World Trade Organization
Economic Research and Statistics Division

This paper appears in the WTO working paper series in the context of the WTO Academic Visiting Program

THE WTO: THEORY AND PRACTICE

Kyle Bagwell: Stanford and NBER
Robert W. Staiger: Stanford and NBER
Manuscript date: October 12, 2009

Disclaimer: This is a working paper, and hence it represents research in progress. This paper represents the opinions of the authors, and is the product of professional research. It is not meant to represent the position or opinions of the WTO or its Members, nor the official position of any staff members. Any errors are the fault of the authors. Copies of working papers can be requested from the divisional secretariat by writing to: Economic Research and Statistics Division, World Trade Organization, Rue de Lausanne 154, CH 1211 Geneva 21, Switzerland. Please request papers by number and title.
The WTO: Theory and Practice*

Kyle Bagwell  Robert W. Staiger
Stanford and NBER  Stanford and NBER

October 12, 2009

Abstract

We consider the purpose and design of the World Trade Organization (WTO) and its predecessor, GATT. We review recent developments in the relevant theoretical and empirical literature. And we describe the GATT/WTO architecture and briefly trace its historical antecedents. We suggest that the existing literature provides a useful framework for understanding and interpreting central features of the design and practice of the GATT/WTO, and we identify key unresolved issues.

*We thank our reviewer, Peter Neary, for very helpful comments. Parts of this Review draw heavily from an earlier review article (Bagwell and Staiger, 2002a) and on excerpts from Chapter 7, pp. 205-234, “Economic Theory and the Interpretation of GATT/WTO” by Kyle Bagwell and Robert W. Staiger in New Frontiers in Economics, edited by Michael Szenberg and Lall Ramrattan Copyright © 2004 Cambridge University Press. Reprinted with permission.
1 Introduction

As a subject for study by economists, the World Trade Organization (WTO) and its predecessor, the General Agreement on Tariffs and Trade (GATT), exhibit a number of interesting and attractive features. First, the GATT/WTO is widely acknowledged to be one of the most successful international institutions ever created: on this basis alone it is important to understand the reasons for its success. Second, though it is a multilateral institution, the GATT/WTO has adopted a bilateral approach to multilateral bargaining according to which reciprocal negotiations (over tariffs) occur on a voluntary basis through time between pairs of countries or among small numbers of countries, with the results of these bilateral negotiations then "multilateralized" to the full GATT/WTO membership by a non-discrimination requirement that tariffs abide by the most-favored nation (MFN) principle. This approach is distinctive among multilateral institutions, raising the question whether it can help account for the success of the GATT/WTO as a negotiating forum; and it has also left in its wake a wealth of bargaining data extending back to the creation of GATT in 1947, the analysis of which could yield important insights for bargaining theory. Third, while commitments negotiated within the GATT/WTO must be self-enforcing to be effective, much like the collusive agreements among firms that have been the subject of a large literature in industrial organization, the GATT/WTO contains its own explicit enforcement/dispute settlement procedure and associated case-law history, the study of which can yield insights into the nature of self-enforcing agreements. Fourth, while the GATT/WTO has in effect served as the constitution of the post-war international trading system, it is a highly incomplete contract, and understanding the nature of this incompleteness is an important task for which many of the lessons from contract theory may be relevant. And finally, most trade-policy decisions that governments face today arise in the context of a variety of international commitments that must be considered; hence, the study of commercial policy in international trade has in effect become the study of trade agreements, where the GATT/WTO plays a central role.

In this Review we describe how recent economic research attempts to understand and interpret the design and practice of the GATT/WTO. We survey both theoretical developments and related empirical work. Our Review proceeds in four broad steps.

First, we survey the existing theories of trade agreements. We organize our discussion here around a simple but fundamental question: What is the problem that a trade agreement might solve? The literature identifies two possibilities: governments may view trade agreements as helping them avoid "beggar-my-neighbor" policies that are unilaterally attractive but mutually destructive; or governments may view trade agreements as helping them avoid "beggar-myself" policies that, although not serving a government's own ex-ante objectives, are nevertheless attractive to the government ex post at the time when it chooses its trade policies. The latter possibility has been

---

1 The GATT was created in 1947, while the WTO came into existence on January 1, 1995, as a result of the Marrakesh Agreement of April 1994, also known as the WTO Agreement. The GATT continues to exist as a substantive agreement within the WTO Agreement, but the WTO Agreement also includes a set of additional agreements that build on and extend GATT principles to new areas. Hoekman and Kosteki (2001) provide an excellent institutional overview of GATT and the WTO.
formalized in the literature, and it has been suggested that trade agreements may indeed play this role. But it is the first possibility – that trade agreements help governments avoid beggar-my-neighbor trade policies – that is embodied in the theoretical attempts to understand and interpret the key design features of the GATT/WTO. And so it is this possibility that makes up the primary focus of our survey. Strikingly, as we describe below, in a wide variety of formal settings the beggar-my-neighbor trade policies which give rise to the need for a trade agreement are driven by a single underlying motivation: terms-of-trade manipulation.

The essential insight that terms-of-trade manipulation can give rise to a problem that a trade agreement might solve dates back to Mill (1844) and Torrens (1844), and the first formal treatment of the idea was provided by Johnson (1953-54). But terms-of-trade manipulation is often dismissed as an empirically irrelevant possibility. Given the central role that terms-of-trade manipulation is now understood to play in theories of trade agreements, it is therefore important to revisit the empirical evidence on this basic question. This is the second step of our Review. Here we survey the growing body of empirical literature and suggest that the empirical relevance of terms-of-trade manipulation is much greater than has been widely believed.

In the third step of our Review, we describe the GATT/WTO architecture and briefly trace its historical antecedents. We suggest that the design of GATT reflects lessons learned from the successes and failures experienced over decades of European and U.S. tariff bargaining.

This sets the stage for our fourth and final step, where we draw on the recent theoretical and empirical literature to interpret the design and practice of the GATT/WTO. Our discussion highlights three of the most central features of the GATT/WTO – reciprocity, non-discrimination and enforcement/dispute settlement.

2 What is the Purpose of a Trade Agreement?

All theories of trade agreements must identify a reason why negotiating governments can gain from the agreement. This involves identifying a “problem” that would arise absent an agreement, when governments make noncooperative trade-policy choices. The purpose of the agreement can then be viewed as providing a “solution” to the problem, and the negotiating governments may share in the associated benefits. It is not just confirming the existence of a problem that is important: a clear understanding of the problem and its structure can also provide important guidance for the design of an institution that can most effectively aid governments in their efforts to find a solution.

One branch of the literature posits that important international externalities arise when governments make trade-policy choices, and that the purpose of a trade agreement is then to internalize these externalities. We survey this approach in section 2.1 The other branch of the literature posits that governments face important commitment problems with regard to the private sector when they make trade-policy choices, and that the purpose of a trade agreement is then to serve as an external commitment device that can “tie the hands” of its members. We briefly survey this approach in
2.1 Addressing Beggar-My-Neighbor Policies

When a government pursues a beggar-my-neighbor policy, a portion of the benefits that it perceives from the policy comes at the expense of other countries. This describes an international externality associated with the government’s policy choice. But what form does the international externality take?

To answer this question, we start with a setting in which markets are perfectly competitive, we describe and interpret the terms-of-trade externality that can arise, and we show that the realistic possibility that governments pursue political/distributional objectives does not introduce additional international externalities. In so doing, we present the central insights from the terms-of-trade theory of trade agreements. We then allow the possibility that markets are not perfectly competitive, we describe the new international externalities that can arise in this alternative setting, and we consider the possibility that these new externalities might give rise to alternative theories of trade agreements.

2.1.1 Terms-of-Trade Externalities

In the benchmark setting of perfectly competitive markets, a famous result holds that unilateral free trade is optimal, when a government maximizes national income and presides over a small country.\footnote{We abstract from domestic distortions for the purposes of this discussion, although the main points of this section carry over to the case of perfectly competitive markets in which domestic distortions are present (see Bagwell and Staiger, 2001a).} For an economist seeking to understand the GATT/WTO, this result is initially discouraging, since in these circumstances governments have no reason to pursue reciprocal tariff liberalization through GATT/WTO negotiations. Nevertheless, the result is instructive. It suggests that a trade agreement might solve a problem that arises because governments have political objectives and/or preside over large countries.

Clearly, real-world governments have political motivations and are interested not just in the size of national income but also in its distribution. And the optimal unilateral policy for a politically motivated government may not be free trade. But this is not the same as saying that political considerations represent a problem that governments might use a trade agreement to solve. Indeed, as we discuss below, in the leading political-economy models of trade policy, governments of small countries do not gain from a trade agreement among themselves; thus, in these models, politics alone cannot explain the appeal of a trade agreement.

What if governments preside over large countries? As we observed in the Introduction, that governments of large countries may succumb to the temptation to engage in terms-of-trade manipulation and thus gain from a trade agreement is not a new insight. Many trade economists, however, have objected to using this theory to explain actual trade agreements. In part these objections are empirical, and we survey the evidence in section 3. But in part these objections are rooted in
the way that the terms-of-trade theory has been traditionally developed. One objection is that the important political constraints under which real-world governments operate are omitted from the theory. A second objection is that the theory does not capture the way that actual governments think. The “terms of trade” are rarely mentioned in actual trade-policy debates, for example. We argue below, though, that these objections are less damaging for the terms-of-trade theory than they might initially appear. The theory can be generalized to include political objectives, and it also may be understood in terms of the market-access language that arises in trade-policy debates.

To develop these points, we first review the textbook two-good general-equilibrium model of trade between two countries. We next define a general family of government preferences. Using this framework, we then characterize and interpret the problem that a trade agreement can solve.

**The Model** Two countries, domestic and foreign, trade two goods which are normal in consumption and produced in perfectly competitive markets under conditions of increasing opportunity costs. We let \( x (y) \) denote the natural import good of the domestic (foreign) country. The local relative price facing domestic (foreign) producers and consumers is defined as \( p \equiv p_x/p_y \) \((p^* \equiv p_x^*/p_y^*)\). We assume that tariffs are non-prohibitive and represent the domestic (foreign) ad valorem import tax as \( t \) \((t^*)\). Letting \( \tau \equiv (1 + t) \) and \( \tau^* \equiv (1 + t^*) \), we then have that \( p = \tau p^w \equiv p(\tau, p^w) \) and \( p^* = p^w/\tau^* \equiv p^*(\tau^*, p^w) \), where \( p^w \equiv p_x^*/p_y^* \) is the “world” (i.e., untaxed) relative price. The foreign terms of trade is given by \( p^w \), and the domestic terms of trade is thus \( 1/p^w \). We interpret \( \tau > 1 \) as an import tax and similarly for \( \tau^* \).

In each country, production levels for \( x \) and \( y \) are determined by the local relative price. Domestic and foreign production functions may then be written as \( Q_i = Q_i(p) \) and \( Q_i^* = Q_i^*(p^*) \) for \( i = \{x, y\} \). National consumption is a function of local and world prices, \( C_i(p, p^w) \) and \( C_i^*(p^*, p^w) \) for \( i = \{x, y\} \): the local relative price defines the trade-off faced by consumers, and it also determines the level and distribution of factor income; and together with the local price, the world price determines tariff revenue (which is distributed lump-sum to consumers). Imports of \( x \) and exports of \( y \) for the domestic country are then respectively defined by \( M_x(p, p^w) \equiv C_x(p, p^w) - Q_x(p) \) and \( E_y(p, p^w) \equiv Q_y(p) - C_y(p, p^w) \). For the foreign country, we similarly define imports of \( y \) and exports of \( x \) as \( M_y^*(p^*, p^w) \) and \( E_x^*(p^*, p^w) \), respectively. For any prices, domestic and foreign budget constraints are represented as

\[ p^w M_x(p, p^w) = E_y(p, p^w), \quad \text{(1)} \]

\[ M_y^*(p^*, p^w) = p^w E_x^*(p^*, p^w). \quad \text{(2)} \]

The equilibrium world price, \( \tilde{p}^w(\tau, \tau^*) \), is determined by the requirement of market-clearing for

---

4 Henceforth, we distinguish foreign variables from domestic by placing an asterisk on the former.

5 Below, we simplify notation and use \( p \) to denote the function \( p(\tau, p^w) \), and \( p^* \) to denote the function \( p^*(\tau^*, p^w) \).

6 Lerner symmetry implies that it is immaterial whether trade taxes and subsidies are depicted as applying to the import good or rather the export good.
good \( y \):
\[
E_y(p(\tau, \tilde{p}^w), \tilde{p}^w) = M^*_y(p^* (\tau^*, \tilde{p}^w), \tilde{p}^w),
\]  
(3)

where we make explicit in (3) the functional dependencies for local prices. Market clearing for good \( x \) is now guaranteed by (1), (2) and (3).

We assume that the Metzler and Lerner paradoxes are ruled out; thus, we assume \( dp/d\tau > 0 \) and \( \frac{\partial \tilde{p}^w}{\partial \tau} < 0 < \frac{\partial \tilde{p}^w}{\partial \tau^*} \). The final two inequalities indicate that each country is “large”: each country can improve its terms of trade by increasing its tariff.

**Government Preferences** The traditional approach to representing government preferences is to impose the assumption that governments maximize national income (Kowalczyk and Riezman, 2009, survey the traditional approach). By contrast, in the political-economy approach, governments are motivated by distributional concerns. Here, we follow Bagwell and Staiger (1999a, 2002b) and adopt a general approach to modeling government preferences, representing the objectives of the domestic and foreign governments with the general functions \( W(p, \tilde{p}^w) \) and \( W^*(p^*, \tilde{p}^w) \), respectively. We thus represent welfare in terms of the prices that the tariffs induce rather than directly in terms of the tariffs themselves.

We place no restrictions on government preferences over local prices: as local prices determine the level and distribution of factor incomes, this allows us to incorporate a wide range of political motivations. In fact, we impose only one assumption on the welfare functions. We assume that, holding its local price fixed, each government experiences an improvement in its welfare when its terms of trade improve:

\[
W_{\tilde{p}^w} < 0 \quad \text{and} \quad W^*_{\tilde{p}^w} > 0.
\]  
(4)

To understand condition (4), it is helpful to consider the underlying tariff changes that could induce a change in the world price while leaving unaltered a country’s local price. Consider the domestic government. If we hold fixed the foreign tariff and increase the domestic tariff, then under our assumptions the world price \( \tilde{p}^w \) falls and the local price \( p \) rises. Thus, a unilateral tariff hike gives the domestic country an improved terms of trade but also changes the domestic local price. We thus cannot understand (4) by considering a domestic tariff change in isolation. Let us then imagine that we increase the domestic tariff and at the same time lower the foreign tariff. This change again leads to a lower world price (i.e., a terms-of-trade gain for the domestic country), but we can calibrate a change of this kind that leaves the domestic local price unaltered. The meaning of (4) is thus that the domestic government values the international income transfer that is implied by an increase in the domestic tariff and a decrease in the foreign tariff that together leave the domestic local price unaltered. An analogous interpretation applies for the foreign government.

The welfare functions presented here are quite general; indeed, governments maximize welfare functions of this form in both the traditional approach and in the leading political-economy approaches to trade policy. Dixit (1987), Johnson (1953-54), Kennan and Reizman (1988) and Mayer (1981) offer important formalizations of the traditional approach. In these formalizations,
governments maximize national income, and the national income of a country improves when it experiences a terms-of-trade improvement. In the political-economy literature, Mayer (1984) shows that, under a representative democracy, the government sets its trade policy to promote the interests of the median voter, whose utility can be represented as a function of this form. As Baldwin (1987) shows, other major approaches to the political economy of trade policy employ government welfare functions of this form as well. Finally, the lobbying models of Grossman and Helpman (1994, 1995) are also included in the framework presented here.

Unilateral Policies We consider now the unilateral policies that governments would select if they were to interact noncooperatively. Suppose, then, that each government sets its tariff policy to maximize its welfare, for any given tariff choice of its trading partner. The associated tariff reaction curves are defined implicitly by

\[ W_p + \lambda W_{\bar{p}w} = 0, \text{ and} \]

\[ W_{p^*} + \lambda^* W_{\bar{p}w}^* = 0, \]

where \( \lambda \equiv [\partial \bar{p}^w / \partial \tau] / [dp/d\tau] < 0 \) and \( \lambda^* \equiv [\partial \bar{p}^w / \partial \tau^*] / [dp^*/d\tau^*] < 0 \). As these expressions highlight, the best-response tariff of each government strikes a balance between the effects on its welfare of the local- and world-price movements induced by its tariff choice.

To gain further insight, let us focus on the domestic government. Starting with an initial tariff pair \((\tau, \tau^*)\), we suppose that the domestic government unilaterally increases its tariff and thus induces a final tariff pair \((\tau^1, \tau^*)\). As noted above, a unilateral hike in the domestic tariff leads to a lower world price \(\bar{p}^w\) and a higher domestic local price \(p\). As suggested by (5), however, we can disentangle this combined change in prices into separate changes in the world and domestic local prices. We do this by imagining that the movement from the initial tariff pair to the final tariff pair is taken in two steps. In the first step, we imagine raising the domestic tariff and lowering the foreign tariff, so as to preserve the domestic local price. As explained above, when tariffs are adjusted in this way, the world price falls. Suppose that we adjust tariffs in this first step until we reach the point at which the world price is reduced to the same level that it achieves at the final tariff pair \((\tau^1, \tau^*)\). At this point, we initiate the second step and raise the domestic and foreign tariffs in a way that preserves the world price. When we adjust tariffs in this fashion, the domestic local price rises. We adjust tariffs in this second step until we reach the final tariff pair \((\tau^1, \tau^*)\).

Across the two steps, we lower the foreign tariff and then raise it back to its initial level, so that the only tariff change that remains in the end is the domestic tariff hike. By breaking the movement into two steps, however, we can isolate the world and local price changes that are identified in (5). In the first step, the domestic local price is unaltered and the world price falls. The effect on

---

\[ \text{See, for example, Olson (1965), Caves (1976), Brock and Magee (1978), Feenstra and Bhagwati (1982), Findlay and Wellisz (1982) and Hillman (1982).} \]
domestic welfare of a change in the world price is captured in (5) with the term \( W_{pw} \). The second step by contrast fixes the world price at its final level and allows the domestic local price to rise from its initial level to its final level. In (5), the effect on domestic welfare of a change in the local price is captured with the term \( W_{p} \). We may interpret (6) for the foreign government similarly.

We now arrive at an important observation. The welfare implications of the local-price movement in the second step are domestic in nature. In particular, they reflect the trade-off for the domestic government between the costs of the induced economic distortions and the benefits of any induced political support. By contrast, the welfare implications of the world-price movement in the first step are international in nature. Specifically, they reflect the benefits to the domestic government of shifting some of the costs of its policy choice onto the foreign government. The cost shifting occurs, since an improvement in the domestic country’s terms of trade corresponds necessarily to a deterioration in the foreign country’s terms of trade.

In a Nash equilibrium, both governments must be on their respective reaction curves, and a Nash equilibrium tariff pair \((\tau^N, \tau^*N)\) thus satisfies (5) and (6). We take this equilibrium to represent the trade-policy decisions that governments would make if there were no trade agreement.

**Trade Agreement** Governments value a trade agreement if it leads to changes in trade policies that generate Pareto improvements for governments relative to the welfare that they experience in the Nash equilibrium. Thus, a trade agreement is potentially valuable if and only if the Nash equilibrium is inefficient, when efficiency is measured relative to government preferences. This discussion motivates further consideration of the Nash equilibrium and its relationship to the efficiency frontier.

Three observations can be stated.\(^8\) The first observation is that Nash tariffs are indeed inefficient. The second observation is that both governments can enjoy welfare gains relative to the Nash equilibrium only if each agrees to set its tariff below its Nash level. The first observation means that a mutually beneficial trade agreement is possible, while the second observation implies that reciprocal trade liberalization is necessary for mutual gains. These first two observations are intuitive and follow from our discussion above. When a government contemplates an increase in its unilateral tariff, it foresees an improvement in its terms of trade; thus, it is in part motivated by the prospect of shifting some of the costs of the tariff hike onto its trading partner, whose terms of trade would deteriorate. The incentive to shift costs naturally leads to inefficient policies. Further, the inefficiency takes a particular form: the possibility of cost shifting leads governments to set tariffs higher than is efficient. Consequently, if both governments are to benefit from a trade agreement, then each must lower its tariff below its Nash level.

As our discussion indicates, the terms-of-trade externality is one reason for the inefficiency of the Nash equilibrium. To see if it is the only reason, we consider a hypothetical world in which governments are not motivated by the terms-of-trade implications of their unilateral trade-policy choices; that is, we consider a hypothetical noncooperative setting in which \( W_{pw} = 0 \) and

\(^8\)Formal proofs of these observations can be found in Bagwell and Staiger (1999a, 2002b).
$W_{pw}^2 \equiv 0$.\(^9\) We therefore define *politically optimal tariffs* as any tariff pair $(\tau^{PO}, \tau^{*PO})$ that satisfies the following two conditions:

$$W_p = 0 \text{ and } W_{p^*}^* = 0.$$  

The key question is whether the politically optimal tariffs are efficient. If the answer is affirmative, then we may conclude that the terms-of-trade externality is the sole rationale for a trade agreement.

The third observation is that the politically optimal tariffs are indeed efficient. To see why, suppose that each government sets its trade policy at its politically optimal level. We note that each government then achieves its preferred local price, given the politically optimal world price. Starting at the political optimum, consider the effects of a small increase in the domestic tariff. The first effect is that the domestic local price rises by a small amount. This effect has no first-order impact on the welfare of the domestic government, however, since the domestic government initially has its preferred local price. The second effect is that the foreign local price falls by a small amount. But the foreign government also initially has its preferred local price; thus, this effect has no first-order impact on the welfare of the foreign government. Finally, the third effect is that the world price falls by a small amount. This terms-of-trade change, however, cannot generate an efficiency gain, since it represents a pure international transfer from the foreign government to the domestic government in the form of tariff revenue. Consequently, starting from the political optimum, a small increase in the domestic tariff cannot generate a Pareto improvement for the governments. The effects of a small increase in the foreign tariff can be similarly analyzed. Thus, once the terms-of-trade motivation is removed from the trade-policy choices of governments, an efficient outcome is achieved, and there is nothing further for a trade agreement to accomplish.

To appreciate the role of our large-country assumption, it is instructive to relax this assumption momentarily and consider an alternative setting in which politically motivated governments preside over small countries. The government of a small country would recognize that it is unable to alter the terms of trade with its trade policy; thus, in this alternative setting, the terms-of-trade motivation is automatically eliminated from the trade-policy decisions of each government. When selecting their unilateral trade policies, the governments of small countries would thus choose politically optimal tariffs; hence, their Nash policies would be efficient. In the leading political-economy models of trade policy, therefore, the governments of small countries have no reason to form a trade agreement among themselves. Our brief detour to the alternative small-country setting thus confirms the general conclusion derived above: the value of a trade agreement is attributable to the terms-of-trade externality that is associated with the trade-policy choices of large countries.

The politically optimal tariffs are of course not the only efficient tariffs. For example, in the special case where governments maximize national welfare, Mayer (1981) shows that efficient tariffs satisfy $\tau = 1/\tau^*$. For this case, the politically optimal tariffs correspond to reciprocal free trade.

---

\(^9\)Importantly, we do not assume here that governments fail to understand the terms-of-trade effects of their tariff choices. Rather, in the context of (5) and (6), we allow that governments understand that $\lambda < 0$ and $\lambda^* < 0$, but we now hypothesize that $W_{p^w} \equiv 0$ and $W_{p^w}^* \equiv 0$. Our approach is to characterize the non-cooperative tariffs that would be selected by governments with these hypothetical preferences and then to evaluate the efficiency of these tariffs with respect to actual government preferences.
and thus rest on the efficiency frontier at the point $\tau = \tau^* = 1$.\textsuperscript{10} A trade agreement enables governments to move from the inefficient Nash tariffs to some point on the contract curve, where the contract curve is that portion of the efficiency frontier on which neither government receives below-Nash welfare. The politically optimal tariffs lie on the contract curve, provided that the countries are not too asymmetric. When the politically optimal tariffs rest on the contract curve, they are focal in the sense that they remedy the terms-of-trade inefficiency in a direct way.

Summarizing, we have argued that the purpose of a trade agreement is to provide governments with an escape from a terms-of-trade driven Prisoners’ Dilemma. This rationale for a trade agreement requires that countries are large, so that they can alter the terms of trade with their trade policies; however, it holds whether or not governments have political/distributional objectives.

**Interpreting the Terms-of-Trade Externality** The discussion above confirms what is at bottom a very simple idea: governments can gain from negotiating a trade agreement, if each would otherwise attempt to shift costs onto the other and as a consequence adopt inefficient unilateral policies. Viewed in this way, the terms-of-trade externality is simply the mechanism through which such cost shifting would occur. As noted in the Introduction, however, many economists are skeptical of the practical relevance of the terms-of-trade theory of trade agreements. We now acknowledge and address two of the main objections to this theory.

The first objection is that the terms-of-trade theory is traditionally advanced under the assumption that governments maximize national income whereas real-world governments have political as well as economic objectives. In fact, we have already addressed this objection above. As we show there, the terms-of-trade rationale for a trade agreement holds as well when governments have political motivations.

The second objection is that the theory does not reflect the way that actual governments think. The theory seems to emphasize abstract general-equilibrium reasoning that might not resonate with practical policy makers, and indeed the “terms of trade” as such are rarely mentioned in trade-policy debates. In response to this objection, we stress that the terms-of-trade theory also may be interpreted in ways which suggest greater practical relevance. First, while we present the theory using a general-equilibrium model above, the theory can also be developed in the context of a partial-equilibrium model. In the partial-equilibrium model, cost shifting occurs through changes in the terms of trade provided that the import tariff is not fully passed through to domestic prices. Intuitively, foreign exporters then bear some of the incidence of the tariff. In this setting, we can thus immediately understand that unilateral tariffs are inefficient, since the domestic government fails to internalize the cost of lost profits that its import tariff imposes on foreign exporters.\textsuperscript{11} Second, real-world trade-policy negotiations are conducted using the language of “market access,” and the terms-of-trade theory can easily be translated into this language. When the domestic government raises

\textsuperscript{10}The politically optimal tariffs generally differ from reciprocal free trade when governments have political/distributional objectives.

\textsuperscript{11}In Bagwell and Staiger (2001b), we employ a partial-equilibrium model, derive the three observations mentioned above, and further develop this interpretation.
its import tariff and thereby shifts in its import demand curve, the resulting “price effect” under which the domestic country enjoys a terms-of-trade improvement is accompanied by a “volume effect” under which the foreign country experiences a reduction in access to the domestic market. Once this link between price and volume effects is forged, the terms-of-trade theory can be developed using the market-access language that trade-policy negotiators adopt.12

2.1.2 Other International Externalities

Our discussion thus far has proceeded from a backdrop in which markets are assumed to be perfectly competitive. The perfectly competitive paradigm offers a valuable benchmark for understanding the purpose of trade agreements, but in many markets firms possess market power. An extensive literature has established that imperfectly competitive markets can give rise to “profit-shifting” and “firm-delocation” effects that provide novel motives for unilateral trade-policy intervention (see, for example, the integrative treatment of this literature in Helpman and Krugman, 1989). This suggests that other international externalities in addition to the terms-of-trade externality might arise in markets with imperfect competition, raising the possibility as well that new rationales for trade agreements beyond the terms-of-trade rationale might then be identified in these settings.

Ossa (2009) explores the role of trade agreements in imperfectly competitive environments. He considers a monopolistically competitive setting of two-way trade in similar products, in which firms produce differentiated products under conditions of free entry and compete for sales both in the home market and abroad, and where exporting the product abroad involves shipping costs. In such an environment, Venables (1987) has shown that a firm-delocation motive for trade policy arises: if the domestic country imposes a tariff on its imports or offers a subsidy to its exports, then foreign firms can be “delocated” to the home market, and domestic consumers save on trade costs and enjoy a lower overall price index; the domestic consumers’ gain, however, comes at the expense of foreign consumers, whose price index rises. Importantly, Ossa shows that the firm-delocation effect represents a beggar-my-neighbor policy that does not travel through the terms-of-trade externality.

In terms of the representation of welfare introduced above, the novel international externality identified by Ossa (2009) takes the form of a local-price externality: in each government’s welfare function, the local price in the other country enters directly, in addition to its own local price and the world prices.13 Intuitively, each government now cares directly about the local price in the other country’s market, because each government would rather have more of the world’s firms (and the production of their individual varieties) located locally rather than abroad to save on transport costs, and the equilibrium pattern of firm location across countries depends on local prices in both countries via the free-entry condition.

As it turns out, this local-price externality arises in other imperfectly competitive environments as well. In particular, Bagwell and Staiger (2009b) explore imperfectly competitive environments

---

12 For a formal definition of market access and further development of the relationship between the terms-of-trade theory and the language of market access, see Bagwell and Staiger (2002b).

13 For a development of this representation in the firm-delocation model, see Bagwell and Staiger (2009a).
in which entry is not free and monopoly/oligopoly profits exist, and where the well-known (Brander
and Spencer, 1981, 1984a,b, 1985) profit-shifting motive for trade-policy intervention arises. Here
as well, governments care directly about the local prices in the markets of their trading partners, in
addition to their own local prices and the world prices. In this case, the reason is that governments
care about the profits of their firms, and firm profits depend in part on the volume of export sales
which is in turn influenced by local prices in the export destination market.

Hence, in the presence of imperfect competition, new international externalities can indeed arise;
in addition to the terms-of-trade externality that travels through the world price, there are also
local-price externalities that travel through domestic and foreign local prices when firm-delocation
or profit-shifting effects are present. This implies that the international policy environment is more
complex than in the case of perfectly competitive markets, and it raises the possibility that the
task of a trade agreement may be more complicated in this environment as a result. To assess this
possibility, however, we must ask whether governments would make unilateral policy choices that
internalize these international externalities in an appropriate fashion from a world-wide perspective,
whatever form these externalities might take, and if not, why not.

It is readily established and unsurprising that unilateral policy choices in this environment are
indeed inefficient. After all, the terms-of-trade (world-price) externality that drives the inefficiency
in the competitive setting is still present here, and now there are additional (local-price) externalities
as well. But when the question of what accounts for the inefficiency is posed as we posed this
question above, and the efficiency properties of politically optimal tariffs are evaluated, a surprising
answer emerges: the sole rationale for a trade agreement in this environment is again to remedy
the inefficiency attributable to the terms-of-trade externality, the same rationale that arises in
perfectly competitive markets. More specifically, if governments could be induced to not value
the pure international rent-shifting associated with the terms-of-trade movements caused by their
unilateral tariff choices, then their tariff choices would be efficient and there would be nothing left
for a trade agreement to do. Intuitively, when governments adopt politically optimal tariffs, they
are not motivated to impose terms-of-trade externalities on each other, and they are thus freed
to use their trade policies to achieve their preferred local prices, the achievement of which then
ensures that the international externalities associated with local-price movements are eliminated
(to the first-order) as well.

It is worth pausing to reflect on this last point. As we have observed, in the presence of imper-
fectly competitive markets, local-price externalities can arise; however, at the political optimum,
where governments would be led if they did not value the world-price externalities, the local-price
externalities disappear and the outcome is efficient. Still, in general these local-price externalities
would be operative away from the political optimum, and they may therefore represent an im-
portant feature for a trade agreement to address, if the agreement is not designed to deliver the

\[14\] A related exploration is contained in Bagwell and Staiger (2002b, Ch. 9), though the representation of welfare
as exhibiting a local-price externality is not developed there. See also the recent analysis of Mrazova (2009).
\[15\] This result is established for the firm-delocation model in Bagwell and Staiger (2009a) and for the profit-shifting
model in Bagwell and Staiger (2009b).
political optimum. For example, as Ossa’s (2009) findings demonstrate, and as discussed further in Bagwell and Staiger (2009a), if governments were constrained in their use of export policies the political optimum could not in general be reached, and in this case the local-price externalities become an important component of the problem that governments must solve in this environment. This observation takes on special relevance in the present context, as the GATT/WTO restricts the use of export subsidies. In this light Ossa’s findings can be interpreted as characterizing a problem that arises when export subsidies are banned. At the same time, this interpretation falls short of delivering a fundamental rationale for a trade agreement, because it appeals to the existence of a trade agreement (on export subsidies) to explain the purpose of a trade agreement.

Finally, we note that in both the competitive paradigm and in each of the imperfectly competitive environments that we have discussed, international (world) prices are determined by anonymous market-clearing conditions. But for many international transactions, the international prices at which the associated goods or services change hands are determined as a result of bilateral bargaining between the domestic purchaser and the foreign supplier. When international prices are determined in this way, Antras and Staiger (2008) show that the mechanism by which countries shift the cost of their policy intervention on to trading partners becomes more complex, and the possibility of a novel “political externality” along the lines suggested by Ethier (2004) may arise and pose an independent problem for a trade agreement to solve. In light of the broad set of market structures in which beggar-my-neighbor problems can be given a terms-of-trade interpretation, the finding of a novel beggar-my-neighbor problem that can arise under certain conditions and cannot be given a terms-of-trade interpretation – and which suggests a distinct rationale for a trade agreement – is significant and warrants further attention, but the literature has not yet gone beyond the identification of this possibility.

2.2 Addressing Beggar-Myself Policies

Thus far we have focused on theories of trade agreements that emphasize the control of beggar-my-neighbor motives. A distinct though possibly complementary approach to the theory of trade agreements can be developed if it is posited that the purpose of a trade agreement is to tie the hands of its member governments against private agents in the economy, and thereby offer an external commitment device. This approach has been formalized in a number of papers (see, for example, Carmichael, 1987, Staiger and Tabellini, 1987, Matsuyama, 1990, Brainard, 1994, Mitra, 2002 and Maggi and Rodriguez-Clare, 1998, 2007, to name a few). To describe the main ideas, we focus on the papers of Maggi and Rodriguez-Clare.

Maggi and Rodriguez-Clare (1998) adopt a small-country perspective so that the terms-of-trade argument for trade agreements is absent, and they focus instead on the possibility that an anticipated trade-policy-lobbying relationship between the government of this small country and producers in one of its sectors could distort the equilibrium allocation of resources in the economy.

\[\text{Regan (2006) also articulates a theory of trade agreements that seems to fit with this approach, although his theory is somewhat informal and hence more difficult to categorize with confidence.}\]
toward the sector with the active lobby. Working with the lobbying model of Grossman and Helpman (1994), Maggi and Rodriguez-Clare confirm that the government will be compensated by the lobby for the ex-post distortions its trade-policy choice imposes on the economy, that is, the distortions given the sectoral allocation of the economy’s resources that are sunk at the time this choice is made. But the lobby will not compensate the government for the ex-ante distortions in the sectoral allocation of resources created by the anticipation of the government’s relationship with the lobby, and this provides an opening for the government to wish to tie its hands ex ante against the possibility of being influenced by ex-post lobbying. As Maggi and Rodriguez-Clare demonstrate, a possible commitment role for a trade agreement is thereby identified.

In Maggi and Rodríguez-Clare (1998), the possibility of lobbying ex ante (i.e., at the time that the decision to form a trade agreement is made) is not considered. In a follow-up paper, Maggi and Rodríguez-Clare (2007) allow for this possibility, and develop a hybrid model that combines both terms-of-trade and commitment arguments for a trade agreement. As they demonstrate, the two motives for a trade agreement can interact in non-trivial ways and generate a number of interesting empirical predictions. For example, the magnitude of trade liberalization delivered by a trade agreement is expected to rise in the degree to which resources are mobile in and out of the sector ex post. Intuitively, when resources are mobile ex post, the lobby correctly anticipates that the rents from trade protection for its members will be low (they will be dissipated by entry into the sector), and so the lobby is not willing ex ante to engage intensively against the formation of a trade agreement that commits the government to low tariffs. And Maggi and Rodríguez-Clare also demonstrate that their hybrid approach can help account for the particular form (ceilings) that tariff bindings take in the GATT/WTO.\footnote{Using models that feature the terms-of-trade externality, Bagwell and Staiger (2005b), Bagwell (2009) and Horn, Maggi and Staiger (forthcoming) also provide interpretations of the use of tariff ceilings in GATT/WTO.}

There is some empirical evidence that commitments made in trade agreements may play this kind of role (see, for example, Staiger and Tabellini, 1999, and Tang and Wei, forthcoming). Nevertheless, as we noted in the Introduction, while the potential commitment role of trade agreements has been identified and formalized in the literature, for the most part the literature adopting this approach has not attempted to understand and interpret the key design features of the GATT/WTO from this perspective (though Maggi and Rodríguez-Clare, 2007, takes an important step in this direction). For this reason, we do not focus on theories that view the purpose of trade agreements as addressing beggar-myself policies for the remainder of our Review.

3 Terms-of-Trade Manipulation: The Evidence

What is the purpose of a trade agreement? As our discussion in the previous section has confirmed, in a wide variety of formal settings the fundamental purpose of a trade agreement is to provide an escape from a terms-of-trade driven Prisoners’ Dilemma. Rarely in economics do so many distinct models answer in unison to a single question, and the robust structure of the problem that is implied by these models offers some hope that an institution with some simple design features
might effectively aid governments in their efforts to find a solution. But there is also an alternative view: these models may answer in unison, but the answer they provide is simply wrong, or at least irrelevant to understanding the central features of trade agreements in general and the GATT/WTO in particular. This, of course, is an empirical issue, and in this section we survey a growing body of empirical work that provides evidence relating to the terms-of-trade argument. We focus here on papers that relate to three specific questions which we pose below, and which are directly related to the essential features of the terms-of-trade argument. We postpone until section 5 a discussion of a wider body of empirical work that relates to various predictions of the terms-of-trade theory.

A first question is whether there is evidence that a country’s tariffs can affect its terms of trade, a clear pre-condition for the empirical relevance of the terms-of-trade theory. Studies by Kreinin (1961), Winters and Chang (2000, 2002), Anderson and VanWincoop (2002) and Bown and Crowley (2006), among others, offer compelling evidence that unilateral tariff changes can significantly affect a country’s terms of trade, even for apparently “small” countries such as Mexico. Moreover, recalling that the terms-of-trade effects of a tariff arise whenever the incidence of the tariff is not fully passed through to domestic prices, the large body of empirical work on exchange rate pass-through surveyed by Goldberg and Knetter (1997) – when combined with the particular empirical findings in Feenstra (1995) that the pass-through associated with exchange rate shocks can be thought of as comparable in magnitude to the pass-through associated with tariff changes – suggests that the terms-of-trade effects of tariff changes are likely to be quantitatively significant and widespread across countries.18

A second question is whether, when governments set their trade policies unilaterally (and hence noncooperatively), they respond to terms-of-trade motives in the way that the theory predicts. According to a recent paper by Broda, Limao and Weinstein (2008), the answer is yes. Focusing on 15 countries that were never GATT members, and that hence set their tariffs in a unilateral fashion prior to joining the WTO, Broda et. al. first estimate the degree of market power that each of these countries was able to exert on the foreign export (world) prices that it faced (as captured by the foreign export supply elasticities faced by these countries). With these estimates, they confirm the answer to the first question posed above that other studies have also provided: most countries, even apparently “small” countries, have significant ability to alter their terms of trade on many imported products with their tariff choices. Broda et. al. then relate this measure of the power to affect world prices to the unilateral tariff choices that each country made over this period. They find that, prior to joining the WTO, these countries on average set tariffs 9 percentage points higher on imports for which they could exert large effects on world prices as compared to imports where their ability to affect world prices was limited, an impact that is very sizable (for example, it is roughly comparable to the size of the average tariffs in these countries). Moreover, they find that this terms-of-trade motive explains more of the cross-industry variation in tariffs than is explained by commonly used political-economy variables. Hence, according to the Broda

18 A more complete survey of the empirical literature relating to this first question is contained in Bagwell and Staiger (2002b, Ch. 11).
et. al. findings, when governments set their trade policies unilaterally and noncooperatively, they respond to terms-of-trade motives strongly and in the way that the theory predicts.\footnote{These findings are reinforced by the recent paper of Dhingra (2009), who finds cross-country empirical support for the median-voter model of tariff determination, but only once terms-of-trade motives are taken into account.}

A third question is whether the tariff cuts negotiated in the GATT/WTO actually reflect the removal of that portion of the noncooperative tariff that embodies the terms-of-trade motive. One way to answer this question is to check whether measures of the power to affect world prices help to predict the levels of noncooperative trade policies but do not help predict the levels of tariffs bound as a result of GATT/WTO negotiations. This is the approach taken by Broda et. al. (2008). Focusing on the United States, they find that U.S. non-tariff barriers and so-called “statutory tariff rates” – which have not been subjected to direct negotiations within the GATT/WTO – are significantly and positively related to the degree of market power which the United States exerts on the world prices of its import products, while the U.S. MFN tariffs – which have been the subject of GATT/WTO negotiations – exhibit no such relationship. A different approach to this question is taken by Bagwell and Staiger (2009c). In that paper, the terms-of-trade theory is used to derive an expression for the component of the noncooperative tariff that embodies the international cost-shifting motive, and this expression is in turn used to predict negotiated tariff levels based on pre-negotiation tariff levels, import volumes and prices, and measures of the power to affect world prices. The implied pattern of negotiated tariff cuts is then confronted with data from the accession negotiations of 16 countries that joined the WTO subsequent to its creation in 1995, and strong and robust support for the predictions of the terms-of-trade theory are found in the observed pattern of negotiated tariff concessions. When viewed together, these two papers paint a reinforcing picture of an emerging message: there is increasing evidence consistent with the view that the tariff cuts negotiated in the GATT/WTO reflect the removal of trade protection that is motivated by cost-shifting incentives, as the terms-of-trade theory predicts.

Finally, we note that the empirical relevance of the terms-of-trade theory does not hinge on all countries being large enough to affect world prices in all products. Instead it simply suggests that the large players in the market should be the most active participants in any particular negotiation. When viewed from this perspective, the empirical pattern that is documented in the studies surveyed above – that most countries have the power to affect their terms of trade in some products, while some countries have the power to affect their terms of trade in most products – seems broadly consistent with the record of tariff bargaining in the GATT/WTO, namely, that most of the participation in tariff negotiations has come from the large industrial countries.

4 The GATT Architecture and its Historical Antecedents

GATT arose in response to the protectionist outbreak of the 1920s and 1930s, which culminated in the U.S. Smoot-Hawley Tariff Act of 1930 and the spate of retaliatory tariffs that followed. From the perspective of our discussion above, we may think of the Nash equilibrium as corresponding to the “tariff war” that is associated with the Smoot-Hawley tariffs. The challenge that governments
faced was then to find some means to reach a more cooperative trade-policy relationship, and thereby move to the contract curve.

4.1 Historical Antecedents

During the decade following World War I, the United States was involved in various multilateral bargaining attempts to address the problem of high and rising tariffs, each largely unsuccessful (Tasca, 1938, p 7). These repeated failures led to a conscious decision on the part of the United States to abandon multilateral tariff bargaining and experiment with bilateral bargaining under the 1934 Reciprocal Trade Agreements Act (RTAA). The RTAA marked the first time that the United States combined bilateral tariff bargaining with unconditional MFN, according to which exports from each country with whom the United States had an agreement under the RTAA would automatically receive the lowest (“most favored nation”) tariff rate that the United States offered to any exporting country.\(^{20}\) Though this approach was novel for the United States, similar approaches had been tried in Europe decades earlier (Tasca, 1938, p. 135), and the design and implementation of the RTAA built on lessons learned from the European experience with bilateral tariff bargaining in at least two important ways.

First, the European experience taught the important lesson that a country’s current bargaining partners would require assurance that any future bilateral deals that it struck with other countries would not erode the value of the concessions being granted, and that the most practical way to provide assurance against such “concession erosion” was with a promise of unconditional MFN (Wallace, 1933, p. 629). The promise of unconditional MFN was included in the RTAA in part to address the concession erosion issue.\(^{21}\) Second, the European experience provided a vivid illustration of the perverse incentive to raise tariffs on the eve of bargaining – and thereby adopt so-called “bargaining tariffs” – to better position oneself for the negotiations to follow (Wallace, 1933, p. 630). As Tasca (1938, p. 179, note 34) observes, this experience provided a second lesson for the United States, which for the purpose of measuring tariff concessions in its bilateral negotiations adopted the strategy of measuring all tariffs, on both sides, at a fixed pre-negotiation date.

These lessons may have helped the United States avoid the twin problems associated with concession erosion and bargaining tariffs that plagued the European efforts before it, and may therefore help explain why bilateral tariff bargaining under the RTAA turned out to be far more successful. But while in the European experience the issue of bargaining tariffs amounted to the unilateral positioning of pre-negotiation tariffs, under the RTAA the analogous issue became how to design bilateral agreements with early negotiating partners to best preserve bargaining power for later agreements with other negotiating partners. This task was made difficult by the unconditional MFN requirement, which automatically granted “for free” to other potential bargaining partners

\(^{20}\) As Tasca (1938, pp. 116-121) describes, the United States had, since 1922, adopted an unconditional MFN approach, but maintained an “autonomous” (i.e., unilateral) tariff up until the RTAA.

\(^{21}\) There were also wider arguments for adopting a policy of unconditional MFN, including the perceived “multilateralization” benefits and a reduction in the risk of war emphasized by U.S. Secretary of State Cordell Hull at the time (see, for example, Culbert, 1987 and Rhodes, 1993).
any tariff concessions granted to early negotiating partners.

The preservation of bargaining power for later negotiations became a major preoccupation of the United States under the RTAA. Beckett (1941, p. 23) and Tasca (1938, pp. 146-147) discuss the tactics used by the United States under the RTAA in this regard. In effect, by granting tariff concessions to a negotiating partner only on those products for which the partner was the principal supplier, possibly combined with product reclassification for tariff purposes to heighten the dominance of the partner in these products, and by splitting the concession into a sequence of partial tariff reductions negotiated with different countries in successive agreements, it was thought that much of the free-rider potential created by unconditional MFN could be eliminated. And where free-riding remained a substantial possibility, two additional tactics were available: attempting to engage groups of countries in simultaneous negotiations; and threatening to withdraw or modify the earlier agreement if free-riding continued.

Summarizing, tariff bargaining under the RTAA exhibited a number of central features: the approach to tariff bargaining was decidedly bilateral, and was chosen only after governments had considered, attempted, and ultimately rejected multilateral tariff bargaining; prior European experience with concession erosion and bargaining tariffs were at the heart of prominent issues that influenced the design and implementation of the RTAA along important dimensions; and unconditional MFN, the principal-supplier rule, split concessions and withdrawal/Modification clauses were understood to be central to the operation of reciprocal tariff bargaining under the RTAA.

Between 1934 and 1947, the United States successfully concluded separate bilateral agreements with 29 countries under the RTAA. Encouraged by this success, the United States pushed for the establishment of a multilateral institution built upon the key components of the RTAA, and in 1946 formal negotiations began for the creation of an International Trade Organization (ITO). In 1947, GATT was negotiated and was intended to serve as an interim agreement, but the ITO was never ratified by the U.S. Congress.

The objectives of the member governments in creating GATT are described in its Preamble, and include “expanding the production and exchange of goods.” The Preamble also states the governments’ belief that “reciprocal and mutually advantageous arrangements directed to the substantial reduction in tariffs and other barriers to trade and to the elimination of discriminatory treatment in international commerce” would contribute toward these objectives.

In total, there were eight rounds of GATT negotiations that together spanned almost 50 years. In the earlier rounds, the primary focus was the reduction of import tariffs on goods. In the final GATT round, known as the Uruguay Round, governments took on a number of new issues (e.g., investment and intellectual property) and formed the WTO. The WTO has sponsored a ninth round, the Doha Round, that is still ongoing.

### 4.2 Architecture

Membership in the GATT/WTO carries with it an obligation to abide by certain rules. In GATT, these rules were contained in a set of 39 articles. These GATT articles have been incorporated into
the WTO, and the WTO has also extended the principles embodied in them to a variety of new issues. Here, we provide a brief overview of the principles embodied in these articles.

GATT articles can be sorted into three broad categories: substantive obligations, exceptions to those obligations, and dispute settlement procedures. The substantive obligations of a GATT/WTO member relate to tariff commitments, MFN treatment and a general “code of conduct” in the international-trade arena. These provisions oblige member-governments to use tariffs rather than non-tariff barriers as protective measures, to apply them on a non-discriminatory basis to other members, and to honor any tariff bindings made in a GATT/WTO negotiation.

At the same time, the GATT/WTO also provides for a variety of exceptions to these obligations. Some exceptions permit a government to suspend an obligation under certain conditions or to withdraw a previous concession through renegotiation. The logic of including exceptions for such “original” actions is that a government is more likely to agree to a tariff commitment if it knows that the legal system has “safeguards” allowing its concessions to be modified or withdrawn under appropriate conditions. But a tariff commitment would lose its meaning if exceptions for original actions were not disciplined in some way, and so GATT/WTO rules also permit exceptions for “retaliatory” actions. Specifically, if a government modifies or withdraws a previous concession, then GATT/WTO rules recognize that a cost may be borne by its trading partner; and if the government fails to compensate the trading partner in an amount that is acceptable to the trading partner, then the partner is allowed to achieve compensation through retaliation. The meaning of retaliation is that the trading partner can withdraw a concession of a “substantially equivalent” nature.

We now come to the third element mentioned above: the GATT/WTO dispute settlement procedures. Here a key issue is the determination of whether the actions by one country “nullify or impair” the benefits that another country expects under the agreement. This occurs when actions are taken by one country which “...harmed the trade of another, and which ‘could not reasonably have been anticipated’ by the other at the time it negotiated for a concession” (Jackson, 1997, p. 115). In the typical complaint, a country is alleged to have failed to comply with its GATT/WTO obligations, as when it imposes quantitative restrictions or violates MFN.22

Every GATT/WTO dispute begins with a consultation phase among the involved parties. Resolution of the dispute may be (and often is) achieved in this first phase. Otherwise, a second phase is initiated in which a GATT/WTO panel (or Appellate Body) conducts an investigation and issues a ruling and recommendation. If the panel finds that nullification or impairment has occurred, then it recommends that the offending country correct any illegal measures, and resolution of the dispute may occur at this point. The offending country may be unwilling to comply with the panel’s recommendation, however, and in this case it might seek a negotiated resolution by offering to compensate the harmed country, perhaps in the form of an MFN tariff reduction on some other good. But if compensation is not offered, or if it is offered but rejected, the harmed

---

22See Hudec (1993) and Bown (2002) for a description of the trade disputes that occurred under GATT, and see Horn and Mavroidis (2001-2007) for legal and economic analyses of the major disputes under the WTO.
country may then retaliate with an authorized and discriminatory suspension of tariff concessions. The number of authorized retaliations has in practice been small, but as Rhodes (1993, p. 109) observes, the threat of authorized retaliation is often the catalyst for resolution in the earlier phases of the dispute. As our discussion in this section confirms, MFN is a pillar of the GATT/WTO architecture, and the enforcement provisions of the GATT/WTO are elaborately developed. The representation of reciprocity in the GATT/WTO, however, may be less apparent from this discussion. We therefore comment briefly on the role of reciprocity in the GATT/WTO.

The GATT/WTO principle of reciprocity refers to the ideal of mutual changes in trade policy which bring about changes in the volume of each country’s imports that are of equal value to changes in the volume of its exports. In our discussion above, the notion of reciprocity arises in two places. First, as we have observed, governments negotiate in GATT/WTO rounds with the stated goal of obtaining mutually advantageous arrangements through reciprocal reductions in tariff bindings: in this context, it is often observed that governments approach negotiations seeking a “balance of concessions,” so that there is a rough equivalence between the market access value of the tariff cuts offered by one government and the concessions won from its trading partner. Second, when a government seeks to renegotiate and modifies or withdraws a previous concession as an original action, and more generally whenever a government takes an action which nullifies or impairs the benefits expected under the agreement by another government, GATT/WTO rules permit affected trading partners to withdraw “substantially equivalent concessions,” and thereby to retaliate in a reciprocal manner.

5 The Design and Practice of the GATT/WTO: A View from the Terms-of-Trade Theory

We now turn to the fourth and final step in our Review. Here we draw on the recent theoretical and empirical literature to interpret the design and practice of the GATT/WTO. We begin with the principle of reciprocity.

5.1 Reciprocity

We have suggested above that the central problem faced by governments as they considered the design of the GATT/WTO can be given a simple interpretation: terms-of-trade manipulation. From this vantage point, we may now pose the question: Why would the principle of reciprocity be attractive to governments? The answer is that reciprocity describes a fixed-terms-of-trade rule to which mutual tariff changes must conform; and in an environment where terms-of-trade manipulation is the problem to be fixed, a fixed-terms-of-trade rule is bound to be attractive.

23 Under GATT, retaliation was authorized only once. Under the WTO, retaliation has been authorized and used in a number of cases. See Mavroidis (2000) and WTO (2001, p. 28) for further discussion.
We now explain these points more fully. We first propose a formal definition of reciprocity. Suppose that, beginning from an initial pair of tariffs, \((\tau^0, \tau^*0)\), a tariff negotiation results in a change to the new pair of tariffs, \((\tau^1, \tau^*1)\). Denoting the initial world and domestic local prices as \(p^w0\) and \(p^0\), and the new world and domestic local prices as \(p^w1\) and \(p^1\), we say that the tariff changes conform to the principle of reciprocity provided that

\[
\bar{p}^w0[M_x(p^1, \bar{p}^w1) - M_x(p^0, \bar{p}^w0)] = [E_y(p^1, \bar{p}^w1) - E_y(p^0, \bar{p}^w0)],
\]

where changes in trade volumes are valued at the existing world price. We next use the balanced trade condition (1) – which must hold at both the initial tariffs and the new tariffs – to establish that (7) may be rewritten as

\[
[p^w1 - p^w0]M_x(p^1, \bar{p}^w1) = 0.
\]

According to (8), reciprocity can be given a simple and striking characterization: mutual changes in trade policy conform to the principle of reciprocity if and only if they leave the world price unchanged. With this characterization in hand, we now consider the application of reciprocity within GATT/WTO practice.

As we have noted above, a first application of reciprocity refers to the balance of concessions that governments seek through a negotiated agreement. Such behavior is hard to reconcile with the standard economic argument that unilateral free trade is in a country’s best interests. The terms-of-trade theory, however, offers a simple interpretation of this application of reciprocity.

To see this, suppose that governments begin with Nash tariffs, and observe that, at the Nash point, we may use (4), (5) and (6) to conclude that \(W_p^* < 0 < W_p\). Next observe that, if governments were to reduce tariffs in a reciprocal manner, then according to (8) the world price would be preserved, while the domestic local price \(p\) would fall and the foreign local price \(p^*\) would rise. But this means that both the domestic-government welfare and the foreign-government welfare would then rise (since \(W_p < 0\) and \(W_p^* > 0\)). Evidently, the structure of international cost-shifting implies that, beginning from their Nash tariff choices, both governments would desire tariff liberalization and the implied greater trade volume if this could be achieved without a decline in the terms of trade. The principle of reciprocity harnesses this desire, and so activates efficiency-enhancing tariff-liberalizing forces in this environment.

As we have noted, a second application of reciprocity in the GATT/WTO arises whenever a government takes an action which nullifies or impairs the benefits expected under the agreement by another government. An important instance of this second application concerns the rules that govern the process by which tariff bindings may be renegotiated. Under GATT Article XXVIII,

\[24\] See Bagwell and Staiger (1999a) where these results are derived formally.

\[25\] This defines reciprocity from the perspective of the domestic country. But in our two-country setting, tariff changes conform to reciprocity for the domestic country if and only if they conform to reciprocity for the foreign country as well.
a country may propose to modify or withdraw a concession agreed upon in a previous round of negotiation, and may offer some compensation to its trading partner in return. But if the country and its trading partner are unable to reach agreement on the appropriate level of compensation, the country is permitted to implement its proposed change anyway, with the understanding that the trading partner may then reciprocate. Here, the principle of reciprocity is used to moderate the response of the trading partner, who is allowed to withdraw substantially equivalent concessions of its own.

In light of the explicit provisions governing renegotiation, market-access negotiations in the GATT/WTO may be viewed as a multi-stage game. Governments first agree to bind their tariffs at specific levels in a round of negotiations. And then each government considers whether it would prefer to raise its tariff above the bound level, anticipating that the outcome of any renegotiation will conform to reciprocity, preserving the balance struck by the original negotiations and thus preserving the world price. In this setting, a pair of initially negotiated tariffs will be renegotiated if, at the initial tariff pair, either government desires less trade volume and the associated change in its local price at the fixed world price. This suggests that many points on the efficiency frontier could be susceptible to renegotiation in this setting. In fact, there is only one efficient tariff pair that is robust to the possibility of renegotiation under the reciprocity rule, and that is the politically optimal tariff pair. If governments were to negotiate to the efficient political optimum, where each government has achieved its preferred local price, then neither government would have any desire to deviate from this point if in so doing it could not alter the terms of trade in its favor. The principle of reciprocity in this second application can be understood to harness this feature, and so to create an environment where the efficient political optimum, once achieved, remains a robust and stable outcome.

This discussion also reveals a broader point: in effect, governments are “penalized” under the GATT/WTO reciprocity rule if they attempt to negotiate an efficient tariff pair other than the political optimum. Suppose, for example, that the domestic government is successful in pushing the initial negotiations to a point on the efficiency frontier which is more favorable to it than the political optimum. At the political optimum, the domestic government’s welfare cannot be enhanced with a change in its local price, so at this alternative preferred point the terms of trade must be more favorable for the domestic government than the terms of trade at the political optimum; and by implication, the terms of trade must then be less favorable for the foreign government at this alternative point relative to the political optimum. Notice, though, that if the initial negotiations lead to this alternative point, then at the fixed but less-favorable terms of trade the foreign government will naturally want less trade volume (and the associated change in its local price) as compared to the political optimum, and the foreign government can achieve this through renegotiation. Therefore, some of the benefit to the domestic government of pushing the initial negotiating point away from the political optimum would be given up in the subsequent renegotiation. And as a result, the domestic government may be less eager to push negotiations away from the political optimum in the first place. As this example illustrates, the reciprocity
rule can help to mitigate the power asymmetries that governments might otherwise display at the bargaining table. In this way, it encourages governments to select the “rules-based” politically optimal tariffs.

Recent empirical evidence confirms the importance of the first GATT/WTO application of reciprocity discussed above. For example, Shirono (2004) finds that the tariff cuts agreed to in the Uruguay Round of GATT negotiations conformed closely to the reciprocity norm, and that the economic significance of the terms-of-trade changes induced by these tariff cuts were quite limited. Limao (2006, 2007) also finds evidence consistent with reciprocity. In particular, focusing on U.S. tariff cuts in the Uruguay Round and constructing a measure of market-access concessions while instrumenting to address the potential endogeneity issues, Limao reports that a decrease in the tariff of a U.S. trading partner that exports a given product leads to a decrease in the U.S. tariff on that product, and that a significant determinant of cross–product variation in U.S. tariff liberalization is the degree to which the United States received reciprocal market-access concessions from the corresponding exporting countries. Finally, Karacaovali and Limao (2008) perform a similar exercise for the EU tariff cutting behavior in the Uruguay Round, and find analogous support for the importance of reciprocity in explaining the pattern of EU tariff cuts: EU tariff reductions were largest for those products exported by countries who themselves granted large reductions in tariffs. While more evidence is needed before the issue is settled, as an empirical matter it does appear that actual tariff bargaining outcomes in the GATT/WTO conform to a reciprocity norm.

Regarding the second application of reciprocity in the GATT/WTO, the empirical questions at issue are somewhat different. It is clear as a legal matter that in this application reciprocity in the GATT/WTO is defined by equivalent trade effects, much as the definition in (7) indicates, and that in practice this concept guides the permissible response to nullification or impairment in the GATT/WTO.26 What is less clear is whether the principle of reciprocity in this application serves to guide governments toward a rules-based outcome such as the political optimum. On the one hand, some features of the political optimum seem to be present in GATT/WTO conventions and bargaining outcomes: the principal-supplier rule creates a presumption that small players in a market will typically not be asked by their trading partners to make significant market-access concessions in that market; special and differential treatment clauses exempt many small countries from a host of other GATT/WTO obligations to which other countries must conform; most of the significant market-access concessions have been made by the large industrialized countries; and even where a country that is small in a given market accepts on paper obligations that apply to that market, the GATT/WTO enforcement procedures operate “on demand,” and so a small player in a market can likely expect to be able to violate obligations in that market without bringing retaliation upon itself in any event. On the other hand, there are many accounts from diplomats

26See, for example, the Appellate Body Opinion in WTO (2004) on the trade-effects interpretation of reciprocity. While it is clear that the trade-effects concept of reciprocity applies in this context, there is still the implementation issue of how trade effects are assessed, and a remaining question is whether the trade-effects approach can be implemented in a meaningful way (see, for example, Spamann, 2006).
and delegates of small and developing countries that run counter to this view.\footnote{See, for example, Jawara and Kwa (2003) and the review of Jawara and Kwa in Staiger (2006).} In any event, this is a critical question, both for the theory and for the policy debate surrounding the performance of the WTO. And to date we are unaware of any empirical work that directly addresses it.

\subsection{5.2 MFN}

We next turn to the nondiscrimination principle, embodied in the GATT/WTO rule that requires all tariffs to be applied on an MFN basis to the trade of other member countries. To begin, we describe a 3-country extension of the benchmark general-equilibrium model developed in section 2.1.1.

The domestic country now exports good $y$ to two foreign countries, denoted by the superscripts ‘*1’ and ‘*2,’ and imports good $x$ from each of these countries (who do not trade with each other). Each foreign country can impose a tariff on its imports of good $y$ from the domestic country (we denote the tariff of foreign-country $i$ by $\tau^i$), while the domestic country can set tariffs on its imports of good $x$ from the two foreign countries. Notice that if the domestic country applies the tariff $\tau^1$ to imports from foreign-country 1 and the discriminatory tariff $\tau^2 \neq \tau^1$ to imports from foreign-country 2, then separate world prices $p^w1$ and $p^w2$ apply to its trade with foreign-countries 1 and 2 respectively: this follows because there can only be one local price in the domestic economy, and the pricing relationships $p = \tau^1 p^w1$ and $p = \tau^2 p^w2$ then imply $p^w1 \neq p^w2$ whenever $\tau^1 \neq \tau^2$.

In this setting, the MFN rule imposes a very simple requirement: the domestic country must apply a common tariff level $\tau^1 = \tau^2 = \tau$ to the imports of $x$, regardless of whether these imports originate from foreign-country 1 or 2. An important implication of the MFN rule is then that a single equilibrium world price, $\bar{p}^w(\tau, \tau^1, \tau^2)$, must prevail; consequently, we may continue to express government preferences with the simple representation $W(p, \bar{p}^w)$, $W^1(p^1, \bar{p}^w)$ and $W^2(p^2, \bar{p}^w)$, the same representation that we used in the 2-country setting.

At a basic level, the appeal of the MFN principle can already be appreciated: in a multilateral world, the MFN principle ensures that the international externality at the root of the problem to be solved continues to exhibit the same structure as in the simpler 2-country setting. This suggests in turn that, in the company of MFN, the attractive properties of reciprocity described above might extend to a multilateral setting. This is indeed the case. Under the MFN principle and beginning from noncooperative tariffs, each country can gain from reciprocal liberalization; and MFN politically optimal tariffs (which are efficient) are robust to renegotiation under the reciprocity rule. And in light of our earlier discussion, in a multilateral world MFN and reciprocity together may be understood as encouraging governments to select the politically optimal tariffs, thereby mitigating the power asymmetries that governments might otherwise wield at the bargaining table.\footnote{These results are established in Bagwell and Staiger (1999a), where it is also established that neither reciprocity nor MFN can by itself exhibit these features in a many-country setting. McCalman (2002) offers a different formalization of the benefits of the MFN rule. In the setting he considers, informational asymmetries play a central role, and provided that the number of small countries is sufficiently great, a large country has diminished capacity under the MFN rule to hold an agreement hostage and extract rents from its small trading partners, enhancing both global efficiency and the payoffs of small countries. Other implications of the MFN rule are explored in Choi (1995), Ludema}
But new issues also arise in a many-country world that are themselves related to the adoption of the MFN principle. One issue concerns concession erosion: as discussed in section 4, in anticipation that trade liberalization would proceed according to a sequence of bilateral negotiations, the MFN principle was included in the RTAA and subsequently in the GATT/WTO in part to address this issue. Another issue involves bargaining tariffs: in the presence of the MFN principle it was feared that free-riding and the implied loss of bargaining power could become a serious obstacle to successful negotiations, as discussed in section 4 as well. At the heart of these issues is a common theme: when countries engage in bilateral tariff bargaining, how can third-country spillovers be minimized?

Strikingly, the twin pillars of MFN and reciprocity can be understood as minimizing third-country spillovers from bilateral tariff bargaining. To see why, consider the case where foreign-country 2 is not involved in the negotiations and keeps its tariff unaltered. In the presence of MFN, the domestic government and the government of foreign-country 1 can still negotiate a reciprocal reduction in their respective tariffs \( \tau \) and \( \tau_{1}^{*} \) which leaves the terms of trade \( \tilde{p}^{w}(\tau, \tau_{1}^{*}, \tau_{2}^{*}) \) unaltered but reduces \( p \) while raising \( p_{1}^{*} \), and which therefore provides these two countries with greater trade volume. But recall now that in foreign-country 2 we have the pricing relationship \( p_{2}^{*} = p^{w}/\tau_{2}^{*} \). It follows that, with \( \tau_{2}^{*} \) held fixed, if the negotiation between the domestic country and foreign-country 1 abides by MFN (so that a single equilibrium world price \( \tilde{p}^{w} \) prevails) and reciprocity (so that \( \tilde{p}^{w} \) is unaltered) then \( p_{2}^{*} \) and therefore \( W_{2}^{*}\left(p_{2}^{*}, \tilde{p}^{w}\right) \) and foreign-country 2’s trade volume are unaltered by these negotiations as well. In abiding by the principles of MFN and reciprocity, the domestic government and the government of foreign-country 1 have thus engineered a bilateral tariff bargain without third-country spillovers.

How can it be that exporters from foreign-country 2 experience a reduced MFN tariff from the domestic country and yet do not enjoy any increase in their export volume? The reason is that these exporters compete for sales in the domestic market with exporters from foreign-country 1, and exporters from foreign-country 1 become more competitive due to the negotiated reduction in foreign-country 1’s import tariff, which releases productive resources from foreign-country 1’s import-competing sector and allows these resources to move into foreign-country 1’s export sector.

It is also interesting to observe that MFN by itself is not enough to accomplish this feat. This can be seen by noting that a non-reciprocal negotiation between the domestic country and foreign-country 1 could also be undertaken in the presence of MFN, but such a negotiation would alter \( \tilde{p}^{w} \) and hence (with \( \tau_{2}^{*} \) held fixed) alter \( p_{2}^{*} \) and therefore in general alter \( W_{2}^{*}\left(p_{2}^{*}, \tilde{p}^{w}\right) \) and foreign-country 2’s trade volume. This suggests a qualification to the position that is sometimes taken in the literature that the MFN principle is by itself sufficient to eliminate the risk of concession erosion.

The preceding discussion leads to an important insight: the MFN rule permits the liberalizing force of reciprocity to be harnessed in an essentially bilateral manner even in a multilateral world.

---

29 These and related points are developed in Bagwell and Staiger (2005a, 2007).
30 This position is suggested by Schwartz and Sykes (1997) and Ethier (2004).
This means that, at least in principle, countries can negotiate split concessions with a sequence of trading partners who need not fear that those concessions will be eroded by later bargains, as long as the bargains abide by the principles of MFN and reciprocity; and where strict reciprocity is not feasible (and hence some spillovers become inevitable), ordering this sequence in accordance with the principal-supplier rule and under the threat of modification or withdrawal can help preserve a country’s bargaining power along the way. This describes a negotiating forum and a set of bargaining tactics that broadly mirror those anticipated from the U.S. experience with the RTAA.\textsuperscript{31}

Are these features borne out in GATT/WTO practice? The existing evidence takes two forms. One form concerns the trade-volume impacts associated with GATT/WTO membership. According to the findings of Subramanian and Wei (2008), GATT/WTO membership is associated with a large and significant increase in trade volumes for developed countries, but developing-country members experience a weak or non-existent impact on their trade volumes. Given that developed countries have been the main participants in GATT/WTO sponsored tariff bargaining while developing countries have been largely inactive in this capacity, the findings of Subramanian and Wei are broadly consistent with the interpretation that the participants in GATT/WTO tariff bargaining have successfully neutralized significant third-country spillovers.\textsuperscript{32} Of course, an alternative interpretation is that developed countries have simply found ways around the MFN principle, and have in their tariff bargaining discriminated against non-participating GATT/WTO members. A paper whose findings weigh against this alternative interpretation is Bown (2004a). In the context of GATT/WTO bilateral dispute settlement negotiations at least, Bown finds that countries do indeed abide by the MFN principle.

This first form of evidence is therefore suggestive, but it is far from conclusive. An interesting avenue for further exploration would be to draw a tighter link between the trade effects experienced by GATT/WTO members and their negotiated tariff commitments. For example, according to the logic of the discussion above, if GATT/WTO bargains stick closely to the MFN and reciprocity principles, it is a country’s own liberalization relative to that of competing exporters, more than the liberalization in the markets to which it exports, that should be decisive in determining the trade effects it experiences from GATT/WTO membership.

A second form of evidence looks directly at the pattern of tariff bindings negotiated in the GATT/WTO, and seeks to uncover possible evidence of free riding and its impacts on bargaining outcomes. Ludema and Mayda (2007, and forthcoming) relate the expected severity of the MFN free-rider problem to a Herfindahl index of the concentration of foreign exporters into a given country’s markets, interacted with cross-country variation in the foreign export supply elasticities faced by a given country to reflect its power over the terms of trade. Using this relationship,

\textsuperscript{31}These points are developed formally in Bagwell and Staiger (2007).
\textsuperscript{32}In fact, in a provocative paper, Rose (2004) reports finding no significant trade-volume impact of GATT/WTO membership whatsoever, but Rose fails to allow for a differential impact across countries. As the findings of Subramanian and Wei (2008) confirm, and as the theory we have discussed implies, allowing for differential trade-volume impacts based on different levels of bargaining participation is crucial for understanding the impacts of GATT/WTO membership. See also Tomz, Goldstein and Rivers (2007) who argue that Rose’s results are overturned when a more complete definition of GATT membership is employed.
they find evidence of free riding in the tariff bargaining that occurs in GATT/WTO negotiating rounds, and suggest that the impact of this free riding on GATT/WTO tariff bargaining outcomes could be substantial. Using the relationship developed by Ludema and Mayda, Bagwell and Staiger (2009c) explore the possibility of a free-rider problem in the context of accession negotiations in the WTO, and find little evidence of a free-rider problem in that setting. The approach of Ludema and Mayda is developed to get around the lack of available data on the actual participants in any given GATT/WTO tariff bargain. Detailed bargaining records do exist, however, and if made available to researchers they could provide a valuable tool for further empirical work on this important question.

Finally, one might wonder whether the fear of concession erosion remains a powerful force in determining GATT/WTO bargaining outcomes. This question takes on special relevance, because the MFN principle is a central means by which the fear of concession erosion was to be allayed, and yet exceptions to the MFN principle under GATT Article XXIV have been widely exercised by governments for the purpose of forming preferential agreements. Here we simply observe that the fear of concession erosion may have been replaced by the related fear of “preference erosion,” which can also become a powerful force—a “stumbling block” in Bhagwati’s (1991) terminology—against further MFN tariff reductions. Support for this position is provided in Limao (2006, 2007) and Karacaoglu and Limao (2008), who model the interaction between preferential and multilateral negotiations and find evidence for a significant stumbling-block effect of the preferential agreements negotiated by the United States and the EU, driven by the incentive of U.S. and EU preferential partners to stop the preference erosion that further U.S. and EU MFN tariff liberalization would imply.33

5.3 Enforcement and Dispute Settlement Procedures

As we discussed in section 4.2, membership in the GATT/WTO carries with it an obligation to abide by a set of rules. But how are these rules enforced? After all, if under the rules governments manage to negotiate from the Nash point to a point on the contract curve, such as the political optimum, the temptation will be substantial for a government to unilaterally select a high tariff and shift costs, and this temptation does not go away simply because an agreement is signed. Rather, as there is no “world jail” into which government leaders are thrown if they violate GATT/WTO rules, an effective GATT/WTO must ensure that such temptations to deviate from the agreement are balanced against the anticipated costs of the retaliatory response by other governments that the deviation would provoke. That is, if it is to be effective, the GATT/WTO must be “self-enforcing.”

As emphasized more generally by McMillan (1986, 1989), Dixit (1987) and Bagwell and Staiger (1990) among others, the enforcement issues that are associated with trade agreements may be

33 We note, though, that this incentive is not always the dominant force determining the impact of preferential agreements on MFN tariffs: focusing on a set of developing countries whose MFN tariffs are relatively high, Estevadordal, Freund and Ornelas (2008) find that membership in a free-trade agreement leads to a reduction in MFN tariffs of the member countries, possibly reflecting the stronger “tariff-complementarity” effects that arise when MFN tariffs are high (on the tariff-complementarity effect, see Bagwell and Staiger, 1999b, Freund, 2000, and Ornelas, 2005 and 2007). Freund and Ornelas (forthcoming) survey the literature on preferential agreements and their relationship to the GATT/WTO more broadly.
analyzed using the theory of repeated games, and a large literature has emerged on this topic.\textsuperscript{34} Here we emphasize just three points.

First, the repeated-game perspective is broadly consistent with the GATT/WTO enforcement provisions as described in section 4.2. In particular, as we have indicated, the creation of the GATT/WTO and its nullification-or-impairment procedures may be viewed as an attempt to replace the Nash outcome with a more efficient equilibrium outcome. To accomplish this, governments agreed to limit the use of retaliation along the equilibrium path and reposition it as an off-equilibrium-path threat that enforces the rules. This view is well-captured in a statement by one of the drafters of GATT’s nullification-or-impairment clause (as quoted in Petersmann, 1997, p. 83):

We have asked the nations of the world to confer upon an international organization the right to limit their power to retaliate. We have sought to tame retaliation, to discipline it, to keep it within bounds. By subjecting it to the restraints of international control, we have endeavored to check its spread and growth, to convert it from a weapon of economic warfare to an instrument of international order.

It should be stressed that there have been numerous GATT/WTO disputes whose resolution has entailed some form of policy response by the disputants, either in the form of explicit, authorized retaliation (rarely in GATT, more often in the WTO) or in the form of a settlement agreement reached in the “shadow” of authorized retaliation or its anticipation (the majority of GATT/WTO cases), and so a limited role for retaliation on the equilibrium path does arise in the GATT/WTO. But from the perspective of the theory of repeated games, this on-equilibrium-path retaliation can be interpreted as adjustments to the bargain that are made in response to shocks which would otherwise upset the balance between the temptation to deviate from the bargain and the costs of the penalty for deviation. Consistent with the theory of repeated games, it is this penalty, which could amount to a breakdown of the entire GATT/WTO system and a return to a Nash “trade war,” and which is the ultimate threat that maintains the international order, that remains off the equilibrium path.\textsuperscript{35}

A second point is the special challenge that smaller countries face in ensuring that their rights are enforced in the GATT/WTO, given that enforcement depends importantly on the ability to retaliate.\textsuperscript{36} This issue has particular relevance for developing country members, who as Hudec (2000) recounts first put forward a proposal on GATT remedies in 1965 suggesting that “collective retaliation” be permitted in cases where a large country violated its obligations to a developing country. Related empirical work includes Prusa and Skeath (2001) and Blonigen and Bown (2003).


\textsuperscript{35}See Bagwell and Staiger (2002b, Ch. 6) for a more extensive discussion of the relationship between retaliation in GATT/WTO and the theory of repeated games.

\textsuperscript{36}See Park (2000) for an analysis of enforcement issues that arise between a small and a large country, and Bown (2004b,c) for evidence on the important effect that the ability to retaliate has on the resolution of GATT/WTO disputes.
country. This raises the interesting question whether permitting multilateral retaliation in the context of a bilateral dispute is warranted in some circumstances. On the one hand, as the quotation above suggests, an important purpose of the GATT/WTO dispute settlement procedures is to restrain retaliation and “check its spread and growth,” and any move toward multilateral retaliation would clearly go against this purpose. On the other hand, it is difficult to see how smaller countries could attain equal footing in a GATT/WTO dispute with their large trading partners without the aid of multilateral retaliation in some form. More recently, Mexico has proposed in the ongoing Doha Round a number of changes to the dispute settlement procedures, among them a variation on the idea of multilateral retaliation according to which the right of retaliation would be made “tradeable” (see WTO, 2002). Maggi (1999) provides a theoretical framework for understanding the potential benefits of multilateral retaliation, and there has been some recent attention to this issue stimulated by the Mexican proposal. But the relevant questions pertaining to this important issue have just begun to be explored.

Finally, while the literature discussed above interprets disputes as fundamentally about enforcement, many GATT/WTO disputes seem to be at least as much about interpreting vague clauses in the agreement or filling gaps where the agreement is simply silent. Moreover, some of the most intense debates about the design of the dispute settlement procedures concern the appropriate degree of “activism” (see, for example, Goldstein and Steinberg, 2007): Should WTO panels be allowed to engage in “judicial lawmaking,” or should they be constrained to rule only on the obligations that are clearly stated in the agreement? Formal analysis of these issues requires that the contractual incompleteness of the GATT/WTO be placed at center stage, and while recent work has begun to explore these issues, such analysis is still in its infancy.

6 Conclusion

We have surveyed recent economic research that attempts to understand and interpret the design and practice of the GATT/WTO. Our Review has focused on three of the most central features of the GATT/WTO – reciprocity, non-discrimination and enforcement/dispute settlement – but we have left out a number of other active areas of research. We conclude by mentioning two.

A first area is the treatment of non-border measures in the GATT/WTO. In our discussion above, we feature a simple setting in which governments negotiate restrictions on import tariffs alone. Of course, in reality, governments select domestic or non-border policies as well, and these policies may also impact trade flows. As emphasized in much recent research, an efficient trade agreement must also restrict governments from using domestic policies to favor import-competing firms. For example, an efficient agreement must place restrictions on the ability of governments

---

37 See, for example, the analysis of auctioning retaliation rights in Bagwell, Mavroidis and Staiger (2007), and the analysis of the exchange of bonds as an enforcement mechanism in Limao and Saggi (2008).

38 A recent paper by Horn, Maggi and Staiger (forthcoming) builds on the writing-costs approach of Battigalli and Maggi (2002) to develop a theory of trade agreements as endogenously incomplete contracts, but that paper does not consider the role of dispute settlement procedures. A paper that does consider the role of dispute settlement procedures in a related incomplete-contracts setting is Maggi and Staiger (2008).
to undermine negotiated market-access commitments by offering new subsidies to their import- 
competing firms. At the same time, restrictions on domestic policies should not be too severe; for 
example, subsidies can be a first-best instrument with which to target a domestic market failure. 
This tension is reflected in recent research that considers the optimal treatment of domestic subsidies 
in a trade agreement, and also in recent work that explores related issues with respect to policies 
that concern domestic labor and environmental standards (when the international externalities 
associated with such standards are pecuniary). A related set of recent research considers whether 
an agreement that links tariffs and standards might enhance efficiency by facilitating improved 
enforcement or information-revelation capabilities. Finally, recent work also indicates that the 
treatment of domestic policies in a trade agreement should balance the additional contracting costs 
associated with new rules for domestic policies against the degree to which such policies represent 
an effective means of manipulating market access.

A second area is the treatment of export policies in the GATT/WTO. While GATT rules 
concerning export subsidies were somewhat permissive, export subsidies are now prohibited in the 
WTO. A theory of how export subsidies should be treated in a trade agreement must begin by 
explaining why a government might use an export subsidy. The theoretical literature emphasizes 
that a government might use an export subsidy in a strategic fashion so as to shift profits or delocate 
firms from competing foreign export sectors. In short-run models where profits can be shifted 
but the location of firms is fixed, the governments of countries that export a given good could 
enjoy greater welfare if they were to reach an agreement under which they limit export subsidies. 
This result appears to offer a potential interpretation for the prohibition on export subsidies in 
the WTO. This interpretation, however, neglects the interests of consumers in importing countries, 
who benefit from the lower prices that export subsidies induce. Global welfare actually may be 
lower when governments of exporting countries agree to limit export subsidies, and so these models 
do not offer a compelling rationale for the prohibition of export subsidies. In long-run models 
where trade policies affect the entry and exit decisions of firms, if transport costs exist, then a 
country may be tempted to use an export subsidy, and such a subsidy would lower the welfare of 
its trading partner. Recent work suggests, though, that a country may not find an export subsidy 
attractive unless its import tariff is also low. This work suggests that a ceiling on export subsidies 
may be efficiency enhancing from a global standpoint once import tariffs have been negotiated 

39 Copeland (1990) offers an early model of the limits of cooperation when some domestic policies are left unre-

stricted. Bagwell and Staiger (2006) develop a model that focuses on the treatment of domestic subsidies in efficient 
trade agreements, and they evaluate GATT and WTO rules from this perspective. For further discussion of the legal 
rules concerning the use of domestic subsidies in GATT and now the WTO, see Sykes (2005). For further discussion 
of the labor and environmental standards and trade-agreement theory, see Bagwell and Staiger (2001a, 2002b Ch. 
8), Bagwell, Mavroidis and Staiger (2002) and Staiger and Sykes (2009).


41 See Horn, Maggi and Staiger (forthcoming).

42 For further discussion of the treatment of export subsidies in GATT and now the WTO, see Sykes (2005).

43 For a survey of the strategic-trade literature, see Brander (1995).

44 The same can be said for models of export subsidies based on political-economy considerations. For further 
discussion, see Brander and Spencer (1985), Bagwell and Staiger (2001c) and Bagwell (2008).
of export subsidies now found in the WTO.\textsuperscript{45} The treatment of export policies in GATT/WTO represents an important direction for further research.

\textsuperscript{45}Venables (1985) offers an initial treatment of the firm-delocation effect and the associated appeal of a unilateral export subsidy. Bagwell and Staiger (2009d) consider a linear version of the model and establish that free trade is then efficient and that a country finds a unilateral export subsidy appealing only once its import tariff is sufficiently low.
References


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: 283-323


Dhingra, S., 2009. Reconciling observed tariffs and the median voter model. Mimeogr., The Univ. of Wisconsin.


Horn H, Maggi G, Staiger RW. Forthcoming. Trade agreements as endogenously incomplete contracts. *American Economic Review*


Ludema RD, Mayda AM. Forthcoming. Do countries free ride on mfn? *Journal of International Economics*


36
Staiger RW. 2006. What can developing countries achieve in the WTO? *Journal of Economic Literature* 44: 779-95


WTO. 2002. Dispute settlement body - special session - negotiations on improvements and clarifications of the dispute settlement understanding - proposal by Mexico: November 4, catalogue record TN/DS/W/23

WTO Appellate Body 2004 United States-Anti-Dumping Act of 1916 (original complaint by the European Communities)—Recourse to arbitration by the United States under 22.6 of the DSU. WT/DS136/ARB, 24 February