

Trade Agreements and Enforcement

Evidence from WTO Dispute Settlement

Chad P. Bown
Kara M. Reynolds



WORLD BANK GROUP

Development Research Group

Trade and International Integration Team

April 2015

Abstract

This paper examines the 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, the paper models 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 to trading partners. The approach is to extend earlier theoretical models of trade agreement accession negotiations to the setting of enforcement negotiations in order to guide the empirical assessment. The paper uses instrumental variables to

estimate the model on trade volume outcomes from World Trade Organization (WTO) disputes over 1995–2009. The evidence is consistent with theoretical predictions that larger import volume outcomes are associated with products that have smaller increases in foreign exporter-received prices (terms-of-trade losses for 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.

This paper is a product of the Trade and International Integration Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at cdown@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

TRADE AGREEMENTS AND ENFORCEMENT: EVIDENCE FROM WTO DISPUTE SETTLEMENT*

Chad P. Bown
Peterson Institute for International
Economics & CEPR

Kara M. Reynolds
American University

JEL: F13, F14

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

* Bown: Peterson Institute for International Economics, 1750 Massachusetts Avenue NW, Washington, DC 20036 USA; tel: +1.202.454.1306, email: cbrown@piie.com.

Reynolds: Department of Economics, American University; 4400 Massachusetts Avenue NW, Washington, DC 20016 USA; Tel: +1.202.885.3768, fax:+1.202.885.3790, email: reynolds@american.edu.

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 Mkrtchyan, participants at seminars at Stanford, Vanderbilt, Illinois, Wesleyan, the World Bank, the Geneva Trade and Development Workshop, the PEIO Conference in Berlin, and two anonymous referees. 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. Bown thanks the World Bank's Development Research Group for its hospitality during the period in which much of the work on this paper was completed. 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 policy making behavior so that subsequent policy outcomes resulting from dispute settlement negotiations are determined by something else?

The purpose of this paper is to 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 examining the subsequent dispute resolution process is motivated by the combination of two insights from the existing literature on trade agreements and dispute settlement that directly speak to the design of the institution established to resolve WTO disputes.

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) similarly 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. 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 agreement levels of import protection and find evidence consistent with the theory that market power affects unilaterally-imposed (noncooperative) tariffs. Bagwell and Staiger (2011) assess 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).¹

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 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. 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 1947-1994.² Moreover, in their survey of the bilateral trade flows of the products that member

¹See also Nicita, Olarreaga, and Silva (forthcoming). A more general survey of the economics literature on trade agreements is Maggi (2014).

²For 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.

countries eventually subject to WTO dispute settlement, Bown and Reynolds (2015) 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 2011.³

In this paper, we investigate empirically the relevance of the terms-of-trade theory for enforcement of international agreements and the negotiated outcomes that arise under dispute settlement. 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 arise 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.⁴

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*.⁵ 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.

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

³These 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 (2016, Section 7), which surveys the economics literature on WTO dispute settlement.

⁴In 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.

⁵Our 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.

across the litigating countries that the previous literature has suggested are also likely to affect dispute settlement outcomes (Bown, 2004a).⁶

We find the empirical results are particularly strong for disputes that are not settled, but that conclude with formal legal decisions. There is also strong evidence that the model helps to explain the trade outcomes for disputes involving high-income countries. This evidence is reassuring as these particular subsamples of disputes are the most aligned with our theoretical framework. Finally, we note 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.

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.

The basic GATT/WTO principle of reciprocity arises in the legal 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 impli-

⁶Bown (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.

cations 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.⁷

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. 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.⁸ 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

⁷In the next section below we formally define the “best response” and “politically optimal” levels of import tariffs in the context of an economic model.

⁸Furthermore, 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.

subject if they refused to comply with WTO rulings.⁹

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 policy 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.¹⁰

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 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.

⁹See 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.

¹⁰Put 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.

Following Bagwell and Staiger (2011), the first 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

$$W_p(p^{PO}, p^{w,PO}) = 0 \quad (2)$$

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

In contrast, 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, including terms-of-trade impacts:

$$W_p \frac{dp}{d\tau} + W_{p^w} \frac{\partial p^w}{\partial \tau} = 0 \quad (3)$$

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:

$$-\frac{W_p}{p^{w,BR}} = \eta^{BR} \quad (4)$$

where $\eta^{BR} \equiv \frac{\sigma^{BR}}{\omega^{*BR}} \frac{M^{BR}}{p^{BR}}$, 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.

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 coordinated 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}).

In particular, Bagwell and Staiger (2011) explains how the term η^{BR} measures the "international cost-shifting" motives embodied in the best response tariff. This term can be thought of as the degree to which the best response tariff would have to be reduced in order to achieve the globally welfare enhancing level of protection, τ^{PO} . Notice that for a given best response volume of imports M^{BR} , this cost-shifting motive is increasing in the elasticity of import demand, so that a given tariff increase generates a larger decrease in the quantity of imports demanded. Second, for a given best response volume of imports M^{BR} , cost shifting motives are *decreasing* in the foreign export supply elasticity. I.e., the more market power that the importing country has vis-a-vis the exporter,

the larger will be the cost-shifting motive embodied in the best response tariff. Put differently, the small importing country case corresponds to $\omega^{*BR} \rightarrow \infty$, in which the cost-shifting motive goes to zero, regardless of M^{BR} or σ^{BR} .

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.¹¹ This country will then face a dispute and we seek to examine determinants of its policy decision to return to the politically optimal tariff.¹² 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.¹³ 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:

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

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

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

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

¹¹We 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 (2013, forthcoming).

¹²That 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.

¹³Put 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.

Finally, imports are defined by $M(p) = D(p) - S(p)$, 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

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

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 $\theta \equiv (-\partial M/\partial p)/(\partial E^*/\partial p^w)$.¹⁴ 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 the determinants of the change in import volumes after the dispute, from best-response to the politically-optimal level, yields

$$M^{PO} - M^{BR} = \psi_1(p^{w, PO} - p^{w, BR}) + \psi_2 M^{BR} \quad (10)$$

where $\psi_1 \equiv -(\delta + \kappa)^2/[\delta + \kappa - (\gamma - 1)\kappa] < 0$, and $\psi_2 \equiv [\theta(\delta + \kappa)]/[\delta + \kappa - (\gamma - 1)\kappa] > 0$, again because $\theta > 0$ and using $[\delta + \kappa - (\gamma - 1)\kappa] > 0$ by equation (A1).

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

Thus far, we have worked under the assumption that the underlying supply and demand models

¹⁴In Bagwell and Staiger (2011), equation (9) is given by equation (12) on p. 1248.

in the theory were linear. This implied that $\theta \equiv (-\partial M/\partial p)/(\partial E^*/\partial p^w)$ was constant and therefore embedded as part of the coefficient ψ_2 in equation (10). Suppose we consider nearby nonlinear models of supply and demand that are continuous with respect to the underlying model parameters and that would yield a θ that is not constant.

In such a case, we can manipulate equation (10) in order to consider an equation of the following form

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

Note that the last term in this equation is equivalent to $[\eta^{BR} p^{w,BR}]$, the interaction between the international cost-shifting motive and the world price received by the foreign exporter.¹⁵ It is straightforward to show that $\xi_1 < 0$ and $\xi_2 > 0$ in this model.¹⁶

Our interpretation of this estimation equation (11) 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 cost shifting motives as captured by $\frac{\sigma^{BR}}{\omega^{*BR}} \frac{M^{BR}}{p^{BR}}$ and the world price received by the foreign exporter in the best response year, $p^{w,BR}$. Intuitively, for a given world price in the best response year, the increase in the volume of imports following resolution of a dispute will be larger the greater the cost shifting motives associated with the initial violation. Put another way, absent any market-based or political shocks, the change in world prices between the politically optimal and best response year should approach zero as the cost shifting motives approach zero, resulting in no increase in the volume of trade from resolution of the dispute.

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

¹⁵This ends up being consistent with the alternative estimation equation (13) in Bagwell and Staiger (2011, p. 1248), that also introduces information from import demand and foreign export supply elasticities as a robustness check to their baseline estimation equation. In our approach, although ξ_2 will be greater than zero for a large class of non-linear demand and supply curves, the magnitudes of our coefficient estimates may not apply to observations that fall far from the sample of data from which we draw our elasticity estimates.

¹⁶Consider, for example, starting from the parameter $\psi_2 \equiv [\theta(\delta + \kappa)]/[\delta + \kappa - (\gamma - 1)\kappa]$ in equation (10). Using the definition $\theta \equiv (-\partial M/\partial p)/(\partial E^*/\partial p^w)$, one can rewrite this as $\psi_2 \equiv \frac{\delta + \kappa}{[\delta + \kappa - (\gamma - 1)\kappa]} \frac{\sigma^{BR}}{\omega^{*BR}} \frac{p^{w,BR}}{p^{BR}}$, where the term $[\delta + \kappa - (\gamma - 1)\kappa]$ must again be positive by equation (A1).

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.¹⁷ While our empirical approach is designed to investigate whether the determinants of trade dispute outcomes are consistent with the core relationships deriving from the terms-of-trade theory and 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.¹⁸

4 Data and Estimation

Although the variables included in equation (11) are technically measured in the same units across all observations (billions of kilograms and dollars per kilogram), in practice the heterogeneity in the types of products under dispute makes it difficult to compare the changes in the levels of volumes and prices across our observations. For example, consider two disputes that each result in an increase in the volume of trade of 0.3 billion kilograms upon resolution. The market impact of a 0.3 billion kilogram increase in wheat will likely be quite different than an increase of the same number of kilograms of steel. To better standardize measurement across observations, we instead rely on theoretically motivated equation (11) to ultimately estimate a model of the form:

$$\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR}) = \xi_1[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})] + \xi_2\left[\frac{\sigma_{gr}^{BR}}{\omega_{gc}^{*BR}} \ln\left(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M_{grc}^{BR}\right)\right] + \nu_{grc}, \quad (12)$$

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 $\xi_2 > 0$.

Because we estimate the model in the percent change in our variables of interest rather than levels, the coefficients that result from this estimation cannot be interpreted as the structural parameters of the model described earlier. Nevertheless, the signs and magnitudes of our coefficients can still be interpreted as evidence of the degree to which terms-of-trade motives can explain the

¹⁷While 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.

¹⁸To our knowledge, Maggi and Staiger (forthcoming) 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, 2016), Maggi and Staiger (2011), Staiger and Sykes (forthcoming) and Park (2011).

outcomes of WTO disputes.

Like other studies of WTO dispute settlement, our coefficient estimates could potentially be impacted by selection bias. Intuitively, disputes arise not at random but rather are chosen strategically by the complainant country, and to the degree that the unobserved characteristics that determine this selection are correlated with the error in our equation, our coefficients could be biased. Some of these concerns are addressed by our focus on a theoretically-motivated model and our inclusion of additional control variables likely associated with the selection process; these include the economic size of the complainant and respondent countries. Nevertheless, unaddressed selection concerns may remain.

In the next sub-section we introduce the data and construction of variables used in the estimation. We then discuss endogeneity issues associated with estimating equation (12) and our particular instrumental variables approach to address these concerns.

4.1 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 2009.¹⁹ We start with 1995 as that was the first year that the WTO and its formal 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 (2015).

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. In most instances, the year of imposition of the best response policy is available from either official WTO dispute documentation or from other official government notifications. However, 17 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 under the terms of its accession. In these cases, we define $t = BR$ simply to be the year of the

¹⁹Overall, 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.

initiation of the formal WTO dispute.²⁰

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), or revelations through other official government documents (e.g., removal of temporary trade barriers), 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.²¹ In 84 percent of our sample, the timing of post-dispute policy change implemented by the respondent is directly observable. However, in 16 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.

Table 1 provides greater detail about the population of disputes that WTO members initiated and concluded between 1995 and 2009, and how it has been reduced to yield the sample used in our estimation exercise. 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.²² 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

²⁰Bown and Reynolds (2015) provide an analysis of the trade flow data associated with products prior to the initiation of the dispute. One result of interest 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.

²¹Article 21.3(c) of the WTO’s Dispute Settlement Understanding specifically makes the following suggestion: “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, time may be shorter or longer, depending upon the particular circumstances.”

²²For 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.

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 152 respondent/complainant pairs, or roughly 61 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 investigate whether the model explains the trade flow outcomes of certain subsamples of disputes better than others, including differences based on legal resolution of the process, the type of policy being challenged, or the countries involved. Table 1 provides information on the number of disputes in these subsamples. For example, the determinants of trade disputes formally resolved through the WTO’s legal system may be different from determinants of disputes that are settled or withdrawn prior to any formal WTO legal rulings. A WTO Panel Report, or the first formal legal ruling of importance in any dispute, was issued in slightly more than half (78 out of 152) of the disputes in our sample.

Next, we 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 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.²³ Table 1 indicates that in our final sample of 152 disputes, slightly less than half are associated with challenges to “global” policies, and the rest are associated with “partial” policies.

Finally, we explore whether our model better fits the high-income respondent subsample of data relative to the low-income importing countries. As Table 1 indicates, our sample contains a significant number of disputes involving both high-income and developing countries as both complainants and respondents.²⁴

4.2 Variable construction and data

Estimation of equation (12) 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, 2015) which matches disputes to the c.i.f. value and volume (as measured by the net

²³To clarify, our categorization as to whether the disputed policy is “global” versus “partial” is based on our judgment of whether its application 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 reach 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).

²⁴Note, 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.

weight in kilograms) of bilateral import data by HS-06 product code from UN Comtrade.²⁵

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 particular dispute.²⁶

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 disputed products’ HS-06 codes. Of the remaining disputes, import volume data is available for almost all HS-06 codes. For the handful of disputes without volume data for all HS-06 codes, 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. This conservative approach forces us to drop 79 disputes from the population of WTO cases because of unobserved import volumes.²⁷

Estimation of equation (12) also requires data on the *domestic* price in the respondent country (p_{grc}^{BR}). Because we do not observe domestic prices directly, we instead proxy for these prices using the average f.o.b. (free on board) export unit value associated with the products under dispute from the respondent country to the world.²⁸ Intuitively, the border price of a product being exported should be identical to the domestic price under the law of one price, which states that identical products sold in different countries should be the same price when expressed in the same currency. This has been used to justify substituting export prices for domestic prices in a number of different contexts. For example, the Canadian government’s export price index was constructed

²⁵Disputes 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.

²⁶While freight and insurance rates are 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 in our discussion of instrument construction.

²⁷Such 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.

²⁸Similar to the calculation of world prices, we calculate this variable using the ratio of the real value of exports from the respondent country to the world to the netweight (kilograms) of these exports, deflating the nominal export value data using the IMF’s world import price index. To avoid any potential bias in this value tied to the dispute itself, we exclude from this calculation the respondent’s exports to the complainant country(s) in the dispute.

using domestic prices as a proxy for export prices for over 30 years.²⁹

Nevertheless, we acknowledge that export prices may deviate from domestic prices for a number of reasons. For example, domestic producers may use different pricing strategies in foreign markets or absorb some of the fluctuations of exchange rates into their mark-ups. Perhaps of greater concern is evidence in research such as Iacovone and Javorcik (2009) that the unit values of goods that are exported are higher than those products sold only domestically, suggesting that exported products are on average higher quality. Should our proxy variable for domestic prices be biased upward, our estimate of the coefficient ξ_2 will also be biased upward. Nevertheless, we have checked that specifications that omit the price wedge term ($\frac{p^{w,BR}}{p^{BR}}$) from equation (11) produce coefficients of similar magnitude to those reported in Tables 4 and 5 below.

Because export data from the respondent country is not available for all HS-06 products, we use the same procedure to calculate our proxy for domestic prices. This forces us to drop an additional 15 disputes in which the respondent’s export volume data is not available for at least 80 percent of the HS-06 codes.³⁰

Finally, the elasticity data used to estimate equation (12) 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 codes 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 (2015). Dispute-level estimates of elasticities are calculated using a trade value weighted average of the HS-06 product elasticities associated with each dispute. The Crowley-Yu elasticities are unavailable for the HS-06 product 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.³¹

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

4.3 Instrumental variables estimation

Ordinary least squares estimates of our parameters will be biased by our exclusion of unobserved shocks to both the export supply and import demand curve of the products under investigation. For example, market-based shocks that occur between the politically optimal and best response year would be correlated with both the change in the world price ($\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$) and volume of imports ($\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$). To correct for this bias, we propose to instrument for our key variables of interest, $\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$ and $\ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M_{grc}^{BR})$, using variables that are correlated with these key variables but uncorrelated with the error term (ν_{grc}).

²⁹Statistics Canada, “New International Canadian Price Index: A Brief Overview,” accessed at <http://unstats.un.org/unsd/tradekb/Knowledgebase/New-Canadian-International-Trade-Price-Index-Brief-Overview>.

³⁰We fail to reject the hypothesis that the means of our key variables of interest in those observations that we drop are identical to those in the final sample, suggesting that this procedure should not introduce unnecessary sample selection bias into our results.

³¹Developing 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.

We ultimately employ three instruments in our analysis. The first is the change in exporter prices for the disputed products due to 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 include the instrument $(\ln(p_{grc}^{w,BR-2}) - \ln(p_{grc}^{w,BR}))$, which should be correlated with the post resolution price change, $(\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR}))$, but unrelated to current market shocks.³²

Second, we include the interaction between the average price wedge $(\frac{\overline{p_{grc}^w}}{p_{grc}})$ and the average bilateral import volume $(\overline{M_{grc}})$ between 1992 and the year prior to the best response year. This variable will be correlated with the volume of trade and price wedge in the best response year. Furthermore, using the historical average prior to the violation fully going into effect should eliminate the correlation between this instrument and unobservables impacting the trade volume postdating resolution of the dispute that are necessarily embedded in our error term.

Third, we follow Khandelwal (2010) by including 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 earlier.

In general, valid instruments must satisfy two requirements. First, they must be highly correlated with the endogenous regressors (even after controlling for the other exogenous regressors). On this score, our statistical tests indicate a valid set of instruments. First stage results reported in Appendix A and B find that both the change in price associated with the original violation and the change in crude oil prices are positive and statistically significant determinants of the change in price associated with resolution of the dispute, while the interaction of the historical average price wedge and average volume is a highly significant, positive determinant of the interaction of these terms in the best response year. The F-statistics associated with the joint tests of whether the excluded instruments are significant determinants of the endogenous variables, reported in Tables 4 and 5, exceed standard thresholds for instrument validity.

More challenging is to satisfy the requirement that the instruments are uncorrelated with the error term. To the extent that the error is associated with shocks to import demand and export supply of the products under dispute between the politically optimal and best response year, the lagged and historical average values of our endogenous regressors taken from prior to this period will be uncorrelated with the error. We acknowledge, however, that to the degree that our sample suffers from selection bias our instruments may still be correlated with the error term. Intuitively, if the selection mechanism is embodied in our error term, the lagged and average variables that potentially govern this selection may not satisfy the exclusion restriction. However, we are reassured by the fact that the Sargan-Hansen J tests for overidentifying restrictions, reported in Tables 4 and

³²For the 17 percent of disputes in which no respondent policy change took place (in which the dispute was triggered by a failure to change policy), foreign-exporter received prices may not have changed prior to the dispute, thus there may not be enough variation in this term for a strong instrument. We therefore discuss the robustness of our baseline results to dropping this subset of observations below.

5, confirm our a priori belief that the instruments are uncorrelated with the error term.

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 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 11 percent and a median of 18 percent. On average, 4.36 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-three percent of all disputes result in positive import growth between $t = BR$ and $t = PO$; but in only 58 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. First, disputes that proceed to a formal legal ruling (through a Panel Report) are no more likely to result in positive import growth or higher levels of import growth than disputes that settle earlier or are withdrawn before any formal legal ruling. This result could be viewed as surprising for any theory predicting that disputes that settle early have “better” economic outcomes.³³ There is also little statistical difference in the trade growth outcomes for disputes that confront “global” (as opposed to “partial”) policies despite the fact that 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.

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 coun-

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.

tries. 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, 45 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 28 percent over the life of the dispute (6 percent per annum). Overall, these potentially unexpected patterns arising from the raw data on trade dispute outcomes motivate the need for a formal model to estimate the impact of theoretical determinants on trade flow outcomes.

5 Econometric Results

5.1 Baseline results for the theoretically-motivated determinants

Column (1) of Table 4 provides our baseline estimates of the determinants of the growth of bilateral import volumes of the disputed product from the complainant to the respondent, $[\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})]$ as defined by equation (12). 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 illustrated in Figure 1. 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 $t = PO$. As discussed in Section 4.3, the Cragg-Donald Wald F-statistics are well over 10 for most of our specifications, and the p-values associated with the Sargan test of overidentifying restrictions confirm that the excluded variables are uncorrelated with our error term suggesting that we have employed a valid set of instruments.³⁴

The coefficient estimates on our key variables of interest in the baseline specification are of the theoretically predicted sign and statistically different from zero.

First, the estimate of -1.017 on $[\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})]$ indicates that disputed-product import volume growth is 1.017 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 that results from the dispute.

Second, the coefficient estimate on $[\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M_{grc}^{BR})]$ of 0.001 implies that a one standard deviation increase in this variable will result in a 13 percentage point increase in disputed product import volume growth; i.e., from the baseline prediction (at the means of the data) of 11 percent to 24 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 sizable impact, leading to a nearly 55 percentage point increase above the mean prediction for import volume growth.

³⁴Appendix Tables A and B provide a complete reporting of the first stage coefficient estimates associated with the model specifications presented in Table 4. Results from specifications that do not control for the potential endogeneity of our key variables of interest are reported in Columns 1-3 of Appendix Table C. Although the magnitude of the coefficient on $[\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M_{grc}^{BR})]$ is virtually identical to those estimated using instrumental variables, failure to control for the endogeneity of $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$ results in a small upward bias in the ξ_1 coefficient.

The estimates on these two key coefficients are robust to some, but not all, alternative specifications. For example, the inclusion of country-level (complainant and respondent) fixed effects in column (2) of Table 4 does not change the qualitative nature of the estimates of the theoretically-motivated determinants of interest. Our results are also robust to a number of other modifications to the sample and variable definitions, which are reported in Appendix C, including changing the definition of the best response year to the initiation of the dispute (column 5), dropping disputes with unobserved post-dispute policy changes (column 6), and dropping all disputes associated with US antidumping use (column 7).³⁵

However, while leaving unaffected the estimate on $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$, inclusion of sector-level fixed effects in column (3) of Table 4 does eliminate the statistical significance on the estimated impact of $[\frac{\sigma_{gr}^{BR}}{\omega_{gc}^{*BR}} \ln(\frac{p_{grc}^{w, BR}}{p_{grc}^{BR}} M_{grc}^{BR})]$.³⁶ Nevertheless, our broad findings are 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.

5.2 The effect of other determinants on disputed product import growth

Our baseline specification also includes a set of covariates that the prior literature suggests would impact the trade volume growth stemming from negotiations under WTO dispute settlement. For ease of exposition, we group these potential determinants into two broad categories: (i) other political-economic determinants 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 variable we introduce 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).³⁷

³⁵This final sensitivity test 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.

³⁶Following the approach suggested by Broda, Limão, and Weinstein (2008) to check for concern over potential measurement error and outliers, the impact of the estimated elasticities ratio variable is sensitive to slight changes in 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 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); one such modification that uses quantiles of this ratio is reported in column 4 of Appendix C. (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.) Again, while the results for the elasticity ratio variable are sensitive, the estimate on $[\ln(p_{grc}^{PO}) - \ln(p_{grc}^{BR})]$ is consistently unaffected by these alternative approaches.

³⁷While 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

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).³⁸ 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).³⁹ 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.

We consistently find evidence of a statistically significant and positive relationship between the retaliation capacity variable and the growth of disputed product import volumes. I.e., the larger 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 0.078 from Column (1) of Table 4 implies that a one standard deviation increase in this variable will result in an 17 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 11 percent to 28 percent. Finally, and as will become clear from the discussion below, we also 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

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).

³⁸Bown 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).

³⁹Our 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).

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 an 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.⁴⁰ 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.⁴¹

For the “global” violation subset of disputes, column (1) of Table 4 presents evidence that the higher is the concentration among 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 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. The coefficient estimate of 0.694 implies that a one standard deviation increase in this variable results in a 23 percentage point increase in disputed product import volume growth; i.e., from the baseline prediction (at the means of the data) of 11 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 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 $t = PO$ in Table 4 and find the expected positive relationship, i.e., importing countries enjoying faster rates

⁴⁰For this reason 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.

⁴¹To 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.

of overall growth are also more likely to experience more disputed product import volume growth. On the other hand, controlling for supply shocks at the aggregate level does not lend evidence of a relationship between complainant country economic growth during the period and disputed product trade growth, as the estimated coefficient is 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) took 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 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, there is little evidence that country size or its level of development has a statistically significant independent effect on the import volume growth.

To summarize the results thus far, we find evidence consistent with the theory that the Bagwell and Staiger (1999, 2011) model of trade agreement negotiations also applies to the trade liberalization negotiations taking place 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-level fixed effects. They 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 and the price wedge between the world and domestic price. Such evidence is also consistent with the theory that import volume growth is expected for products in which the international cost-shifting motives are high. Nevertheless, this result is weak and sensitive to variable definition and measurement of the underlying trade elasticities.

Finally, 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). We find that trade dispute outcomes are also 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.

5.3 The impact of institutional features on dispute settlement outcomes

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. 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 the exogenously-determined samples help clarify the extent to which our basic modeling approach to trade liberalization negotiations under dispute settlement can potentially explain patterns in the data.

The estimates in columns (4) and (5) of Table 4 assess a 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 may enjoy an implicit preference (relative to the complainant country) in the respondent country’s disputed product import market. Column (4) reveals that the subsample of disputes associated with challenges to policies applied on a “global” basis have estimates that are consistent with theoretical predictions and are statistically significant. In contrast, the model performs less well in column (5), though the coefficient estimates for the theoretically-motivated covariates are so imprecisely estimated that we cannot rule out that they are equivalent to those reported in column (4).

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. Column (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, forthcoming) contain different predictions for policy outcomes in cases that reach a legal ruling versus those that settle early. Second, 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 4 utilize the current modeling approach to explore whether the complainant country’s capacity to retaliate influences 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.⁴² While 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 (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 itself 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 had little additional effect.

5.4 Dispute settlement outcomes and economic characteristics

Table 5 provides a last set of estimates of our preferred, baseline model specification on alternative subsamples of data defined by their (exogenously determined) economic characteristics. Column (1) provides estimates for the baseline specification from the entire sample of disputes as a point of reference.

The first specifications consider subsamples based on income categories for the countries involved in the disputes. Columns (2) and (3) split the sample 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) split the sample

⁴²In 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.

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) and the ratio of elasticities interacted with pre-dispute import volumes and price wedge between the world and domestic price are both 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 (column 3) or complainant (column 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. However, a separate and theoretically-motivated explanation is that developing countries 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.⁴³

Finally, specifications (6) and (7) split the sample 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.⁴⁴ 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.

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

⁴³Our 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).

⁴⁴In 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.

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 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.

These results also suggest an important avenue for the increasingly prominent theoretical literature on the design of dispute resolution procedures (e.g., Maggi and Staiger 2011, forthcoming). Our evidence is that the successful resolution to a trade dispute - e.g., the importing country's relaxation of an excessively restrictive policy - may be made more difficult, the larger is the (within-market) change in the terms-of-trade associated with the post-dispute policy change. That is, terms-of-trade incentives continue to affect trade policy decisions even after a country is a member of a trade agreement and bound by its legal provisions. Future work might explore the implications of embedding these new theoretical frameworks for dispute resolution into more complete models of trade agreements, especially those motivated by international cost-shifting concerns (e.g., Bagwell and Staiger 1990, 1999).

Finally, 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 (2016) "Is the WTO Passé?" *Journal of Economic Literature* 54(4).
- 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. (2016) "Arbitration and Renegotiation in Trade Agreements," *Journal of Law, Economics, and Organization* 32(3): 586-619.
- 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 (2015) "Trade Flows and Trade Disputes," *Review of International Organizations* 10(2): 145-177.

- 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. 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.
- Iacovone, L., and B. Javorcik (2008) "Shipping good tequila out: Investment, domestic unit values and entry of multi-product plants into export markets." Unpublished manuscript.
- 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 P. Tovar (2011) "Policy Choice: Theory and Evidence from Commitment via International Trade Agreements," *Journal of International Economics* 85(2): 186-205.
- 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 (forthcoming) "Trade Disputes and Settlement," *International Economic Review*.
- Nicita, A., M. Olarreaga, and P. Silva (forthcoming) "Cooperation in WTO's Tariff Waters," *Journal of Political Economy*.
- 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. (2015) "Estimating Import Supply and Demand Elasticities: Analysis and Implications," *Journal of International Economics* 96(1): 1-17.

Staiger, R.W. and A.O. Sykes (2013) “Non-Violations,” *Journal of International Economic Law* 16(4): 741-775.

Staiger, R.W. and A.O. Sykes (forthcoming) “How Important Can the Non-Violation Clause Be for the GATT/WTO?” *American Economic Journal: Microeconomics*.

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

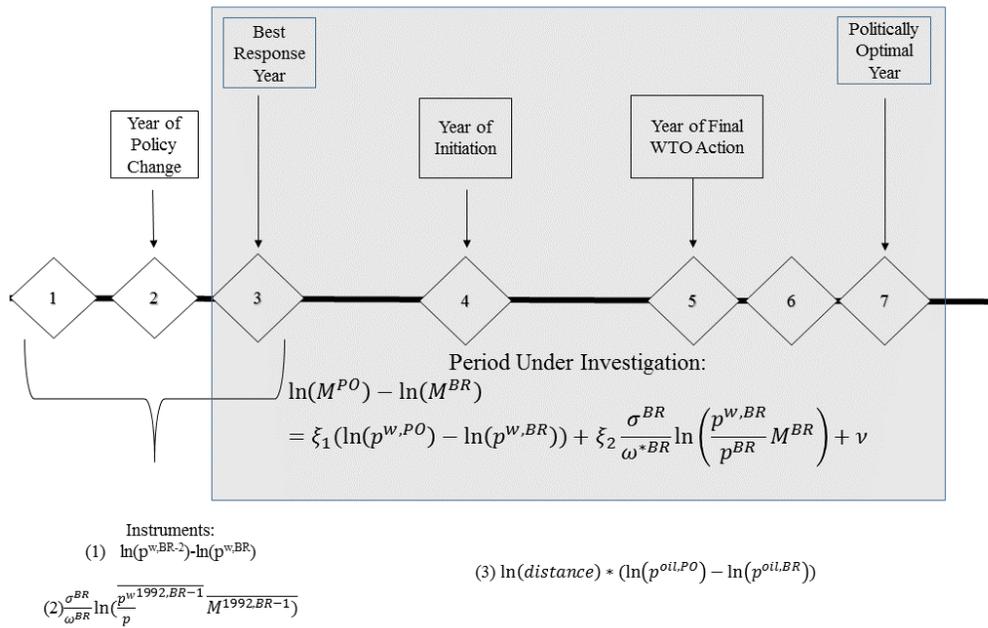


Figure 2: Average Import Volumes and Foreign-Exporter Received Prices for Products Subject to WTO Dispute, 1995-2009

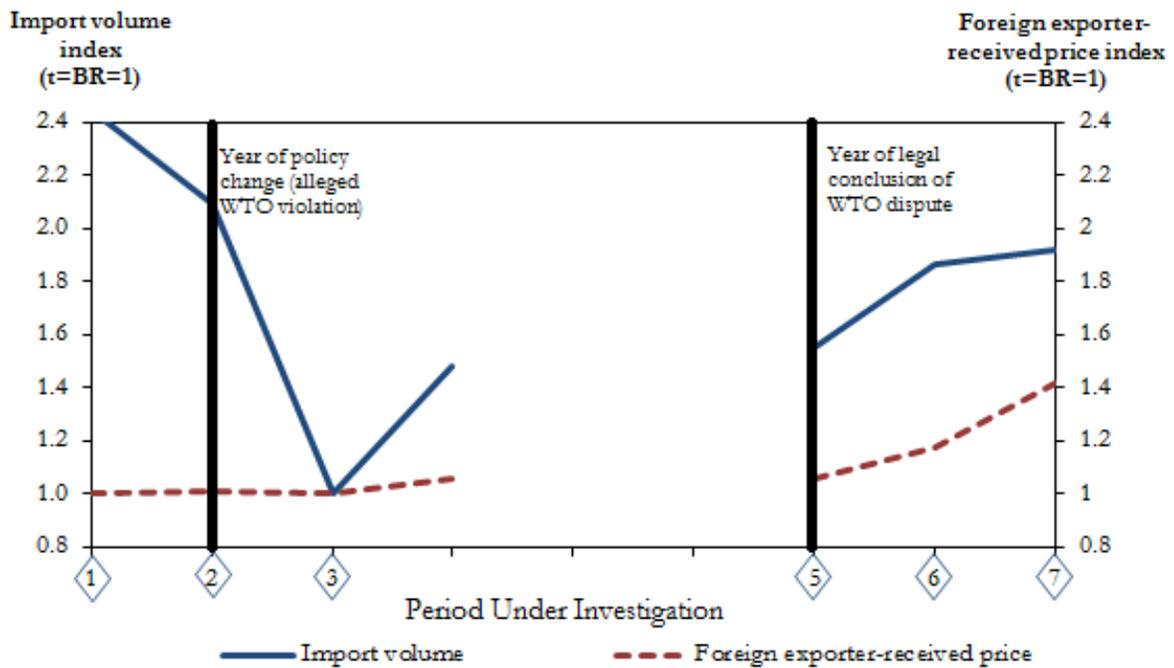


Table 1: WTO Disputes Initiated and Concluded, 1995-2009

	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>	94	
- <i>Disputes in which we are unable to observe elasticities</i>	3	
Final Sample	152	100.0
- <i>Disputes in which the alleged WTO violation is of a policy change</i>	126	82.9
- <i>Disputes in which the alleged WTO violation is of no policy change</i>	26	17.1
- <i>Disputes which result in circulation of at least a Panel Report</i>	78	51.3
- <i>Disputes in which no legal ruling was issued</i>	74	48.7
- <i>Disputes over “global” policies that apply to all trading partners</i>	72	47.4
- <i>Disputes over “partial” policies in which some third country exporters are excluded from application</i>	80	52.6
- <i>Disputes in which the respondent (importer) is high income</i>	96	63.2
- <i>Disputes in which the respondent (importer) is developing</i>	56	36.8
- <i>Disputes in which the complainant (exporter) is high income</i>	89	58.6
- <i>Disputes in which the complainant (exporter) is developing</i>	63	41.4

Table 2: Summary Statistics

	Mean	Std. Dev.	Min	Max
Dependent variable:				
$\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$	0.11	1.18	-6.33	2.91
Explanatory variables:				
<i>Theoretical model determinants:</i>				
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	0.06	0.55	-2.95	3.73
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M_{grc}^{BR})$	-36.65	131.69	-1430.98	38.02
<i>Other political-economic determinants:</i>				
$\ln(\text{Complainant's 'actionable' imports from respondent} / \text{Respondent's imports at stake in dispute})$	11.08	2.17	5.41	17.18
Global disputes	0.47	0.50	0.00	1.00
Global disputes x Exporter concentration	0.23	0.33	0.00	0.99
<i>Other disputed product, timing, or disputant determinants:</i>				
Respondent's real GDP growth between <i>BR</i> and <i>PO</i>	0.12	0.13	-0.20	0.74
Complainants's real GDP growth between <i>BR</i> and <i>PO</i>	0.13	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	0.14	0.00	1.00
$\ln(\text{Respondent's real GDP (billions)})$	7.62	1.96	3.15	9.60
$\ln(\text{Complainant's real GDP (billions)})$	7.19	2.03	2.78	9.60
$\ln(\text{Respondent's real GDP per capita})$	9.59	1.13	6.85	10.72
$\ln(\text{Complainant's real GDP per capita})$	9.36	1.32	6.36	11.09

Note: 152 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	Partial	Global	t- test	High Income Respondent	Emerging Respondent	t- test	High Income Complainant	Emerging Complainant	t- test
Average Import Growth between BR and PO (median)	0.11 (0.18)	0.18 (0.26)	0.03 (0.11)		0.05 (0.10)	0.17 (0.25)		0.19 (0.19)	-0.03 (0.06)		-0.01 (0.08)	0.28 (0.27)	
Average Import Growth per Annum between BR and PO (median)	0.04 (0.05)	0.06 (0.05)	0.01 (0.04)		0.02 (0.04)	0.05 (0.06)		-0.06 (0.05)	0.01 (0.02)		0.01 (0.04)	0.07 (0.07)	
Share of disputes with positive import growth between BR and PO	0.63	0.67	0.58		0.56	0.69		0.69	0.52	***	0.55	0.73	***
Share of disputes with import growth larger than RGDP growth	0.58	0.62	0.54		0.55	0.61		0.64	0.48		0.51	0.68	***
Average years between BR and PO	4.36	4.79	3.91	***	4.39	4.33		4.22	4.61		4.27	4.49	
Disputes	152	78	74		80	72		96	56		89	63	

Note: *** denotes t-test statistically different from zero at the 1 percent levels.

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

Dependent variable is the growth in import volumes in products under dispute: $\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	-1.017	-0.911	-1.062	-0.708	-1.588	-1.194	-0.250	-0.781	-1.842
	(0.241)	(0.205)	(0.188)	(0.242)	(0.406)	(0.255)	(0.588)	(0.227)	(0.879)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M_{grc}^{BR})$	0.001	0.002	-0.000	0.002	-0.001	0.002	-0.001	0.002	0.000
	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
ln(Complainant's 'actionable' imports from respondent/ Respondent's imports at stake in dispute)	0.078	0.078	0.065	0.084	0.104	0.091	0.062	0.078	0.039
	(0.041)	(0.041)	(0.047)	(0.045)	(0.069)	(0.062)	(0.055)	(0.090)	(0.085)
Global Disputes	-0.389	-0.282	-0.019			-0.566	-0.266	-0.169	-0.773
	(0.223)	(0.231)	(0.270)			(0.348)	(0.312)	(0.303)	(0.346)
Global Disputes x Exporter Concentration	0.694	0.405	0.159	0.539		0.893	1.011	1.127	0.752
	(0.340)	(0.344)	(0.373)	(0.285)		(0.483)	(0.591)	(0.449)	(0.519)
Respondent's real GDP growth between <i>BR</i> and <i>PO</i>	1.937	0.764	-0.386	2.649	1.757	1.318	3.452	0.034	3.783
	(0.721)	(1.539)	(1.672)	(0.762)	(1.405)	(0.998)	(1.270)	(0.999)	(1.038)
Complainant's real GDP growth between <i>BR</i> and <i>PO</i>	-1.406	-1.117	-0.919	-1.357	-0.847	-0.039	-3.566	1.118	-2.833
	(0.963)	(1.592)	(1.711)	(1.503)	(1.644)	(1.347)	(1.673)	(1.437)	(1.251)
Dispute's <i>PO</i> year is during trade collapse of 2008-9	-0.208	-0.390	-0.197	-0.816	0.061	-0.614	0.050	-0.111	0.074
	(0.261)	(0.235)	(0.238)	(0.349)	(0.384)	(0.363)	(0.403)	(0.328)	(0.499)
Dispute's <i>BR</i> year is during trade collapse of 2008-9	-0.511	-1.211	-1.229	-0.560	-0.288		-0.924	-0.521	0.626
	(0.563)	(0.537)	(0.603)	(0.818)	(0.793)		(0.673)	(0.674)	(1.909)
ln(Respondent's real GDP)	0.072	-3.386	3.329	0.033	-0.066	0.081	0.077	0.127	0.031
	(0.065)	(5.390)	(5.508)	(0.075)	(0.130)	(0.116)	(0.084)	(0.103)	(0.091)
ln(Complainant's real GDP)	-0.092	10.518	9.723	-0.076	-0.156	-0.111	-0.098	-0.113	0.002
	(0.051)	(3.492)	(3.879)	(0.062)	(0.079)	(0.076)	(0.073)	(0.070)	(0.074)
ln(Respondent's real GDP per capita)	0.070	1.881	-5.166	-0.001	0.479	-0.039	0.150	0.057	0.124
	(0.117)	(6.202)	(6.127)	(0.128)	(0.247)	(0.192)	(0.160)	(0.181)	(0.161)
ln(Complainant's real GDP per capita)	-0.027	-12.244	-12.555	0.013	0.007	0.027	-0.157	-0.195	0.05
	(0.083)	(3.732)	(4.222)	(0.108)	(0.131)	(0.110)	(0.147)	(0.138)	(0.113)
Country Fixed Effects	No	Yes	Yes	No	No	No	No	No	No
Sector Fixed Effects	No	No	Yes	No	No	No	No	No	No
							No	High	Low
Sample	Full	Full	Full	Global	Partial	Panel	Panel	Retaliation	Retaliation
Observations	152	152	152	72	80	78	74	76	76
Root MSE	0.915	0.644	0.596	0.732	0.982	0.917	0.847	0.842	0.905
Sargan Test	0.232	0.0620	0.166	0.246	0.432	0.491	0.338	0.510	0.192
Cragg-Donald	24.33	16.69	17.90	31.72	7.422	16.97	5.068	20.30	2.365

Note: Standard errors in parentheses. Coefficients from constant term not reported. 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 Level of Economic Development and Trade Outcomes

Dependent variable is the growth in import volumes in products under dispute: $\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	-1.017 (0.241)	-0.864 (0.273)	-1.569 (0.338)	-0.950 (0.192)	-5.193 (6.433)	-0.710 (0.186)	-1.621 (0.378)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M_{grc}^{BR})$	0.001 (0.001)	0.002 (0.001)	-0.002 (0.002)	0.002 (0.001)	-0.004 (0.005)	-0.001 (0.001)	0.001 (0.001)
ln(Complainant's 'actionable' imports from respondent/ Respondent's imports at stake in dispute)	0.078 (0.041)	0.147 (0.046)	-0.033 (0.074)	0.041 (0.045)	0.022 (0.274)	0.061 (0.036)	0.032 (0.063)
Global Disputes	-0.389 (0.223)	-0.456 (0.248)	-0.177 (0.444)	-0.002 (0.253)	-1.125 (0.923)	-0.396 (0.195)	-0.296 (0.300)
Global Disputes x Exporter Concentration	0.694 (0.340)	0.573 (0.372)	0.480 (0.680)	0.581 (0.359)	0.423 (1.213)	0.029 (0.296)	0.703 (0.496)
Respondent's real GDP growth between <i>BR</i> and <i>PO</i>	1.937 (0.721)	2.462 (1.328)	2.629 (1.101)	1.152 (0.764)	4.240 (2.791)	2.631 (0.713)	1.929 (1.089)
Complainant's real GDP growth between <i>BR</i> and <i>PO</i>	-1.406 (0.963)	-0.787 (1.110)	-3.214 (1.714)	-3.533 (1.433)	-0.450 (3.451)	0.821 (0.802)	-3.825 (1.472)
Dispute's <i>PO</i> year is during trade collapse of 2008-9	-0.208 (0.261)	-0.459 (0.313)	0.136 (0.418)	-0.014 (0.313)	0.525 (1.396)	0.159 (0.244)	-0.503 (0.340)
Dispute's <i>BR</i> year is during trade collapse of 2008-9	-0.511 (0.563)	-0.550 (0.848)	0.211 (0.760)	-0.691 (0.854)	2.545 (6.312)	0.075 (0.638)	0.567 (0.703)
ln(Respondent's real GDP)	0.072 (0.065)	0.059 (0.100)	0.039 (0.143)	0.178 (0.078)	-0.355 (0.261)	-0.049 (0.070)	0.090 (0.083)
ln(Complainant's real GDP)	-0.092 (0.051)	-0.096 (0.057)	-0.093 (0.115)	-0.119 (0.060)	0.133 (0.201)	-0.068 (0.040)	0.004 (0.081)
ln(Respondent's real GDP per capita)	0.070 (0.117)	0.140 (0.342)	0.243 (0.233)	-0.066 (0.123)	1.404 (1.093)	0.118 (0.138)	-0.002 (0.137)
ln(Complainant's real GDP per capita)	-0.027 (0.083)	0.010 (0.082)	0.020 (0.210)	-0.334 (0.269)	0.335 (0.570)	-0.012 (0.062)	0.012 (0.129)
Sample	All	High Income Respondent	Emerging Respondent	High Income Complainant	Emerging Complainant	High Growth Disputes	Small Growth Disputes
Observations	152	96	56	89	63	76	76
Root MSE	0.915	0.809	0.904	0.769	1.646	0.529	0.855
Sargan Test	0.232	0.575	0.0102	0.361	0.854	0.00719	0.0463
Cragg-Donald	24.33	22.08	7.851	25.27	0.154	24.71	7.009

Note: Standard errors in parentheses. Coefficients from constant term not reported. 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: $\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\ln(p_{grc}^{w,BR-2}) - \ln(p_{grc}^{w,BR})$	0.858	0.887	0.985	0.796	1.039	0.961	0.628	0.977	0.511
	(0.101)	(0.127)	(0.134)	(0.084)	(0.238)	(0.138)	(0.162)	(0.126)	(0.195)
$\ln(distance_{rc})x(\ln(p^{oil,PO}) - \ln(p^{oil,BR}))$	0.020	0.028	0.016	-0.011	0.047	0.043	0.011	0.017	0.003
	(0.013)	(0.018)	(0.020)	(0.013)	(0.021)	(0.023)	(0.015)	(0.019)	(0.017)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,1992,BR-1}}{p_{grc}} - \frac{M_{grc}^{1992,BR-1}}{M_{grc}})$	-0.000	-0.000	0.001	-0.000	0.000	-0.000	-0.001	-0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
ln(Complainant's 'actionable' imports from respondent/ Respondent's imports at stake in dispute Global Disputes	0.021	0.018	0.065	0.034	0.012	0.024	0.016	0.084	0.008
	(0.020)	(0.028)	(0.035)	(0.019)	(0.036)	(0.035)	(0.024)	(0.051)	(0.036)
	-0.184	0.041	-0.159			-0.209	-0.141	-0.221	-0.188
	(0.108)	(0.161)	(0.196)			(0.189)	(0.137)	(0.171)	(0.142)
Global Disputes x Exporter Concentration	-0.063	-0.428	-0.073	0.142		-0.027	-0.145	0.034	0.102
	(0.165)	(0.228)	(0.274)	(0.122)		(0.265)	(0.256)	(0.261)	(0.223)
Respondent's real GDP growth between BR and PO	0.205	1.244	1.196	-0.486	0.706	0.210	0.170	-0.181	0.017
	(0.355)	(1.040)	(1.202)	(0.323)	(0.727)	(0.547)	(0.568)	(0.567)	(0.457)
Complainant's real GDP growth between BR and PO	-0.134	-1.146	-0.621	0.419	-0.659	-0.621	0.072	0.537	-0.293
	(0.500)	(1.132)	(1.253)	(0.659)	(0.893)	(0.804)	(0.745)	(0.907)	(0.580)
Dispute's PO year is during trade collapse of 2008-9	0.021	0.026	-0.080	0.092	-0.004	-0.197	0.201	0.009	0.215
	(0.128)	(0.162)	(0.175)	(0.155)	(0.201)	(0.207)	(0.170)	(0.188)	(0.179)
Dispute's BR year is during trade collapse of 2008-9	0.549	0.748	0.917	-0.459	0.916		0.705	-0.270	1.823
	(0.268)	(0.359)	(0.408)	(0.344)	(0.384)		(0.234)	(0.395)	(0.411)
ln(Respondent's real GDP)	-0.017	6.888	4.481	0.069	-0.094	-0.069	-0.009	0.046	-0.009
	(0.032)	(3.326)	(3.749)	(0.031)	(0.064)	(0.065)	(0.037)	(0.058)	(0.039)
ln(Complainant's real GDP)	-0.007	-0.055	-1.724	0.019	-0.011	-0.018	-0.030	0.005	0.010
	(0.025)	(2.414)	(2.871)	(0.026)	(0.042)	(0.042)	(0.032)	(0.040)	(0.032)
ln(Respondent's real GDP per capita)	0.073	-7.441	-3.884	-0.061	0.191	0.093	0.067	-0.067	0.082
	(0.058)	(3.878)	(4.215)	(0.053)	(0.119)	(0.108)	(0.068)	(0.103)	(0.071)
ln(Complainant's real GDP per capita)	0.067	-0.501	1.134	-0.014	0.086	0.005	0.165	0.030	0.055
	(0.040)	(2.581)	(3.110)	(0.046)	(0.066)	(0.061)	(0.055)	(0.078)	(0.043)
Country Fixed Effects	No	Yes	Yes	No	No	No	No	No	No
Sector Fixed Effects	No	No	Yes	No	No	No	No	No	No
							No	High	Low
Sample	Full	Full	Full	Global	Partial	Panel	Panel	Retaliation	Retaliation
Observations	152	152	152	72	80	78	74	76	76
Root MSE	0.445	0.445	0.432	0.310	0.513	0.504	0.369	0.475	0.386

Note: Coefficient estimates associate with the first stage results from the instrumental variable regressions associated with Table 4. Standard errors in parentheses. Coefficients from constant term not reported.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\ln(p_{grc}^{w, BR-2}) - \ln(p_{grc}^{w, BR})$	0.461 (3.340)	1.804 (3.667)	1.457 (3.309)	-1.955 (5.142)	0.303 (2.919)	-2.247 (4.435)	3.931 (3.858)	1.198 (4.525)	-0.457 (3.824)
$\ln(distance_{rc})x(\ln(p^{oil, PO}) - \ln(p^{oil, BR}))$	-0.784 (0.416)	-0.112 (0.519)	-0.132 (0.501)	-1.257 (0.808)	-0.462 (0.258)	-1.486 (0.746)	-0.663 (0.354)	-1.340 (0.691)	-0.823 (0.332)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln\left(\frac{p_{grc}^{w, 1992, BR-1}}{p_{grc}^{BR}} \frac{M_{grc}^{1992, BR-1}}{M_{grc}^{BR}}\right)$	0.950 (0.009)	0.964 (0.009)	0.883 (0.017)	0.963 (0.013)	0.847 (0.013)	0.973 (0.011)	0.886 (0.017)	0.967 (0.012)	0.822 (0.013)
ln(Complainant's 'actionable' imports from respondent/ Respondent's imports at stake in dispute Global Disputes	-1.877 (0.663) -1.377 (3.582)	-2.286 (0.819) -10.183 (4.648)	-1.316 (0.852) -7.038 (4.840)	-2.380 (1.182)	-0.548 (0.440)	-0.824 (1.118) -0.772 (6.048)	-1.549 (0.575) -4.811 (3.259)	-3.734 (1.841) -4.767 (6.113)	-0.495 (0.714) 7.120 (2.780)
Global Disputes x Exporter Concentration	12.365 (5.459)	12.103 (6.606)	10.551 (6.760)	12.952 (7.501)		19.118 (8.488)	13.256 (6.119)	18.256 (9.343)	-1.307 (4.386)
Respondent's real GDP growth between BR and PO	-16.057 (11.760)	-0.017 (30.102)	-52.941 (29.661)	-19.695 (19.867)	-4.129 (8.916)	-42.253 (17.540)	6.024 (13.566)	-25.960 (20.320)	2.542 (8.978)
Complainant's real GDP growth between BR and PO	31.776 (16.585)	18.670 (32.761)	80.622 (30.930)	55.582 (40.560)	16.150 (10.954)	71.377 (25.778)	16.262 (17.789)	59.191 (32.501)	24.321 (11.386)
Dispute's PO year is during trade collapse of 2008-9	-2.851 (4.233)	-1.429 (4.692)	3.750 (4.308)	-7.635 (9.524)	-0.358 (2.459)	-5.535 (6.641)	1.540 (4.051)	-1.114 (6.744)	-3.608 (3.516)
Dispute's BR year is during trade collapse of 2008-9	-12.102 (8.900)	-2.154 (10.395)	6.451 (10.080)	-36.270 (21.164)	-5.277 (4.714)		-8.187 (5.585)	-25.118 (14.138)	1.257 (8.071)
ln(Respondent's real GDP)	3.206 (1.061)	-39.718 (96.287)	118.157 (92.505)	4.510 (1.900)	0.584 (0.786)	8.788 (2.091)	0.274 (0.872)	3.244 (2.078)	3.220 (0.774)
ln(Complainant's real GDP)	2.012 (0.822)	-32.453 (69.881)	-72.167 (70.840)	2.343 (1.601)	1.026 (0.516)	1.610 (1.341)	2.150 (0.775)	2.767 (1.424)	1.339 (0.619)
ln(Respondent's real GDP per capita)	-7.297 (1.907)	38.159 (112.251)	-121.301 (104.010)	-12.223 (3.287)	-0.106 (1.460)	-18.937 (3.469)	1.037 (1.618)	-9.148 (3.683)	-5.009 (1.386)
ln(Complainant's real GDP per capita)	0.143 (1.325)	30.235 (74.711)	9.867 (76.753)	-0.049 (2.810)	-0.137 (0.814)	2.352 (1.970)	-2.168 (1.318)	-0.871 (2.794)	0.653 (0.853)
Country Fixed Effects	No	Yes	Yes	No	No	No	No	No	No
Sector Fixed Effects	No	No	Yes	No	No	No	No	No	No
							No	High	Low
Sample	Full	Full	Full	Global	Partial	Panel	Panel	Retaliation	Retaliation
Observations	152	152	152	72	80	78	74	76	76
Root MSE	14.76	12.87	10.66	19.05	6.291	16.15	8.813	17.02	7.585

Note: Coefficient estimates associate with the first stage results from the instrumental variable regressions associated with Table 4. Standard errors in parentheses. Coefficients from constant term not reported.

Table Appendix C: Sensitivity Analysis

	Dependent variable is the growth in import volumes in products under dispute: $\ln(M_{grc}^{PO}) - \ln(M_{grc}^{BR})$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	-1.092 (0.284)	-1.019 (0.241)	-1.092 (0.142)	-0.982 (0.235)	-1.265 (0.302)	-1.044 (0.262)	-1.007 (0.241)
$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M^{BR})$	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)		0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
$\frac{\sigma_{gr}^{BR} q}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M^{BR})$				0.020 (0.011)			
ln(Complainant's 'actionable' imports from respondent/ Respondent's imports at stake in dispute Global Disputes Global Disputes x Exporter Concentration Respondent's real GDP growth between BR and PO Complainant's real GDP growth between BR and PO Dispute's PO year is during trade collapse of 2008-9 Dispute's BR year is during trade collapse of 2008-9 ln(Respondent's real GDP) ln(Complainant's real GDP) ln(Respondent's real GDP per capita) ln(Complainant's real GDP per capita)	0.080 (0.041) -0.394 (0.241) 0.679 (0.266) 1.946 (0.789) -1.376 (1.104) -0.205 (0.466) -0.470 (0.274) 0.074 (0.070) -0.094 (0.044) 0.069 (0.121) -0.020 (0.080)	0.079 (0.041) -0.383 (0.223) 0.690 (0.340) 1.922 (0.721) -1.376 (0.963) -0.212 (0.261) -0.512 (0.563) 0.073 (0.065) -0.093 (0.051) 0.067 (0.117) -0.025 (0.083)	0.079 (0.041) -0.401 (0.221) 0.683 (0.338) 1.963 (0.718) -1.410 (0.962) -0.201 (0.260) -0.468 (0.551) 0.073 (0.065) -0.093 (0.051) 0.072 (0.117) -0.022 (0.082)	0.118 (0.049) -0.421 (0.230) 0.656 (0.345) 1.643 (0.736) -2.176 (0.979) -0.372 (0.267) -0.323 (0.676) 0.060 (0.068) -0.116 (0.053) 0.060 (0.120) -0.075 (0.089)	-0.013 (0.045) -0.028 (0.243) 0.155 (0.372) 2.335 (0.792) -0.986 (1.024) 0.120 (0.288) -0.560 (0.600) 0.012 (0.069) -0.043 (0.054) 0.156 (0.126) 0.021 (0.088)	0.086 (0.048) -0.315 (0.262) 0.669 (0.400) 1.804 (0.809) -1.109 (1.095) -0.254 (0.303) -0.794 (1.106) 0.069 (0.078) -0.101 (0.058) 0.073 (0.138) -0.016 (0.093)	-1.007 (0.241) 0.001 (0.001) 0.076 (0.042) -0.378 (0.225) 0.726 (0.341) 1.752 (0.754) -1.753 (1.070) -0.492 (0.291) -0.598 (0.700) 0.073 (0.066) -0.089 (0.053) 0.047 (0.119) -0.032 (0.086)
Best Response Period	3	3	3	3	4	3	3
Endogenous Variables	None	$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	$\frac{\sigma_{gr}^{BR}}{\omega_{gr}^{*BR}} \ln(\frac{p_{grc}^{w,BR}}{p_{grc}^{BR}} M^{BR})$	$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$	$\ln(p_{grc}^{w,PO}) - \ln(p_{grc}^{w,BR})$
Sample	All	All	All	All	All	Resolved	Non-AD
Observations	152	152	152	139	148	127	137
Root MSE	0.959	0.915	0.914	0.905	0.962	0.978	0.914
Cragg-Donald		36.48	10726	28.54	12.02	22.48	26.05
Sargan Test				0.147	0.553	0.351	0.0790

Note: Standard errors in parentheses. Coefficients from constant term not reported. Note that g indexes the disputes (products), r indexes respondent (importing) countries, c indexes complainant (exporting) countries, BR denotes the best response year, 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.