate the enforcement provisions of

trade agreements so as to neutralize terms-of-trade externalities, these results are not surprising.

On the other hand, the model is admittedly less successful in using the terms-of-trade theory to explain variation across growth in disputed product imports for the developing country respondent (specification 3) or complainant (specification 5) subsamples. The smaller sizes for the developing economy samples may partially explain the lack of statistical significance of the determinants of interest, as may the fact that the elasticities are less precisely measured in these subsamples, as discussed above. However, a separate and theoretically-motivated explanation is that developing countries in particular may not seek entry into trade agreements for terms-of-trade reasons in the first place, but instead if for time-consistency or bargaining reasons (Maggi and Rodriguez-Clare, 1998; Limão and Tovar, 2011) they lack the ability to unilaterally commit their private sectors to a policy of more liberal trade. While our approach does not provide a formal test of the commitment theory, our failure to find supportive evidence of the terms-of-trade theory for the subsample of developing countries is at least consistent with these countries pursuing trade agreements and its enforcement provisions for other motives.⁴⁵ Other possible contributing explanations include Ludema and Mayda (2013), that the lack of overall exporter concentration in key products of interest to developing country exporters may also help explain why terms-of-trade incentives may not have been neutralized (through tariff-cutting trade liberalization) upon entry into the agreement in the first place.

Finally, specifications (6) and (7) split the sample exactly in two depending on whether the dispute involved high or low growth in import volumes between the best response year at the beginning of the dispute and the politically optimal year at the dispute’s conclusion.⁴⁶ This is one way to investigate whether our results are being driven by the relatively “successful” (high import growth) or “unsuccessful” (low import growth) outcomes. The estimates of the key coefficients are not statistically different across the two subsamples. If anything, the slightly larger coefficient estimates on the retaliation capacity and exporter concentration variables in the “unsuccessful” subsample suggest that the model may be particularly useful in identifying important determinants behind the relatively less successful negotiated outcomes (at least as measured by this metric), where little import growth occurred.

6 Conclusion

This paper examines implications of the terms-of-trade theory for the enforcement of international trade agreements. We extend the Bagwell and Staiger (1999, 2011) theoretical and empirical approach from the setting of original trade agreement negotiations to the setting of formal trade *dispute* negotiations. While the size of the policy movements associated with disputes are notoriously difficult to measure accurately because disputed policies are typically applied as non-tariff

⁴⁵Our evidence from WTO disputes is also broadly consistent with separate results in the gravity literature that the WTO as an *institution* can have differential trade effects for developing country members relative to high-income country members (Subramanian and Wei, 2007).

⁴⁶In particular, we split the sample in two based at median level of the disputed product’s import growth per annum that took place over the period of the dispute.

barriers, the timing of such policy changes are readily observable. We therefore use an explicit theory to map what we can better observe and measure, namely data on trade volumes and prices, to an empirical framework to estimate the determinants of trade dispute outcomes.

We estimate the model on a new data set of the trade volume outcomes deriving from formal WTO disputes initiated and legally concluded between 1995 and 2009, and we provide evidence that determinants of the trade flows that result from dispute settlement negotiations are consistent with theoretical predictions. In particular, larger post-dispute trade volume outcomes are associated with products that have smaller increases to exporter-received prices (terms-of-trade losses) resulting from the conclusion of the dispute, larger pre-dispute trade volumes, larger import demand and smaller foreign export supply elasticities. The resulting estimates on these theoretical predictions are economically significant. Finally, we are also able to confirm that results from related research apply to this setting of trade policy negotiations; i.e., that negotiated trade policy outcomes arising under dispute settlement are also impacted by asymmetries in bilateral retaliation capacities and variation in exporting country concentration, the latter of which contributes to identifying the potential severity of the free rider problem of collective organization.

Overall, our results provide further evidence not only of the relevance of the terms-of-trade theory of trade agreements, but also that terms-of-trade motives are not completely extinguished when countries take on binding commitments under trade agreements such as the WTO. The determinants of the trade volume response associated with respondent importing countries changing their policies after formal WTO dispute settlement include concern for direct, within-market terms-of-trade losses (increases in exporter-received prices for the product under dispute) as well as the potential desire to avoid terms of trade losses in other markets (decreases in own exporter-received prices) under potentially WTO-authorized trading partner retaliation. A final implication of our results is to identify an area for continued exploration for the emerging theoretical literature on dispute settlement provisions in trade agreements; i.e., the role of enforcement provisions for trade agreements that are modeled as both incomplete contracts *and* motivated by the desire to coordinate policy changes in order to address terms-of-trade externalities.

Nevertheless, our evidence does leave unanswered the question of what determines the negotiated outcomes for the significant number of disputes involving developing countries. While the lack of relevance of the terms-of-trade model for the developing country disputes may suggest that such countries are using the WTO as a commitment device vis-a-vis their private sectors rather than to neutralize terms-of-trade externalities (Maggi and Rodriguez-Clare, 1998; Limão and Tovar, 2011), it does not resolve the question of what specific purposes developing countries have in mind when they sign onto trade agreements like the WTO. In particular, it remains unclear how the external enforcement of this commitment device motive is implemented in practice if not through dispute settlement (Bown and Hoekman, 2008). A better understanding of these questions is an important area for future research.

References

- Bagwell, K., C.P. Bown, and R.W. Staiger (2015) "Is the WTO Passé?" Unpublished manuscript, Dartmouth College, April.
- Bagwell, K. and R.W. Staiger (1990) "A Theory of Managed Trade," *American Economic Review* 80(4): 779-795.
- Bagwell, K. and R.W. Staiger (1999) "An Economic Theory of GATT," *American Economic Review* 89(1): 215-248.
- Bagwell, K. and R.W. Staiger (2002) *The Economics of the World Trading System*. Cambridge, MA: The MIT Press.
- Bagwell, K. and R.W. Staiger (2011) "What Do Trade Negotiators Negotiate About? Empirical Evidence from the World Trade Organization," *American Economic Review* 101(4): 1238-73.
- Beshkar, M. (2010) "Trade Skirmishes and Safeguards: A Theory of the WTO Dispute Settlement Process," *Journal of International Economics* 82(1): 35-48.
- Beshkar, M. (2014) "Arbitration and Renegotiation in Trade Agreements," Unpublished manuscript, Indiana University, November.
- Bown, C.P. (2002) "The Economics of Trade Disputes, the GATT's Article XXIII and the WTO's Dispute Settlement Understanding," *Economics and Politics* 14(3): 283-323.
- Bown, C.P. (2004a) "On the Economic Success of GATT/WTO Dispute Settlement," *The Review of Economics and Statistics* 86(3): 811-823.
- Bown, C.P. (2004b) "Trade Disputes and the Implementation of Protection under the GATT: An Empirical Assessment," *Journal of International Economics* 62(2): 263-294.
- Bown, C.P. and M.A. Crowley (2013) "Self-Enforcing Trade Agreements: Evidence from Time-Varying Trade Policy," *American Economic Review* 103(2): 1071-1090.
- Bown, C.P. and B.M. Hoekman (2008) "Developing Countries and Enforcement of Trade Agreements: Why Dispute Settlement Is Not Enough," *Journal of World Trade* 42(1): 177-203.
- Bown, C.P. and P.C. Mavroidis (Eds.) (2013) *The WTO Case Law of 2011*. Cambridge, UK: Cambridge University Press.
- Bown, C.P. and J. Pauwelyn (Eds.) (2010) *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*. Cambridge, UK: Cambridge University Press.
- Bown, C.P. and K.M. Reynolds (forthcoming) "Trade Flows and Trade Disputes," *Review of International Organizations*.

- Bown, C.P. and M. Ruta (2010) "The Economics of Permissible WTO Retaliation," in C.P. Bown and J. Pauwelyn (eds.), *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*. Cambridge, UK: Cambridge University Press, 149-193.
- Broda, C. and D. Weinstein (2006), "Globalization and the Gains from Variety," *Quarterly Journal of Economics* 121(2): 541-585.
- Broda, C., N. Limão, and D. Weinstein (2008) "Optimal Tariffs and Market Power: The Evidence," *American Economic Review* 98(5): 2032-65.
- Busch, M.L. and E. Reinhardt (2000) "Bargaining in the Shadow of the Law: Early Settlement in GATT/WTO Disputes," *Fordham International Law Journal* 24(1): 158-172.
- Crowley, M.A. and M. Yu (2013) "Market-Specific Cost Shocks and Firm Export Behavior," *University of Cambridge working paper*, September.
- Evenett, S. (2010) "Sticking to the rules: quantifying the market access protected by WTO retaliation," in C.P. Bown and J. Pauwelyn (eds.), *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*. Cambridge, UK: Cambridge University Press, 198-231.
- GATT. (1994) *Uruguay Round Agreement Establishing the World Trade Organization, Including the Agreement on Safeguards and the Uruguay Round Understanding on Rules and Procedures Governing the Settlement of Disputes*. Geneva: GATT.
- Grinols, E.L. and R. Perrelli (2006) "The WTO Impact on International Trade Disputes: An Event History Analysis," *The Review of Economics and Statistics* 88(4): 613-624.
- Horn, H. (2006) "National Treatment in the GATT," *American Economic Review* 96(1): 394-404.
- Horn, H., G. Maggi, and R.W. Staiger (2010) "Trade Agreements as Endogenously Incomplete Contracts," *American Economic Review* 100(1): 394-419.
- Horn, H. and P.C. Mavroidis (2001) "Economic and Legal Aspects of the Most-Favored-Nation Clause," *European Journal of Political Economy* 17(2): 233-279.
- Horn, H. and P.C. Mavroidis (Eds.) (2003) *The WTO Case Law of 2001*. Cambridge, UK: Cambridge University Press.
- Horn, H. and P.C. Mavroidis (2008) "The WTO Dispute Settlement Dataset: 1995-2006," *World Bank Working Paper*.
- Horn, H., P.C. Mavroidis, and H. Nordstrom (2005) "Is the Use of the WTO Dispute Settlement System Biased?" in P.C. Mavroidis and A.O. Sykes (eds.) *The WTO and International Trade Law/Dispute Settlement*. Cheltenham, UK: Edward Elgar.
- Jackson, J.H. (1997) *The World Trading System: Law and Policy of International Economic Relations*. (2nd edition) Cambridge, MA: The MIT Press.

- Kee, H.L., A. Nicita, and M. Olarreaga (2008) "Import Demand Elasticities and Trade Distortions," *The Review of Economics and Statistics* 90(4): 666-682.
- Khandelwal, A. (2010), "The Long and Short (of) Quality Ladders," *Review of Economic Studies* 77(4): 1450-1476.
- Limão, N. and K. Saggi (2008) "Tariff Retaliation versus Financial Compensation in the Enforcement of International Trade Agreements," *Journal of International Economics* 76(1): 48-60.
- Limão, N. and K. Saggi (2013) "Size Inequality, Coordination Externalities and International Trade Agreements," *European Economic Review* 63: 10-27.
- Limão, N. and P. Tovar (2011) "Policy Choice: Theory and Evidence from Commitment via International Trade Agreements," *Journal of International Economics* 85(2): 186-205.
- Ludema, R. (2001) "Optimal International Trade Agreements and Dispute Settlement Procedures," *European Journal of Political Economy* 72(2): 355-376.
- Ludema, R. and A.M. Mayda (2009) "Do Countries Free Ride on MFN?" *Journal of International Economics* 77(2): 137-150.
- Ludema, R. and A.M. Mayda (2013) "Do Terms-of-Trade Effects Matter for Trade Agreements? Theory and Evidence from WTO Countries," *Quarterly Journal of Economics* 128(4): 1837-1893.
- Maggi, G. (1999) "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review* 89(1): 190-214.
- Maggi, G. (2014) "International Trade Agreements," in G. Gopinath, E. Helpman and K. Rogoff (Eds.), *The Handbook of International Economics*, vol. 4. Amsterdam, NL: Elsevier.
- Maggi, G. and A. Rodriguez-Clare (1998) "The Value of Trade Agreements in the Presence of Political Pressures," *Journal of Political Economy* 106 (2): 574-601.
- Maggi, G. and R.W. Staiger (2011) "The Role of Dispute Settlement Procedures in International Trade Agreements," *Quarterly Journal of Economics* 126(1): 475-515.
- Maggi, G. and R.W. Staiger (2015a) "Optimal Design of Trade Agreements in the Presence of Renegotiation," *American Economic Journal: Microeconomics* 7(1): 109-43.
- Maggi, G. and R.W. Staiger (2015b) "Trade Disputes and Settlement," Dartmouth College manuscript, February.
- Nicita, A., M. Olarreaga, and P. Silva (2013) "Cooperation in WTO's Tariff Waters," *CEPR Discussion Papers No. 9529*.
- Olson, M. (1965) *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

- Palmeter, D. and P.C. Mavroidis (2004) *Dispute Settlement in the World Trade Organization: Practice and Procedure*. Cambridge, UK: Cambridge University Press.
- Park, J.H. (2011) “Enforcing International Trade Agreements with Imperfect Private Monitoring,” *Review of Economic Studies* 78(3): 1102-34.
- Rauch, J.E. (1999) “Networks versus markets in international trade,” *Journal of International Economics* 48(1): 7-35.
- Schwartz, W.F. and A.O. Sykes (2002) “The Economics Structure of Renegotiation and Dispute Resolution in the WTO/GATT System,” *Journal of Legal Studies* 31(1): S179-S204.
- Soderbery, A. (forthcoming) “Estimating Import Supply and Demand Elasticities: Analysis and Implications,” *Journal of International Economics*.
- Staiger, R.W. and A.O. Sykes (2013a) “How Important Can the Non-Violation Clause Be for the GATT/WTO?” *NBER Working Paper* No. 19256, July.
- Staiger, R.W. and A.O. Sykes (2013b) “Non-Violations,” *Journal of International Economic Law* 16(4): 741-775.
- Subramanian, A. and S.-J. Wei (2007) “The WTO Promotes Trade, Strongly But Unevenly,” *Journal of International Economics* 72(1): 151-175.

Figure 1: Timeline Mapping Policy Changes and the WTO Dispute Process to the Empirical Model

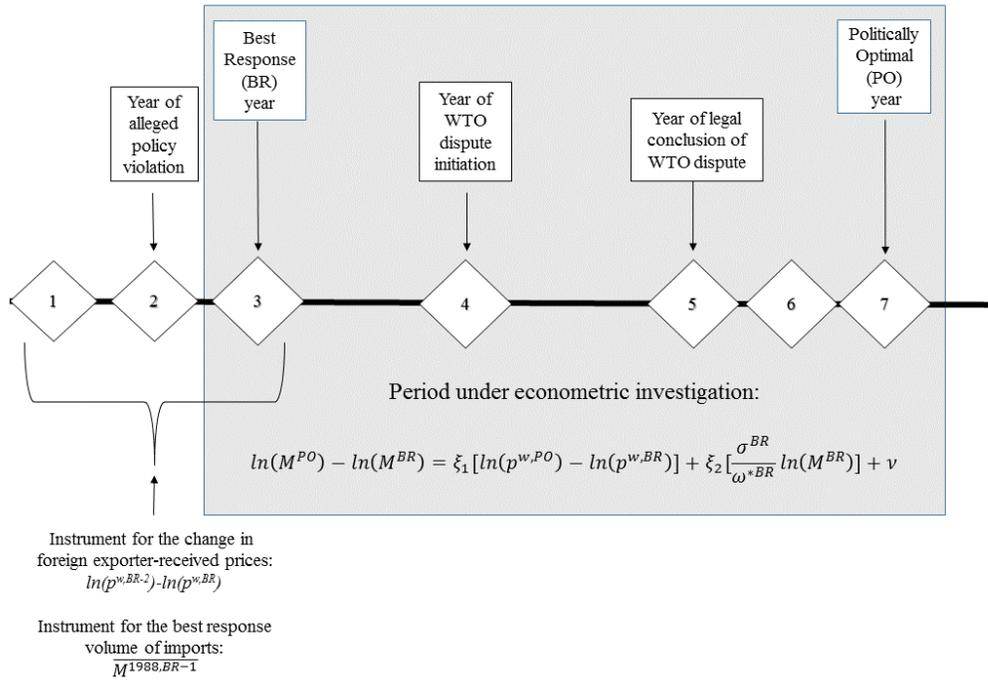


Figure 2: Average Import Volumes and Foreign-Exporter Received Prices for Products Subject to WTO Dispute, 1995-2009

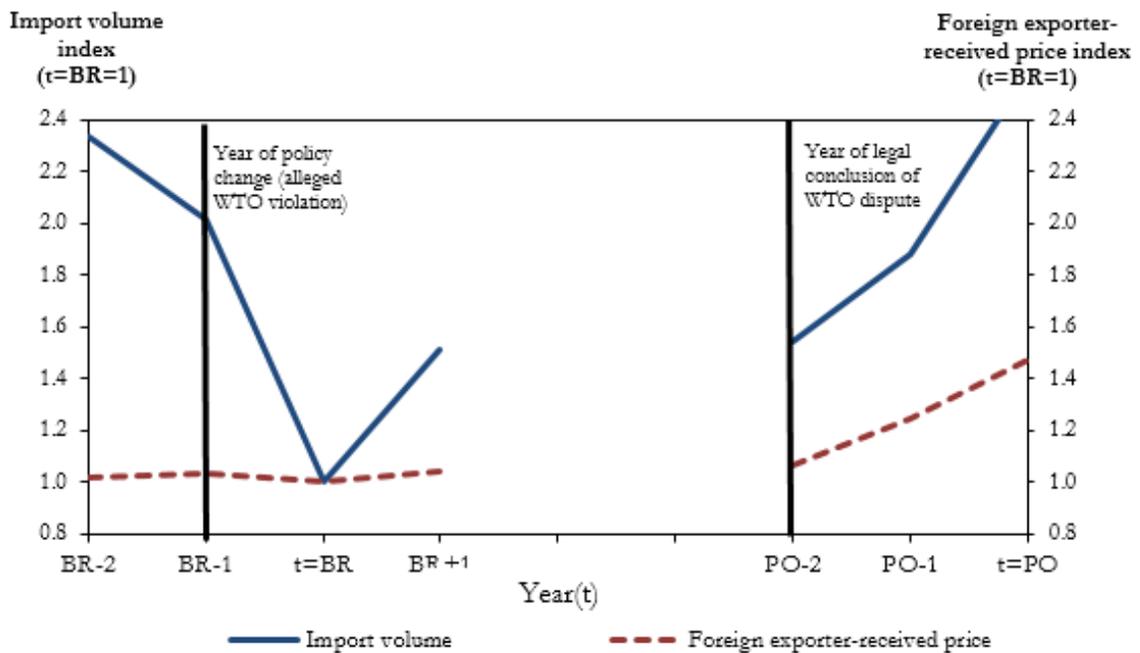


Table 1: WTO Disputes Initiated and Concluded, 1995-2009

	Number of disputes	Share of disputes in final sample
Total WTO disputes (bilateral pair, non-redundant definition)	347	
- <i>Disputes over policies that primarily affect exports</i>	35	
Disputes over policies that primarily affect imports	312	
- <i>Disputes over policies that primarily affect services imports</i>	7	
Disputes over policies that primarily affect goods imports	305	
- <i>Disputes over policies that affect general imports (no specific products listed)</i>	56	
Disputes over policies that target imported products	249	
- <i>Disputes in which we are unable to observe quantities and unit prices</i>	79	
- <i>Disputes in which we are unable to observe elasticities</i>	3	
Final Sample	167	100.0
- <i>Disputes in which the alleged WTO violation is of a policy change</i>	137	82.0
- <i>Disputes in which the alleged WTO violation is of no policy change</i>	30	17.9
- <i>Disputes in which the resolution results in an observed policy change</i>	139	83.2
- <i>Disputes in which the resolution results in no observed policy change</i>	28	16.7
- <i>Disputes which result in circulation of at least a Panel Report</i>	83	49.7
- <i>Disputes in which no legal ruling was issued</i>	84	50.2
- <i>Disputes over “global” policies that apply to all trading partners</i>	78	46.7
- <i>Disputes over “partial” policies in which some third country exporters are excluded from application</i>	89	53.3
- <i>Disputes in which the respondent (importer) is high income</i>	105	62.9
- <i>Disputes in which the respondent (importer) is developing</i>	62	37.1
- <i>Disputes in which the complainant (exporter) is high income</i>	97	58.1
- <i>Disputes in which the complainant (exporter) is developing</i>	70	41.9

Table 2: Summary Statistics

	Mean	Std. Dev.	Min	Max
Dependent variable:				
$\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$	0.12	1.24	-6.33	4.77
Explanatory variables:				
<i>Theoretical model determinants:</i>				
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	0.05	0.67	-4.44	3.73
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(M_{grc}^{BR})$	-31.83	107.25	-1152.83	51.64
<i>Other political-economic determinants:</i>				
$\ln(\text{Complainant's 'actionable' imports from respondent} / \text{Respondent's imports at stake in dispute})$	11.06	2.18	5.41	17.18
Global disputes	0.47	0.50	0.00	1.00
Global disputes x Exporter concentration	0.22	0.33	0.00	1.00
<i>Other disputed product, timing, or disputant determinants:</i>				
Respondent's real GDP growth between <i>BR</i> and <i>PO</i>	0.13	0.14	-0.20	0.74
Complainants's real GDP growth between <i>BR</i> and <i>PO</i>	0.14	0.10	-0.02	0.46
Dispute's <i>PO</i> year is during trade collapse of 2008-9	0.09	0.29	0.00	1.00
Dispute's <i>BR</i> year is during trade collapse of 2008-9	0.02	0.13	0.00	1.00
$\ln(\text{Respondent's real GDP (billions)})$	7.62	2.01	2.93	9.60
$\ln(\text{Complainant's real GDP (billions)})$	7.17	2.01	2.41	9.60
$\ln(\text{Respondent's real GDP per capita})$	9.58	1.16	6.47	10.72
$\ln(\text{Complainant's real GDP per capita})$	9.35	1.32	6.36	11.09

Note: 167 bilateral trade dispute observations; *g* indexes the disputes (products), *r* indexes respondent (importing) countries, *c* indexes complainant (exporting) countries, *BR* denotes the best response year (defined as one year after the violation sparking the dispute, and *PO* denotes the politically optimal year (two years after the legal conclusion of the dispute).

Table 3: Descriptive Statistics for Import Volume Growth Resulting from WTO Disputes, 1995-2009

	All disputes	Panel	No Panel	t- test	Global	Partial	t- test	High Income Respondent	Emerging Respondent	t- test	High Income Complainant	Emerging Complainant	t- test
Average import growth	0.12	0.12	0.12		0.12	0.12		0.22	-0.06		-0.01	0.30	
between <i>BR</i> and <i>PO</i> (median)	(0.18)	(0.23)	(0.11)		(0.24)	(0.08)		(0.18)	(0.07)		(0.05)	(0.26)	
Average import growth per annum	0.04	0.05	0.04		0.04	0.05		0.08	-0.01	*	0.03	0.07	
between <i>BR</i> and <i>PO</i> (median)	(0.05)	(0.05)	(0.04)		(0.05)	(0.04)		(0.05)	(0.03)		(0.02)	(0.07)	
Share of disputes with positive import growth between <i>BR</i> and <i>PO</i>	0.62	0.65	0.58		0.68	0.56		0.67	0.53	**	0.54	0.73	**
Share of disputes with import growth larger than RGDP growth	0.56	0.59	0.54		0.59	0.54		0.62	0.47	**	0.49	0.66	**
Average years between <i>BR</i> and <i>PO</i>	4.57	5.07	4.08	***	4.45	4.69		4.58	4.56		4.57	4.59	
Disputes	167	83	84		78	89		105	62		97	70	

Note: ***, **, * denote t-test statistically different from zero at the 1, 5, and 10 percent levels, respectively.

Table 4: Model Estimates of Import Volume Growth Resulting from WTO Disputes

Dependent variable is the growth in import volumes in products under dispute: $\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$								
	OLS (1)	IV: Instrument for Prices (2)	IV: Instrument for Volume (3)	IV: Full IV Estimates (4)	IV: Alternative Definition of BR Year (5)	IV: Baseline (6)	IV: Add Complainant/ Respondent FE (7)	IV: Add Sector FE (8)
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	-1.080*** (0.157)	-0.966*** (0.283)	-1.080*** (0.115)	-0.962*** (0.283)	-1.188*** (0.214)	-0.997*** (0.266)	-0.975*** (0.227)	-1.014*** (0.231)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{BR}} \ln(M_{grc}^{BR})$	0.002* (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.001* (0.001)	0.001* (0.001)	0.002*** (0.001)	0.000 (0.001)
$\ln(\text{Complainant's 'actionable' imports from respondent} / \text{Respondent's imports at stake in dispute})$						0.084** (0.039)	0.076* (0.040)	0.087* (0.046)
Global disputes						-0.376* (0.206)	-0.269 (0.214)	0.055 (0.250)
Global disputes x Exporter concentration						0.649** (0.328)	0.373 (0.338)	0.071 (0.364)
Respondent's real GDP growth between BR and PO						1.912*** (0.667)	0.561 (1.364)	0.654 (1.392)
Complainants's real GDP growth between BR and PO						-1.076 (0.903)	-0.710 (1.510)	-1.404 (1.576)
Dispute's PO year is during trade collapse of 2008-9						-0.241 (0.251)	-0.245 (0.233)	-0.163 (0.244)
Dispute's BR year is during trade collapse of 2008-9						-0.537 (0.562)	-0.877 (0.534)	-0.977 (0.624)
$\ln(\text{Respondent's real GDP})$						0.053 (0.061)	-0.708 (5.390)	0.664 (6.081)
$\ln(\text{Complainant's real GDP})$						-0.066 (0.048)	9.552*** (3.219)	9.476** (3.708)
$\ln(\text{Respondent's real GDP per capita})$						0.089 (0.108)	0.307 (6.346)	-2.194 (6.763)
$\ln(\text{Complainant's real GDP per capita})$						-0.065 (0.080)	-12.002*** (3.479)	-11.519*** (3.962)
Constant	0.230*** (0.079)	0.223*** (0.082)	0.226*** (0.080)	0.219*** (0.082)	0.187** (0.081)	-0.939 (0.951)	52.833* (30.857)	63.010** (31.374)
Observations	167	167	167	167	165	167	167	167
Root MSE	0.999	0.993	0.990	0.993	0.994	0.906	0.652	0.609
Cragg-Donald		16.18	6518	10.76	18.16	10.12	6.326	5.313
Sargan Test				0.934	0.423	0.279	0.506	0.755

Note: Standard errors in parentheses, ***, **, * denote statistically different from zero at the 1, 5, and 10 percent levels, respectively. Note that g indexes the disputes (products), r indexes respondent (importing) countries, c indexes complainant (exporting) countries, BR denotes the best response year (defined as one year after the violation sparking the dispute in all specifications except (5) in which it is defined as the year of the initiation of the dispute), and PO denotes the politically optimal year (two years after the legal conclusion of the dispute). Cragg-Donald is the F-statistic associated with the Cragg-Donald test for weak instruments and Sargan Test is the p-value associated with the Sargan test for overidentifying restrictions.

Table 5: Model Estimates of Import Volume Growth, by Institutional Features

Dependent variable is the growth in import volumes in products under dispute: $\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$									
	IV: Baseline (1)	IV: Drop disputes with no observed change in policy or legal ruling (2)	IV: Drop US AD/CVD (3)	IV: Global only (4)	IV: Partial only (5)	IV: Panel Report only (6)	IV: No Legal Decision Only (7)	IV: High Retaliation Capacity (8)	IV: Low Retaliation Capacity (9)
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	-0.997*** (0.266)	-1.040*** (0.280)	-0.996*** (0.274)	-0.749*** (0.285)	-1.269*** (0.474)	-1.207*** (0.268)	0.368 (1.188)	-0.705*** (0.228)	-4.687 (4.662)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{BR}} \ln(M_{grc}^{BR})$	0.001* (0.001)	0.001* (0.001)	0.001* (0.001)	0.002** (0.001)	-0.001 (0.002)	0.002*** (0.001)	0.000 (0.003)	0.002*** (0.001)	0.001 (0.005)
$\ln(\text{Complainant's 'actionable' imports from respondent / Respondent's imports at stake in dispute})$	0.084** (0.039)	0.094** (0.046)	0.080* (0.041)	0.100** (0.045)	0.102 (0.064)	0.116** (0.059)	0.040 (0.072)	0.091 (0.083)	-0.143 (0.314)
Global disputes	-0.376* (0.206)	-0.309 (0.242)	-0.379* (0.213)			-0.533* (0.316)	-0.330 (0.387)	-0.179 (0.273)	0.098 (1.401)
Global disputes x Exporter concentration	0.649** (0.328)	0.614 (0.380)	0.667** (0.335)	0.558* (0.287)		0.964** (0.466)	0.913 (0.746)	1.158*** (0.435)	-0.167 (1.830)
Constant	-0.939 (0.951)	-1.185 (1.113)	-0.814 (1.026)	-1.550 (1.272)	-2.894 (1.796)	-0.534 (1.475)	0.801 (2.423)	0.467 (1.707)	-0.075 (3.904)
Additional control variables of Table 4, specification (6)	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	167	139	147	78	89	83	84	84	83
Root MSE	0.906	0.968	0.915	0.763	0.948	0.909	1.134	0.835	2.503
Sargan Test	0.279	0.444	0.0932	0.177	0.462	0.380	0.514	0.387	0.922
Cragg-Donald	10.12	15.94	9.259	11.96	2.461	12.18	0.711	20.80	0.184

Note: Standard errors in parentheses, ***, **, * denote statistically different from zero at the 1, 5, and 10 percent levels, respectively. Note that g indexes the disputes (products), r indexes respondent (importing) countries, c indexes complainant (exporting) countries, BR denotes the best response year (defined as one year after the violation sparking the dispute), and PO denotes the politically optimal year (two years after the legal conclusion of the dispute). Cragg-Donald is the F-statistic associated with the Cragg-Donald test for weak instruments and Sargan Test is the p-value associated with the Sargan test for overidentifying restrictions. Coefficient estimates of the additional control variables (not reported in the Table) are available upon request.

Table 6: Model Estimates of Import Volume Growth, by Level of Economic Development and Trade Outcomes

Dependent variable is the growth in import volumes in products under dispute: $\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$							
	IV: Baseline (1)	IV: High Income Respondent (2)	IV: Emerging Respondent (3)	IV: High Income Complainant (4)	IV: Emerging Complainant (5)	IV: High Growth Disputes (6)	IV: Small Growth Disputes (7)
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	-0.997*** (0.266)	-0.796*** (0.291)	-1.641*** (0.345)	-0.912*** (0.206)	-5.194 (7.540)	-0.668*** (0.219)	-1.589*** (0.401)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(M_{grc}^{BR})$	0.001* (0.001)	0.002*** (0.001)	-0.002 (0.002)	0.002*** (0.001)	-0.004 (0.005)	-0.001 (0.001)	0.002** (0.001)
$\ln(\text{Complainant's 'actionable' imports from respondent} / \text{Respondent's imports at stake in dispute})$	0.084** (0.039)	0.157*** (0.044)	-0.028 (0.069)	0.051 (0.044)	0.028 (0.284)	0.065* (0.034)	0.042 (0.058)
Global disputes	-0.376* (0.206)	-0.373 (0.228)	0.026 (0.419)	-0.021 (0.246)	-1.090 (0.946)	-0.417** (0.186)	-0.135 (0.278)
Global disputes x Exporter concentration	0.649** (0.328)	0.546 (0.360)	0.097 (0.637)	0.639* (0.364)	0.169 (1.091)	0.048 (0.284)	0.424 (0.470)
Constant	-0.939 (0.951)	-2.997 (2.987)	-1.612 (2.234)	2.525 (2.830)	-13.211 (13.729)	-0.062 (1.006)	-1.408 (1.227)
Additional control variables of Table 4, specification (6)	Y	Y	Y	Y	Y	Y	Y
Observations	167	105	62	97	70	84	83
Root MSE	0.906	0.794	0.909	0.790	1.581	0.543	0.865
Sargan Test	0.279	0.668	0.0437	0.268	0.739	0.00528	0.0774
Cragg-Donald	10.12	6.714	5.906	9.398	0.118	4.977	5.254

Note: Standard errors in parentheses, ***, **, * denote statistically different from zero at the 1, 5, and 10 percent levels, respectively. Note that g indexes the disputes (products), r indexes respondent (importing) countries, c indexes complainant (exporting) countries, BR denotes the best response year (defined as one year after the violation sparking the dispute), and PO denotes the politically optimal year (two years after the legal conclusion of the dispute). Cragg-Donald is the F-statistic associated with the Cragg-Donald test for weak instruments and Sargan Test is the p-value associated with the Sargan test for overidentifying restrictions. Coefficient estimates of the additional control variables (not reported in the Table) are available upon request.

Table Appendix A: First Stage Estimates: $\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$

	OLS (1)	IV: Instrument for Prices (2)	IV: Instrument for Volume (3)	IV: Full IV Estimates (4)	IV: Alternative Definition of BR Year (5)	IV: Baseline (6)	IV: Add Complainant/ Respondent FE (7)	IV: Add Sector FE (8)
$\ln(p_{grc}^{w,BR-2}) - \ln(p_{grc}^{w,BR})$		0.693*** (0.128)		0.692*** (0.128)	0.692*** (0.128)	0.705*** (0.130)	0.691*** (0.168)	0.722*** (0.180)
$\ln(\text{Real price of crude oil}^{PO}) - \ln(\text{Real price of crude oil}^{BR}) \times$ $\frac{\log(\text{distance})}{\omega_{gr}^{BR}} (\ln(M_{grc}^{1988,BR-1}))$		0.030** (0.014)		0.030** (0.014)	0.030** (0.014)	0.027 (0.017)	0.045* (0.024)	0.025 (0.028)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{BR}} \ln(M_{grc}^{BR})$		-0.000 (0.000)		-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.001)	0.000 (0.001)
$\ln(\text{Complainant's 'actionable' imports from respondent /}$ $\text{Respondent's imports at stake in dispute})$						0.016 (0.026)	0.039 (0.040)	0.045 (0.049)
Global disputes						0.034 (0.140)	0.167 (0.216)	0.112 (0.273)
Global disputes x Exporter concentration						-0.306 (0.215)	-0.589* (0.318)	-0.307 (0.392)
Respondent's real GDP growth between BR and PO						0.225 (0.455)	0.992 (1.355)	-0.320 (1.516)
Complainants's real GDP growth between BR and PO						0.024 (0.659)	-1.032 (1.570)	0.558 (1.751)
Dispute's PO year is during trade collapse of 2008-9						0.066 (0.171)	0.111 (0.233)	0.214 (0.252)
Dispute's BR year is during trade collapse of 2008-9						0.614 (0.371)	0.809 (0.524)	1.055* (0.609)
$\ln(\text{Respondent's real GDP})$						0.023 (0.041)	7.752 (4.759)	10.265* (5.386)
$\ln(\text{Complainant's real GDP})$						-0.008 (0.033)	0.954 (3.252)	-0.947 (4.031)
$\ln(\text{Respondent's real GDP per capita})$						-0.030 (0.073)	-10.350* (5.502)	-11.058* (5.986)
$\ln(\text{Complainant's real GDP per capita})$						0.071 (0.053)	-1.150 (3.517)	-0.540 (4.328)
Constant		-0.030 (0.068)		-0.028 (0.069)	-0.028 (0.069)	-0.679 (0.650)		
Observations		167		167	167	167	167	167
Root MSE		0.617		0.617	0.617	0.617	0.659	0.659

Note: Coefficient estimates associate with the first stage results from the instrumental variable regressions associated with Table 4. Standard errors in parentheses, ***, **, * denote statistically different from zero at the 1, 5, and 10 percent levels, respectively.

Table Appendix B: First Stage Results: $\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(M_{grc}^{BR})$

	OLS (1)	IV: Instrument for Prices (2)	IV: Instrument for Volume (3)	IV: Full IV Estimates (4)	IV: Alternative Definition of BR Year (5)	IV: Baseline (6)	IV: Add Complainant/ Respondent FE (7)	IV: Add Sector FE (8)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(M_{grc}^{1988, BR-1})$			0.916*** (0.011)	0.917*** (0.011)	0.917*** (0.011)	0.926*** (0.011)	0.944*** (0.012)	0.843*** (0.017)
$\ln(p_{grc}^{w, PO}) - \ln(p_{grc}^{w, BR})$			-0.211 (1.959)					
$\ln(p_{grc}^{w, BR-2}) - \ln(p_{grc}^{w, BR})$				5.945* (3.466)	5.945* (3.466)	4.618 (3.227)	5.722 (3.742)	6.073* (3.103)
$\ln(\text{Real price of crude oil}^{PO}) - \ln(\text{Real price of crude oil}^{BR}) \times$ $\log(\text{distance})$				-0.478 (0.370)	-0.478 (0.370)	-0.746* (0.412)	0.036 (0.545)	0.493 (0.491)
$\ln(\text{Complainant's 'actionable' imports from respondent /}$ $\text{Respondent's imports at stake in dispute})$ Global disputes						-1.993*** (0.656)	-2.737*** (0.898)	-2.035** (0.848)
Global disputes x Exporter concentration						-3.020 (3.482)	-11.345** (4.819)	-5.170 (4.720)
Respondent's real GDP growth between BR and PO						14.878*** (5.328)	14.977** (7.089)	8.252 (6.778)
Complainants's real GDP growth between BR and PO						-9.353 (11.275)	27.090 (30.194)	4.526 (26.193)
Dispute's PO year is during trade collapse of 2008-9						25.781 (16.343)	-24.550 (34.976)	6.934 (30.255)
Dispute's BR year is during trade collapse of 2008-9						-0.979 (4.232)	-2.236 (5.190)	2.555 (4.360)
$\ln(\text{Respondent's real GDP})$						-9.479 (9.207)	-4.681 (11.687)	7.151 (10.520)
$\ln(\text{Complainant's real GDP})$						2.631** (1.029)	-61.713 (106.047)	7.873 (93.077)
$\ln(\text{Respondent's real GDP per capita})$						2.399*** (0.819)	8.336 (72.475)	13.334 (69.673)
$\ln(\text{Complainant's real GDP per capita})$						-5.645*** (1.818)	74.597 (122.604)	-39.107 (103.453)
Constant			3.785*** (1.385)	5.566*** (1.859)	5.566*** (1.859)	39.433** (16.124)	-5.294 (78.365)	-34.771 (74.791)
Observations			167	167	167	167	167	167
Root MSE			16.90	16.70	16.70	15.29	14.69	11.38

Note: Coefficient estimates associate with the first stage results from the instrumental variable regressions associated with Table 4. Standard errors in parentheses, ***, **, * denote statistically different from zero at the 1, 5, and 10 percent levels, respectively.