Is the WTO passé?*

Kyle Bagwell
Stanford and NBER

Chad P. Bown
The World Bank and CEPR

Robert W. Staiger
Dartmouth and NBER

Preliminary: November 2014

Abstract

Membership in the World Trade Organization (WTO) is leading to a very different set of policy outcomes for governments than those emerging from the most recent wave of preferential trade agreements (PTAs), including current negotiations under the “mega-regional” agreements spear-headed by the United States, such as the Trans-Pacific Partnership (TPP) and the Trans-Atlantic Trade and Investment Partnership (TTIP). Should economists see these developments as an efficient institutional hand-off, where the WTO has carried trade liberalization as far as it can manage, and is now passing the baton for PTAs to finish the job? We survey a growing theoretical and empirical economics literature on international trade agreements to provide an answer to this question. We argue that this literature does not support the view that the WTO is passé, and that based on the literature economists would be hard-pressed to make a compelling case that it is now PTAs rather than the WTO that should be entrusted with the rules of globalization.

*Bagwell thanks the Center for Advanced Studies in the Behavioral Sciences for support and hospitality. Bown acknowledges financial support from the World Bank’s Multi-Donor Trust Fund for Trade and Development. Staiger gratefully acknowledges financial support from the NSF (SES-1326940). We thank Daniel Trefler and Douglas Irwin for useful discussions. Semira Ahliiyih provided outstanding research assistance.
1 Introduction

It is now clear that membership in the World Trade Organization (WTO), and its predecessor organization the General Agreement on Tariffs and Trade (GATT), is leading to a very different set of policy outcomes than those emerging from the most recent wave of preferential trade agreements (PTAs). Over a period spanning nearly 50 years, the GATT concluded 8 successful rounds of multilateral trade negotiations, a feat that led to a reduction in the average ad valorem tariff on industrial goods to below 4 percent and an expanding GATT/WTO membership from 23 to nearly 160 countries. But the GATT/WTO tariff liberalization process has ground to a halt with the 9th round, the so-called Doha Development Round, now moribund. And the scope of GATT/WTO liberalization, with its traditional focus on border measures, is and always has been shallow. By contrast, PTAs have now emerged as the vehicle by which countries reduce their tariffs from current WTO levels down to zero, albeit on a discriminatory basis, with the number of PTAs expanding from roughly 100 in 1990 to nearly 400 today. And the intended scope of PTA liberalization, which reaches further and further behind the border, is increasingly deep.

Should economists see the current state of affairs as an efficient institutional “hand-off,” with the GATT/WTO having carried trade liberalization as far as it could manage, and now passing the baton on to PTAs so that PTAs can finish the job and help governments arrive at the international efficiency frontier?¹ And if so, can PTAs also be entrusted with the task of seeing that governments remain at the international efficiency frontier with their own systems of dispute resolution? If these questions can be answered in the affirmative, then economists should view PTAs as the legitimate successor to the GATT/WTO and could reasonably conclude that “the WTO is passé.” But there are alternative interpretations of these developments.

One possibility is that PTAs are needed to complete the liberalization process and move governments to the international efficiency frontier, but that a central role for dispute resolution should continue to reside at the WTO. Under this view PTAs and the WTO are complementary to an efficient multilateral trading system, and both deserve the support of economists.

More ominously, the current state of affairs might be seen as the ultimate proof that PTAs are stumbling blocks to the multilateral trading system. According to this interpretation, the WTO still has important liberalization work to do, but it has stalled short of its goal at least in substantial part because of the existence and ready availability of PTAs for its member governments. From this perspective it is true that the liberalization process under the GATT/WTO has ground to a halt short of the international efficiency frontier, but rather than as its legitimate successor PTAs should be seen as a root cause of the WTO’s current woes.

Or governments may have in fact already achieved the international efficiency frontier under the GATT/WTO liberalization process – or if they have not yet achieved it, at least could achieve

¹By “international efficiency frontier” we mean policy choices that could not be adjusted to generate Pareto improvements across countries when each country’s welfare is judged by the preferences of its government. This focus on government preferences generally follows the literature, though a distinction is made in some of the literature between ex-ante and ex-post government preferences. We discuss these points further in our survey.
it with some selective fixes to the GATT/WTO process – so that the WTO is actually in far better shape than would first appear; and rather it is the rise of PTAs and the nature and kind of additional liberalization associated with PTAs that represents the big ongoing failure of efficient international trade policy cooperation. With this interpretation PTAs are liberalization run amok.

In this paper we make use of a growing economics literature on international trade agreements to sort through these interpretations and provide answers to the questions posed above. To facilitate our discussion, we adopt a simple organizing principle: we group papers in the literature by their stance on what makes a trade agreement valuable to its member governments, that is, by the nature of the “problem” that a trade agreement is supposed to “solve” for its member governments. According to this organizing principle, there are four strands of the literature.

The oldest and most established strand of the literature is the “terms-of-trade” theory of trade agreements. The terms-of-trade theory posits that governments use trade agreements to undo the policy inefficiencies that are associated with unilateral policy choices when those choices can shift the costs of intervention onto trading partners through movements in foreign exporter prices (terms of trade). In this theory, addressing an international externality (which travels through the terms of trade) is the central purpose of a trade agreement. A second strand of the literature, the “commitment” theory, also has a well-established history in the literature, but here the central role for an international externality is absent. Instead, according to the commitment theory, governments value trade agreements as a way to tie their hands (make commitments) against their own lobbies and citizens. The two remaining strands of the literature, what we call the “delocation/profit-shifting” theory of trade agreements and the “offshoring” theory of trade agreements, are more recent arrivals. They can be viewed as attempts to identify international externalities beyond the terms-of-trade externality that give rise to and shape international trade agreements. In these theories the international externalities take forms that can be interpreted as extending beyond the terms of trade to include the domestic-market prices in each country. The delocation/profit-shifting theory argues that such non-terms-of-trade externalities have been important for understanding real-world trade agreements all along, while the offshoring theory suggests that non-terms-of-trade externalities may only have become prominent with the recent rise of offshoring and international supply chains.

As might be anticipated, the strength of the literature’s support for the various interpretations of recent developments depends to some extent on which purposes are central to real-world trade agreements and hence which strands of the literature are considered most relevant for interpreting these developments. While we discuss below evidence that lends support to all four strands, a growing body of this evidence points to the terms-of-trade theory of trade agreements as central for understanding the actual trade agreements that we see. We therefore first evaluate these developments from the perspective of the terms-of-trade theory, surveying both the relevant theoretical and empirical literature to assess the various interpretations and establish some initial answers to these questions. We then survey the commitment, delocation/profit-shifting and offshoring strands of the literature, describing where they yield different assessments of these interpretations, and we
utilize this description in combination with a survey of the relevant empirical literature relating to these theories to suggest qualifications to the answers provided by the terms of trade theory.

To preview our bottom line, the literature we survey does not support the view that the WTO is passé. On the contrary, from the perspective of the terms-of-trade-theory strand of the literature, if anything economists should adopt a cautious view of PTAs and reserve their strongest support for the WTO, which with some selective fixes to the GATT/WTO process could according to this literature more credibly be relied on than PTAs to bring governments to the international efficiency frontier and keep them there. The commitment, delocation/profit-shifting and offshoring theories do raise important caveats to unqualified support for the WTO, and there are features of PTAs that these other strands of the trade-agreements literature would support. But at least until more empirical evidence suggests otherwise, these theories and the empirical evidence related to them do not make a compelling case that it is now PTAs rather than the GATT/WTO that should be entrusted with the rules of globalization.

The remainder of our survey is structured as follows. To set the stage, in the next section we provide a brief overview of the main institutional features of the world trading system, focusing on its two principal components, the GATT/WTO multilateral trading system and the current state of PTAs. In section 3 we review the basics of the terms-of-trade theory of trade agreements and survey the empirical literature that relates to its essential tenets. In sections 4 and 5 we then use the terms-of-trade theory and the strand of the literature which builds upon it as a lens through which to evaluate the GATT/WTO and PTA approaches to trade liberalization, and from this perspective we interpret the recent developments in the world trading system and provide initial answers to the questions posed above. In section 6 we survey the literature on the commitment, delocation/profit-shifting and offshoring theories of trade agreements, and we identify insights from each of these strands of the literature that suggest qualifications to the answers provided by the terms of trade theory. Then in section 7 we turn to an evaluation of dispute settlement in the world trading system, and we augment our answers with this evaluation. Finally section 8 concludes.

2 The World Trading System: a brief overview

Individuals are the ultimate drivers of globalization, but governments set the rules of the game, and the rules can be very important to the outcome. Here we provide a brief summary of the two main sets of rules for the world trading system: the policy commitments and their enforcement under the GATT/WTO multilateral system, and the sets of rules associated with the web of PTAs currently in force. We describe how the recent wave of PTAs is changing the rules of globalization along a number of important dimensions relative to the rules established by the GATT/WTO, and we suggest that there are important choices embedded in these two institutional forms.

2.1 The GATT/WTO Multilateral System

We begin with some background on the GATT/WTO Multilateral System.
A short history of tariff liberalization under the GATT and WTO  The GATT began in 1947 with 23 Contracting Parties and grew in membership over the next five decades before ultimately being consolidated into the WTO in 1995. As of 2014, the WTO counts 160 member economies – including both the European Union and each of the 28 EU member states individually.\(^2\)

The trans-Atlantic economies of the United States, Canada, and a number of European countries were not only a driving force behind the creation of GATT, but they provide what is perhaps the most familiar story line for how the key GATT/WTO countries achieved multilateral trade liberalization and sustained periods of low most-favored-nation (MFN) tariffs. These countries used the GATT forum to negotiate reciprocal MFN tariff reductions through periodic negotiating rounds (WTO, 2007), and they then locked in those tariffs through legally binding commitments which led to a gradual, 50 year period of multilateral liberalization. Table 1 illustrates the multilateral trade liberalization process of negotiations covering 1947-1994, and Table 2 documents the resulting average applied tariff rates for a number of these countries in 1952 and again in 2005 after eight “rounds” of multilateral GATT negotiations.\(^3\)

The trans-Atlantic GATT experience is not, however, how countries have universally liberalized their MFN tariffs or even entered into the GATT/WTO system. There are two prominent deviations from this experience that can be seen in the GATT/WTO experience of other member countries.

First, some countries did not enter the multilateral system at its inception; indeed, many such “latecomers” did not seek or were not admitted entrance into the agreement until well after the initial set of GATT Contracting Parties had already negotiated substantial levels of MFN tariff liberalization. As such, the GATT/WTO has had the flexibility to accommodate major economy accessions such as West Germany in 1951, Japan in 1955, China in 2001 and Russia in 2012.\(^4\)

Second, a number of developing countries chose not to participate in significant aspects of the reciprocal tariff liberalization negotiations that took place under successive GATT rounds even though they were party to the agreement; instead, these countries requested the application of a “special and differential treatment” exemption from reciprocity in order to pursue import substitution policies including application of persistently high tariffs and other nontariff barriers. This

\(^2\)The European Union (EU) is a member of the WTO; for legal reasons it was officially known until 2009 as the European Communities. The 28 individual countries of the EU are also WTO members in their own right. The EU is a single customs union with a single trade policy and tariff, and the European Commission “speaks” on behalf of the EU member states in most WTO matters. Nevertheless, it is important to point out that other customs unions are not represented in the WTO in this manner. E.g., MERCOSUR is not a WTO member, despite it being a customs union between Argentina, Brazil, Paraguay and Uruguay - each of which is a WTO member individually.

\(^3\)By 1952, average import tariffs had already fallen substantially from Great Depression peak levels in the 1930s and 1940s due to a combination of inflation, as many tariffs were imposed as specific duties, and the negotiated liberalization of the first three GATT rounds. Irwin (1995, Table 5.2) reports average tariff rates in 1931 (after the US imposition of its Smoot-Hawley tariff) for France, Germany, and Italy of 38, 40, and 48 percent, respectively. Irwin (2011, 2012) provides important accounts of the political-economy of the import protection that increased sharply during the Great Depression period, and Irwin, Mavroidis, and Sykes (2008) describe the negotiations that ultimately led to establishment of the GATT in the late 1940s. For a more extensive analysis tracking the multilateral trade liberalization that took place across countries over the 60 year period following the GATT 1947 inception, see WTO (2007).

\(^4\)China was an original Contracting Party to the GATT but withdrew in 1950. The other two original contracting parties to subsequently withdraw from the GATT were Lebanon and Syria.
includes 1947 GATT founders such as India and Brazil. To the extent that these countries currently apply relatively low (in historical terms) MFN tariffs, their liberalization episodes frequently were not undertaken reciprocally, but instead unilaterally (e.g., India) or in concert with a period of preferential liberalization (e.g., Brazil). These countries also did not follow the trans-Atlantic approach of gradually lowering their MFN tariffs over a span of decades; instead, their period of low and sustained multilateral tariffs began suddenly and not until the 1990s. And while the currently applied MFN tariffs of these countries may be relatively low, as we further detail below, their applied rates have not been legally bound under the WTO at similarly low levels, unlike the record of the United States and the European Union.

Finally, the WTO’s 160 members notwithstanding, there are at least three dozen countries that are not yet members.5 While most of these are developing countries and some of them may have liberalized their own economies (through either preferential or unilateral tariff liberalization), there remain roughly 500 million people - or 7 percent of the global population - that reside in countries that are entirely outside of the WTO system. These countries have not made legally binding commitments on their applied MFN import tariffs. And potential exporters in these countries do not benefit from legally binding MFN treatment from trading partners, nor do they have access to the WTO’s dispute settlement provisions to self-enforce the market access rights defined by receipt of MFN treatment.

Contemporary tariff commitments under the WTO Table 3 summarizes many of the salient features resulting from the GATT/WTO’s “shallow” integration approach to trade liberalization, including information on contemporary multilateral tariffs across and within the major economies. The table splits countries into three groups – the high income members of the Group of 20 (G20), the emerging economy members of the G20 (which includes the BRICS, namely, Brazil, Russia, India, China and South Africa), and a selected sample of other major developing countries with 2012 populations of over 50 million - some of which are not (yet) WTO members, as they are currently only WTO “observers.”6 Overall, the data reveal a substantial degree of heterogeneity across countries and across industries, and include many examples within and across countries where applied MFN import tariffs, as well as the bindings that have been legally negotiated to constrain them, are not close to free trade.

Consider first the data for the United States. The simple average MFN tariff that the US applies to imports from any other WTO member is 3.4 percent. One hundred percent of the US

---

5At the time of writing, this includes the following WTO observers: Afghanistan, Algeria, Andorra, Azerbaijan, Bahamas, Belarus, Bhutan, Bosnia and Herzegovina, Comoros, Equatorial Guinea, Ethiopia, Holy See (Vatican), Iran, Iraq, Kazakhstan, Lebanese Republic, Liberia, Libya, Sao Tomé and Príncipe, Serbia, Seychelles, Sudan, Syria and Uzbekistan. The following are recognized by the UN as countries but are not WTO members: Eritrea, Kiribati, Marshall Islands, Federated States of Micronesia, Monaco, Nauru, North Korea, Palau, San Marino, Somalia, South Sudan, Timor-Leste, Turkmenistan, and Tuvalu.

6Governments that have “observer” status in the WTO are non-members that are granted certain rights and are expected to uphold certain obligations. Rights include the ability to attend certain WTO meetings and obtain access to information; obligations include minimal contributions to the WTO’s budget. For more information, see WTO (2008, Chapter 4).
tari€ lines are bound at some level, and the simple average binding rate is 3.5 percent. The binding
is the formal legal commitment that a country submits to the WTO membership and entails the
promise of the rate above which it will not raise its applied tariff. The WTO permits countries to
offer applied tariff rates below their tariff binding rates, provided that such offerings are made to
all other members on a nondiscriminatory (MFN) basis. The fact that, for the United States, the
average applied MFN import tariff is pushing up against the average binding rate is an indicator
that the US has very little scope to unilaterally increase its applied MFN import tariffs without
running afoul of WTO rules.

While the average US applied and bound MFN tariff rates are quite low, multilateral trade
policy exhibits considerable heterogeneity across countries, even within high-income economies.
Most of the major industrialized economies have almost universal binding coverage and applied
rates that are relatively close to their tariff bindings. Nevertheless, while these economies’ applied
MFN tariffs may be low in historical terms, they range from an average of 2.7 percent (Australia)
to 13.3 percent (South Korea). And while average tariffs in high income economies may be low,
there are still important examples of outliers or tariff “peaks.” For example, 2.7 percent of US tariff
lines have applied MFN rates that “peak” at higher than 15 percent, with the highest rate being
350 percent. Canada, the European Union, and South Korea each have more than 5 percent of
MFN tariff lines with rates higher than 15 percent, and maximum applied rates in these economies
are greater than 500 percent.

The tariff data exhibit even more heterogeneity across emerging and developing economies.
First, while average applied MFN tariffs may also be relatively low for these countries in historical
terms, even the rates applied by the relatively advanced (G20) emerging economies remain, on
average, much higher than the rates applied by industrialized countries. Second, some countries
(e.g., India) have not committed to legally binding a significant share of their tariff lines at any
level. Third, within the set of products that countries have committed to bind, there is a significant
differential between the applied rate and the binding commitment. This last point holds for all of
the G20 emerging economies (including Argentina, Brazil, India, and Mexico) with the exception
of the relatively new WTO accession countries of China (2001) and Russia (2012), for which the
existing membership demands included relatively low levels of MFN tariff bindings.7

Heterogeneity across the tariff data can be even more extreme for other major (but poorer)
developing countries. Some WTO members (e.g., Bangladesh, Burma, Nigeria) have committed
to upper limits for tariff bindings on fewer than 20 percent of their import tariff lines. Even on
products for which these WTO members bind their tariffs, the average binding rates may be more
than 100 percentage points higher than the applied rates.

There are also important differences in applied MFN tariff heterogeneity within countries across
sectors. As one important example, Table 3 shows many instances of sharp differences between
average applied tariffs in agricultural products relative to overall rates of protection. Within the

7 As of 2012, Russia had not yet fully phased in its MFN applied tariff cuts under its WTO accession terms and
thus its average applied rate was still above its average binding commitment.
G20, a few countries such as Argentina, Australia, and Brazil offer lower average import tariffs for agriculture than they do for other products. For most others, however, the rates in agriculture are substantially higher (Anderson, Rausser, and Swinnen, 2013).

Finally, applied MFN tariffs are not the only important trade policy instrument within the multilateral WTO system. An increasing number of countries since the early 1990s have begun to invoke GATT/WTO exceptions to their negotiated tariff bindings and use the temporary trade barrier (TTB) policies of antidumping, safeguards, and countervailing duties.

Before 1990, industrialized economies such as Australia, Canada, the European Union and the United States dominated overall use of TTBs, and especially the most predominant antidumping policy (Blonigen and Prusa, 2003). Since the early 1990s, a number of emerging economies have subsequently become major users of TTBs (Bown, 2011a) as they reduced their applied import tariffs. The last column of Table 3 provides data on the import coverage of the TTBs cumulatively applied in 2012. As examples, eleven different G20 economies had more than 1 percent of their tariff lines also subject to an imposed TTB in 2012 – including Argentina, Brazil, China, Indonesia, Mexico, and Turkey – a number of which did not even have an antidumping law in place 25 years earlier. Many of these countries also had one or more episodes during that 25 year period during which TTB import coverage rose to as high as 4-6 percent.

Table 3 reveals at least two other insights from the TTB statistics. First, not all WTO member countries are users of these TTB policies. Indeed most of the poorest WTO members have never implemented a formal antidumping or safeguard investigation. This is at least partially due to their tariff bindings being sufficiently above their applied rates so that they can adjust applied rates upward in response to shocks. Second, even members of a customs unions – i.e., countries that drop their tariffs on internal trade with each other to zero and share a common MFN tariff policy applied toward non-members, examples of which we describe in more detail below – do not necessarily apply a common set of TTB policies. In 2012, for example, customs union partners Argentina and Brazil had different shares of product lines covered by TTBs, as did the partners Turkey and the European Union.8

**GATT/WTO commitments relating to behind-the-border measures** The GATT traditionally eschewed efforts to negotiate restrictions on the use of behind-the-border measures of its member governments. The WTO has attempted to venture into this realm of “deeper” integration, most substantively first with the so-called Singapore issues originally included as part of the Doha Round negotiations and more recently with the much more modest Doha efforts to negotiate an agreement on Trade Facilitation, but as yet neither of these attempts has led to success. It is important to point out, however, that many areas of deeper cooperation over non-tariff barriers of focal interest to the new PTAs (that we describe below) are not entirely absent from consideration

---

8 Across customs unions, for example, there are different rules regarding whether members can apply TTBs against other customs union members - e.g., a country may not apply a safeguard against another customs union member under MERCOSUR whereas it is possible to apply such a safeguard under the EU-Turkey customs union. Finally, within customs unions there may even be different rules depending on TTB policy instruments - e.g., in MERCOSUR, members can impose antidumping on one another but not a safeguard.
by the WTO Agreements (WTO, 2012). It is simply that the GATT/WTO shallow integration approach has addressed such behind the border and non-tariff barriers differently.

For example, the WTO does permit governments some leeway to use public policy interventions - including those that will affect trade flows - in order to allow for product standards and, for example, to protect plant, animal and human health. Furthermore, the WTO provides some guidance for those interventions in an attempt to minimize the possibility that the interventions might become non-tariff barriers to trade that governments apply without a legitimate public policy or market failure motive. For trade in goods, these exceptions and guidelines are captured in the GATT’s basic rules on nondiscrimination and national treatment found in Article III, and they have been further elucidated under the GATT’s original Article XX, WTO’s Technical Barriers to Trade (TBT) Agreement, and the Agreement on the Application of Sanitary and Phytosanitary (SPS) Measures. Furthermore, countries can bring such non-tariff issues to the attention of the WTO membership by filing “specific trade concerns” with relevant WTO standing committees. Finally, the GATT/WTO has always had available a basic market access “preservation” rule that allowed member countries the right to initiate non-violation complaints under formal dispute settlement proceedings, an area to which we turn next.9

**Dispute Settlement under the WTO and GATT** Access to formal dispute settlement procedures has always been part of the GATT/WTO system. The 1947 GATT’s Article XXIII established the basic provisions whereby government-to-government dispute resolution would take place, and over the subsequent fifty years, Contracting Parties initiated more than 250 disputes under the GATT fora in attempts to formally resolve a variety of trading frictions that arose.10 Legal scholars generally characterize the GATT-provided mediation that took place during this period as a “diplomacy-based” approach to dispute resolution.

Many elements of the system changed dramatically in 1995 with the inception of the WTO, as the current system is much more “legalistic” than its GATT predecessor. Under the WTO, members initiated nearly 500 formal disputes against one another between 1995 and 2014, or nearly twice as many formal disputes to date as during the GATT period of 1947-1994. The literature identifies a number of contributing explanations, including that more countries are now actively involved in the trading system, there is substantially more trade, and countries have taken on more legally binding commitments.

Over time, more and more WTO members have found themselves involved in formal disputes. To date, nearly 50 out of the 160 WTO members have initiated a case as a “complainant” (i.e., the plaintiff) and more than 50 members have faced a dispute as a “respondent” (i.e., the defendant).11

---

9 Countries can file nonviolation complaints under the GATT’s Article XXIII. For an introduction, see Staiger and Sykes (2013). Our survey further elucidates a number of implications of this feature of the GATT/WTO.
10 See Bown (2002; Table 1).
11 The size of the total WTO membership is larger than the set of potentially active litigants given that each of the current 28 EU member states is an independent member of the WTO in addition to the EU’s own separate membership. It is relatively rare for an EU member state to be involved in a WTO dispute that does not involve the entire European Union.
Furthermore, more than half of the membership has been formally involved in at least one dispute via the legal status as an “interested third party.” This is a potentially important role even for countries without trade stakes involved in a particular bilateral dispute, given that jurisprudence arising from a dispute between any two countries – e.g., a policy dispute pitting Colombia versus Panama or Moldova versus Ukraine – could have policy implications for the entire WTO membership – including the United States and European Union.

The United States and the European Union are the two most frequent WTO litigants; combined they have initiated roughly 40 percent of all disputes, and roughly 50 percent of all disputes involve one or the other as a respondent. A significant share of the caseload also involves one of these economies challenging the other. Nevertheless, a number of other industrialized countries have also been frequent litigants in formal WTO disputes, including Australia, Canada, Japan, New Zealand, and South Korea. On the other hand, the share of WTO disputes involving developing country members, with the exception of least developed countries, has risen over time. There are at least three contributors to this phenomenon. First, the share of developing country trade in total trade has increased over time as has the share of the WTO membership made up by developing countries, so there is more developing country trade covered by the system over which disputes might arise. Second, some of the relative increase in developing country use is due to the sharp decline in absolute use of the system by the United States and the European Union especially after 2000. Third, much of the increased involvement of developing countries can be associated with China’s WTO accession, and the fact that China has been especially involved in litigation since 2007. The list of other developing countries that have been frequent WTO litigants includes Argentina, Brazil, Chile, Colombia, Guatemala, Honduras, India, Indonesia, Mexico, Panama, Peru, Philippines, and Thailand.

2.2 PTAs

While a central pillar of the GATT/WTO system is the MFN principle, GATT Article XXIV provides an exception to the MFN principle to allow GATT/WTO members to form PTAs that satisfy certain features. The key Article XXIV stipulations are that the PTA must eliminate tariffs on “substantially all” trade among the member countries, and that the external MFN tariffs that the member countries continue to apply to imports from outside the PTA not be increased as a result of the PTA formation.

For decades during the post-World War II period, much of the analysis of PTAs centered primarily on one successful experience of regional integration - i.e., the continuing and ongoing evolution of western Europe. The 1951 Treaty of Paris established the European Coal and Steel Community (ECSC) which was expanded with the 1957 Treaty of Rome to create the six-country European

---

12 The poorest and least developed country members of the WTO system – of which there are dozens – are almost entirely absent from participation in formal WTO dispute settlement. Bown and Hoekman (2008) provide a discussion of the political-economic hurdles facing the poorest and smallest countries in the WTO, including the difficulties they face in finding trading partners willing to use the (dispute settlement) system to hold them accountable to their trade agreement commitments.
Economic Community (EEC). Today’s European Union is the result of continued integration over the subsequent five decades, including numerous country accessions (28 member countries, as of 2014) as well as substantial “deepening” of negotiations and agreements beyond trade preferences and toward factor market, economic, monetary, and even political integration.

Beginning in the late 1980s, a number of other potentially economically meaningful PTAs have arisen and subsequently been sustained. These PTAs include the 1987 CUSFTA (Canada-US Free Trade Agreement) that was subsequently expanded into the NAFTA (North American Free Trade Agreement) through the addition of Mexico in 1994. There are also increasingly important developing country PTAs, including the MERCOSUR (Mercado Común del Sur) customs union involving Argentina, Brazil, Paraguay and Uruguay in the early 1990s, ASEAN (Association of Southeast Asian Nations) Free Trade Area involving Brunei, Indonesia, Malaysia, Philippines, Singapore and Thailand in the early 1990s, and CAFTA-DR (Central American Free Trade Area – Dominican Republic) involving the US and five Central American economies in the mid-2000s. As of 2014, the WTO reports that it has been notified of nearly 600 reciprocal trade agreements in existence, and nearly 400 agreements are currently in force (WTO, 2014a).

**Tariffs and behind-the-border measures in PTAs** There are two key areas in which PTAs push beyond the multilateral commitments countries undertake at the WTO. The first is by reducing import tariffs even lower than WTO levels, albeit on a discriminatory basis. The second is by introducing negotiations and potential cooperation beyond import tariffs and toward new policy instruments that are applied behind the border.

In a recent study, WTO (2011) provides a relatively comprehensive characterization of the global patterns of tariffs and trade taking place under PTAs. They begin by noting that with the sharp increase in PTAs since 1990, the value of trade between PTA members has grown faster than the world average; not surprisingly the share of intra-PTA trade in world trade has nearly doubled from 18 percent in 1990 to 35 percent in 2008. Including intra-EU trade flows in these statistics indicates that intra-PTA trade as a share of world trade increased from 28 percent to 51 percent over this period. However, their data analysis, based on a matching of product-level trade flows to tariffs and preferential tariffs to MFN tariffs, reveals a number of other important stylized facts, some of which question the conventional wisdom of the impact of PTAs thought to arise through additional discriminatory tariff liberalization.

First, while economists typically model PTAs as resulting in zero applied tariffs between partners, real-world PTAs do not always lead to zero tariffs on all intra-PTA goods trade. The data indicate that many negotiated exceptions within PTAs result in a significant number of PTA tariffs remaining at levels above zero, including eight percent of tariffs for the major PTAs of the US.

---

13 The number of notifications and trade agreements in force differ for a number of reasons. One is because notifications include not only new agreements, but also the accession of new countries to existing agreements - e.g., Croatia’s accession to the European Union in 2013. Second, some agreements notified to the GATT/WTO later become “inactive” (or no longer in force), when they become superseded by a subsequent agreement that was later notified and which is currently in force - e.g., CUSFTA is no longer in force as it was superseded by NAFTA.

14 The data reported here and below derives specifically from section II of WTO (2011, pp. 47-86).
Canada, EU, and Japan (Damuri, 2012). Indeed, in an analysis of the PTAs involving 85 countries and 90 percent of world trade in 2007, the WTO (2011, pp. 124-125) finds that roughly 66 percent of tariff lines with MFN tariff “peaks” (MFN rates defined as greater than 15 percent) have not been reduced at all through PTAs. Hence, while existing PTAs should be viewed as a significant force in eliminating (roughly one third of, and on a discriminatory basis) the tariff peaks that remain among WTO members, a majority of these tariff peaks still remain.

Second, while a large and increasing share of world trade takes place between PTA members, this share substantially overstates the amount of preferential trade between members. In many instances in which PTA tariffs are zero, the MFN tariffs are also zero, resulting in zero preference margin. Furthermore, even where positive preference margins exist, exporters may not utilize available preferences because of both the resource costs (to sourcing inputs from less efficient suppliers in PTA markets) and bureaucratic costs (to proving legal compliance) due to rules of origin and local value-added requirements needed to gain access to the lower preferential rates.\textsuperscript{15}

How much trade is therefore really taking place under preferential tariffs as opposed to MFN tariffs? First, between 49 percent (including intra-EU trade) and 65 percent (excluding intra-EU trade) of world trade takes place between countries that are not part of a common PTA. Second, excluding intra-EU trade flows, the WTO estimates that only 16 percent of global trade is eligible for any preferential tariffs; furthermore, less than 2 percent of global trade is eligible to receive preferences with margins above 10 percentage points.\textsuperscript{16} Including intra-EU trade in these statistics results in 30 percent of global trade being eligible for any preferential tariffs, 4 percent being eligible for margins over 10 percentage points. Combining these and other factors, and despite the explosive increase in PTA adoption since the late 1980s especially, the WTO estimates that, excluding (including) intra-EU trade, 84 percent (70 percent) of world merchandise trade still takes place on an MFN basis.

These observations suggest that, while existing PTAs have served as the most important conduit for the introduction of tariff discrimination within the WTO system, and while PTA member countries have agreed to some significant tariff reductions beyond their WTO tariff commitments, at least to date PTAs have played a far smaller role than might at first appear to be the case in actually delivering discriminatory tariff reductions on a wide scale that would lead to trade diversion (Viner, 1950).\textsuperscript{17} A second and increasingly important characteristic of PTA negotiations, however, is that they are no longer simply or even primarily about tariff liberalization. Instead, many PTAs are pushing negotiations toward deeper integration agreements that address non-tariff policies, including many policies that government policymakers impose behind the border.\textsuperscript{18} Some of the

\textsuperscript{15}See, however, Keck and Lendle (2014) for a recent challenge to the position that preferences often go unutilized.

\textsuperscript{16}Considering these figures with and without intra-EU trade flows may be important depending on the context, given that the European Union is a “unique” PTA in that it is also a customs union that has undertaken deeper integration - including factor markets and monetary integration for a substantial subset of member countries - and also steps toward political integration.

\textsuperscript{17}See also the discussion in Bhagwati (2008) and the survey in Panagariya (2000).

\textsuperscript{18}Some of these non-tariff policies under negotiation are applied at the border. For example, policies like antidumping and safeguards are frequently applied as quotas or price undertakings, and thus even though they are applied at the border are non-tariff barriers. Second, other potential non-tariff barriers that arise at the border may include
ongoing PTA negotiations that we describe below are even addressing areas that traditionally have been described as “domestic” regulation - labor and environmental standards, product standards, investment and tax provisions, etc.

As a definitional matter, the literature currently lumps issue areas arising under these new PTAs into one of two categories. The first category contains “WTO-plus” PTA provisions - i.e., those that also exist under the WTO, but where the PTA members in their agreement take on commitments to go further. Tariffs are the clearest example; e.g., WTO members make legally binding and enforceable MFN tariff commitments under the WTO, and PTAs typically involve partners choosing to lower at least some of those tariffs toward each other even further. Other examples include services, intellectual property rights, and product standards - each of which has at least some basic coverage through the WTO.

The second category for PTA provisions are “WTO-extra” areas, and these involve issues that are not yet explicitly addressed by the WTO. Examples of WTO-extra areas are labor standards, environmental standards, foreign direct investment provisions, movement of capital, competition policy, data protection, and even potential cooperation over other domestic regulations in order to help achieve improved levels of “regulatory coherence” across PTA member countries.

Horn, Mavroidis, and Sapir (2010) provide an early attempt at characterizing the “depth” associated with the provisions of PTAs by first examining the many US and EU PTAs in existence as of 2008. The initial evidence was that the EU’s PTAs tend to have many more WTO-extra provisions than do the US’s PTAs. However, the pattern is reversed when the analysis conditions on the legal enforceability (under dispute settlement) of the provisions, as US PTAs contain more legally enforceable WTO-extra provisions than do the EU’s agreements. Furthermore, the WTO (2011, Section D) has extended this approach of characterizing the nature of 14 different “WTO-plus” provisions and 38 different “WTO-extra” provisions to a wider sample of PTAs beyond those involving the US and EU so as to also include a number of PTAs involving developing countries. While it is beyond the scope of our survey to summarize the many characterizations of the data on these provisions contained in the WTO report, suffice it to say that the report’s analysis of a rich set of databases reveals substantial heterogeneity in application of these provisions across different PTAs. And the report identifies a number of open questions for research to tackle regarding the linkages between the non-tariff provisions of these deeper PTAs and important economic and policy outcomes; a number of initial research steps in this vein are described in more detail below.

Dispute settlement under PTAs In contrast to the WTO, there is very little empirical record of sustained and effective dispute resolution taking place under the major PTAs. With the exception of the European Union, dispute settlement provisions in most PTAs have rarely been used, and

---

19 Other recent contributions characterizing and assessing such PTA provisions include work by WTO Secretariat legal staff (Chase et al, 2013) and political scientists (Elsig and Allee, forthcoming). Note that the latter assess a larger coverage of dispute settlement provisions in preferential agreements in a publicly available “design of trade agreements” (DESTA) database (Dür, Baccini, and Elsig, 2014). See also WTO (2011).
their record of resolving disputes when they have been triggered is mixed at best. In fact, it is not uncommon for PTA dispute settlement procedures when they are used to generate third-country spillovers which lead to wider disputes that must be settled within the WTO dispute settlement system, or for disputes among PTA members to simply be taken to the WTO for resolution. We illustrate with two examples.

First, Tallberg and Smith (2014) report that very few (roughly 20) disputes were initiated under MERCOSUR between 1993 and 2005. But one high profile MERCOSUR dispute ended with Brazil imposing a new import restriction on retreaded tires from non-MERCOSUR partners but not on its MERCOSUR partners. Citing a similar MERCOSUR rule, Argentina imposed a new import restriction on footwear from non-MERCOSUR partners but not on its MERCOSUR partners. Because MERCOSUR rules apparently required that imports from MERCOSUR partners be exempted from the policies, Brazil’s and Argentina’s newly imposed import-restricting policies provided an additional implicit tariff preference to PTA partners relative to non-partners, and the resulting diversion of trade flows from third-countries led several non-MERCOSUR countries to challenge the MERCOSUR rule and ruling under WTO dispute settlement procedures.

Second, like MERCOSUR, NAFTA dispute settlement has rarely been triggered; e.g., fewer than 15 disputes were initiated under NAFTA between 1994 and 2010, and NAFTA dispute settlement fell into disuse after 2001 (Tallberg and Smith, 2014). But the small number of NAFTA disputes should not be interpreted as evidence that the PTA partners are living in harmony. Instead, it turns out that the three NAFTA partners (US, Canada, and Mexico) have taken more than twice as many disputes against one another to formal WTO dispute settlement since NAFTA’s inception than they have taken to the NAFTA forum. And some of these bilateral frictions - e.g., US-Canada trade in softwood lumber; US-Mexico trade in the related products of sugar, corn, high-fructose corn syrup, and ultimately soft drinks - actually started as formal NAFTA disputes but could not be resolved under the NAFTA forum. The disputes escalated and ultimately spilled over to resolution through formal WTO litigation.

---

20The EU has a different institutional design that includes a supra-national framework that initiates disputes against member states from within, and thus does not rely exclusively on the “state-to-state” framework of dispute resolution found in the WTO and many other PTAs. The dispute settlement provisions within the European Union have thus led to thousands of disputes. Tallberg and Smith (2014, p. 126) report that the supranational European Commission initiated more than 30,000 cases over 1978-2009 against its member states, with the area of greatest conflict being the internal market and environmental issues. Furthermore, the Commission only referred 11.5 percent of these initiated disputes to the European Court of Justice for a legal decision. On the other hand, EU member states have initiated only a handful of disputes against one other. Finally, the European Union’s free trade agreement with Iceland, Liechtenstein, and Norway under EFTA contains a supranational Surveillance Authority (SA) modeled similarly to the European Commission; Tallberg and Smith (2014, p. 138) report that the SA initiated roughly 400 disputes against the three EFTA member states over 1994-2008 and that the member states filed zero EFTA disputes against one another during this period.

21These disputes are described in greater detail in section 7.3 below.

2.3 The Choice

Summarizing, GATT/WTO membership is leading to a very different set of policy outcomes for its member governments than the outcomes that are emerging from the current wave of PTAs. While average MFN tariffs have been dramatically lowered through the process of GATT/WTO negotiations, MFN tariffs of WTO members have not been eliminated, and in certain industries in most countries and most industries in certain countries significant tariff peaks still exist, with the WTO membership unwilling or unable to negotiate further MFN liberalization. Furthermore, the WTO’s attempts to move even modestly beyond its traditional shallow integration approach have largely failed. Meanwhile, the action in international trade agreements has increasingly moved to PTAs, where remaining tariffs are being eliminated - albeit selectively and on a preferential basis among the member countries - and where deep integration appears to be the primary goal. Thus far, while the WTO and PTAs are clearly moving along different trajectories, the extent of their divergence may still be modest. But the “mega-regional” PTAs currently under negotiation, such as the Transatlantic Trade and Investment Partnership (TTIP) between the US and EU, or the Trans Pacific Partnership (TPP) between the US, Japan, and other Pacific trading partners, could alter this landscape dramatically.

In short, globalization is looking increasingly different under these two sets of rules, and it is therefore important for economists to adopt a reasoned and informed position as to which set of rules to support. Perhaps nowhere is this better illustrated by current affairs than the question whether the United States and the EU should throw their weight behind the WTO and a re-energized Doha Round, or rather put their efforts into negotiating new PTAs. Consider the likely differences in globalization’s outcomes depending on which of these strategies is pursued.

If the United States and the EU were to put their full support behind a successful Doha Round, even the most ambitious conclusion of the round would by all accounts entail relatively small further cuts in average tariffs and more substantial but still modest reductions in the remaining tariff peaks (and agricultural export subsidies). And any nod to deep integration (under the rubric of “trade facilitation”) would be very modest at best. After all, as Robert E. Hudec well describes in his depiction of the genesis of GATT’s shallow-integration approach, while governments understood that behind-the-border measures could have trade effects, the GATT (and to largely the same degree, the WTO today) never had its heart in deep integration:

“...The standard trade policy rules could deal with the common types of trade policy measure governments usually employ to control trade. But trade can also be affected by other ‘domestic’ measures, such as product safety standards, having nothing to do with trade policy.” [When GATT was created in 1947,]...governments would never have agreed to circumscribe their freedom in all these other areas for the sake of a mere trade agreement.” Hudec (1990)

Contrast Hudec’s description of the early GATT and our (most-ambitious) description of a successful WTO Doha Round with what has been leaked about the current ongoing major PTA initiatives
of the United States.

Consider first the TTIP negotiations between the United States and the EU. The TTIP has adopted as its main focus the streamlining of domestic standards across the Atlantic. The BBC puts it this way:

“Direct tariffs on goods and services between the two are already low, but there are other barriers such as regulatory and safety standards, inspection procedures, and preferences for domestic business. Removing these could significantly reduce the costs for companies doing transatlantic business.” BBC News (7/8/2013)

And as part of this streamlining, the TTIP would provide firms (as opposed to other governments) with legal rights to challenge government policies, a right that does not exist in the GATT/WTO. As the BBC reported:

“The European Commission says it will consult on part of a far-reaching EU-US free trade deal amid concern that hard-won social protections in Europe might be undermined. The trade negotiations began last year but now the Commission has launched a three-month public consultation on the proposed investment rules for firms. There are fears that they could give big firms too much power to block unfavourable government policies.” BBC News (1/22/2014)

Consider next the TPP that the United States is currently negotiating with a set of its Pacific trading partners. Here again the focus is on harmonizing domestic standards. As Marketplace Morning Report (1/28/2014) put it, “The Trans-Pacific Partnership has been called NAFTA on steroids.” The New York Times continues:

“If successful, the TPP agreement would eliminate most remaining tariffs on nearly $2 trillion in goods and services exchanged between the United States, Australia, Brunei, Canada, Chile, Japan, Malaysia, Mexico, New Zealand, Peru, Singapore and Vietnam. [But the TPP negotiations] “...would go far beyond lowering tariffs, with provisions requiring countries to maintain compatible regulatory regimes, facilitate corporate financial transactions, establish copyright and patent protections to govern intellectual property rights and to safeguard foreign investors.” NYTimes (2/5/2014)

And in a Report for Congress, the Congressional Research Service states:

“The TPP arguably provides the United States with the opportunity to project its trade interests by negotiating a ‘comprehensive and high-standard’ FTA with provisions that build off those in FTAs the United States concluded throughout the 2000s, especially the most recent ones...That, by itself is not new; the United States has often conducted asymmetrical negotiations with countries of differing levels of development in which it has dominated. This time, however, with more players at varied levels of
development, the United States may not be able simply to impose its vision or standards on those countries, and they are likely to make demands for concessions from the United States.” CRS Report for Congress (8/21/2013)

Evidently, the choice between the WTO and PTAs is a choice that is likely to matter. To the extent that economists’ opinions matter, it is important that they back the right one.

3 The Terms-of-Trade Theory

In this section we offer a brief review and empirical assessment of the essential predictions of the terms-of-trade theory of trade agreements, the oldest and most established strand of the trade agreements literature. According to the terms-of-trade theory, addressing an international externality (which travels through the terms of trade) is the central purpose of a trade agreement. More specifically, the theory posits that governments use trade agreements to undo the policy inefficiencies that arise with unilateral policy choices when those choices can shift the costs of intervention onto trading partners through movements in foreign exporter prices (terms of trade).

We begin with a presentation of the theory, which we develop in a benchmark model. We then review the evidence related to the central tenets of the theory. And finally, we compare at a broad level the nature of efficient policy commitments according to the terms-of-trade theory with the existing WTO policy commitments, and thereby use the terms-of-trade theory as a lens through which to offer an initial consideration of the possibility that the WTO membership might have already arrived at the international efficiency frontier.

3.1 Theory

To present the essential predictions of the terms-of-trade theory, we develop a benchmark two-country two-good perfectly competitive general equilibrium trade model. We first develop the model under the assumption that governments have only tariffs as instruments of policy intervention, and consider the purpose of a trade agreement. We then extend the benchmark model to allow that governments also have behind-the-border policy instruments at their disposal, and consider whether the purpose of a trade agreement is changed in this richer policy setting.

Trade model We assume that readers are familiar with the basics of the two-country two-good perfectly competitive general equilibrium trade model, and we focus here on presenting only the essential notation and key equations of the model. A more complete description of the trade model can be found in any undergraduate International Economics textbook (or see Bagwell and Staiger, 2010a, for a recent development of the model in the context of the trade agreements literature). The two countries, Domestic (no *) and Foreign (*), produce, consume and trade two goods, x and y, under conditions of perfect competition, with Domestic the natural importer of x and Foreign the natural importer of y. The relative price facing Domestic producers and consumers in their
local market is \( p \equiv p_x/p_y \), while the relative price facing Foreign producers and consumers in their local market is \( p^* \equiv p_{x^*}/p_{y^*} \). With non-prohibitive Domestic and Foreign ad valorem tariffs denoted respectively by \( t \) and \( t^* \), and with \( \tau \equiv (1+t) \) and \( \tau^* \equiv (1+t^*) \), international goods-market arbitrage then implies \( p = \tau p^w \equiv p(\tau, p^w) \) 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 \( 1/p^w \).

Production (as well as the distribution and level of factor incomes) is fully determined in each country by the local-market relative price in that country, while each country’s consumption depends on both the local-market relative price in that country and the terms of trade (with the latter, together with the local-market price, determining the tariff revenue collected by the country and distributed to its consumers). And each country’s trade is simply the difference between its consumption and production. Hence, for any local-market and world prices, Domestic imports of \( x \) can be written as \( M(p, p^w) \) and Domestic exports of \( y \) can be written as \( E(p, p^w) \), where for notational simplicity we have suppressed the good-specific subscript on the import and export functions. The analogous functions for Foreign are \( M^*(p^*, p^w) \) and \( E^*(p^*, p^w) \). And for any prices, the Domestic and Foreign national budget constraints are then given by the respective trade balance equations

\[
p^w M(p, p^w) = E(p, p^w), \quad \text{and} \quad M^*(p^*, p^w) = p^w E^*(p^*, p^w), \tag{1,2}
\]

with the equilibrium world price \( \tilde{p}^w(\tau, \tau^*) \) then determined by the market clearing condition for \( x \)

\[
M(p(\tau, p^w), p^w) = E^*(p^*(\tau^*, p^w), p^w), \tag{3}
\]

and with Walras’ Law ensuring that the \( y \)-market clears as well. Finally, with market-clearing local and world prices written as \( p = p(\tau, \tilde{p}^w) \), \( p^* = p^*(\tau^*, \tilde{p}^w) \) and \( \tilde{p}^w = \tilde{p}^w(\tau, \tau^*) \), we impose the following standard price assumptions to rule out the Metzler and Lerner Paradoxes:

\[
\frac{dp(\tau, \tilde{p}^w(\tau, \tau^*))}{d\tau} > 0 > \frac{\partial \tilde{p}^w(\tau, \tau^*)}{\partial \tau} \quad \text{and} \quad \frac{dp^*(\tau^*, \tilde{p}^w(\tau, \tau^*))}{d\tau^*} < 0 < \frac{\partial \tilde{p}^w(\tau, \tau^*)}{\partial \tau^*} \tag{4}
\]

According to (4), each country’s tariff is “protective” of its import-competing sector (i.e., the imposition of a tariff raises the local-market price of the import good), and each country is “large” in world markets and can improve its terms of trade with an increase in its tariff.

**Government preferences** When it comes to the goals of trade policy, real-world governments have diverse sets of preferences, in some cases adopting trade policies that would seem to promote aggregate national income while in other cases adopting trade policies with a clear distributional goal in mind. This diversity is reflected in the trade policy literature, where assumed government
preferences range from national income maximization (see Dixit, 1987, Johnson, 1953-54, Kennan and Reizman, 1988 and Mayer, 1981 for important formalizations, and Kowalczuk and Reizman, 2009, for a recent survey of this approach) to those of a representative democracy as reflected in the preferences of the median voter (see Mayer, 1984, for the initial formulation of this approach, and Dutt and Mitra, 2002, and Dhingra, forthcoming for important follow-up work) to those of a government influenced by lobbies (for early formalizations of this approach, see Olson, 1965, Caves, 1976, Brock and Magee, 1978, Feenstra and Bhagwati, 1982, Findlay and Wellisz, 1982, Hillman, 1982, and Baldwin, 1987; and see Grossman and Helpman, 1994 and 1995a, for the canonical treatment of this approach in the more recent literature).

The diversity of government preferences, both in the real world and in the formal trade policy literature, raises the question whether these preference differences across governments might translate into different purposes across the trade agreements that governments negotiate. To ensure that our answers regarding the purpose of a trade agreement are not dependent on adopting a particular formulation of government preferences from among this diverse set, we follow Bagwell and Staiger (1999a, 2002) and adopt a general “reduced from” 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, and we place no restrictions on a government’s preferences over its local-market prices. As local-market prices determine the level and distribution of factor incomes, this allows us to incorporate all of the formal models of trade policy motives mentioned above (which amount to specific models of how governments feel about changes in their local-market prices and the redistribution of factor incomes that this induces). We do impose one assumption on the government welfare functions, namely, that holding its local price fixed, the welfare of a government increases when its terms of trade improve:

$$W^{w}(p, \tilde{p}^{w}) < 0 < W^{*w}(p^*, \tilde{p}^{w}). \quad (5)$$

This assumption, which amounts to a statement that each government would like more tariff revenue if it could achieve this extra revenue without experiencing any change in its local-market price, is met by each of the formal models of trade policy determination that we mentioned above.\(^{23}\)

**Nash tariffs** In the absence of a trade agreement, the two governments choose their tariffs unilaterally and non-cooperatively, and we assume that these choices are characterized by an interior Nash equilibrium. The first-order conditions that define the Nash tariff choices are

$$\frac{dW(p, \tilde{p}^{w})}{dT} = W_{p} \frac{dp}{dT} + W_{\tilde{p}^{w}} \frac{\partial \tilde{p}^{w}}{\partial T} = 0, \quad \text{and}$$

$$\frac{dW^*(p^*, \tilde{p}^{w})}{dT^*} = W^*_{p} \frac{dp^*}{dT^*} + W^*_{\tilde{p}^{w}} \frac{\partial \tilde{p}^{w}}{\partial T^*} = 0. \quad (6)$$

\(^{23}\)We also assume sufficient concavity of $W$ and $W^*$ so that the second-order conditions for the optimization problems that we consider below are satisfied.
The top equation of (6) defines the Domestic government’s best-response tariff, while the bottom equation defines the Foreign government’s best-response tariff, with the Nash tariffs defined where both governments are on their respective reaction curves.

Notice that, as the two terms in the top equation of (6) highlight, in the Nash equilibrium the Domestic government strikes a balance between the effects on its welfare of the local-market-price and world-price movements induced by its tariff choice. The welfare implications of the local-market-price movement (the first term) are domestic in nature, reflecting the trade-off for the Domestic government between the benefits of any distributional changes (e.g., induced political support) and the costs of the induced economic distortions. By contrast, the welfare implications of the world-price movement (the second term) are international in nature, as 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, because an improvement in the Domestic country’s terms of trade corresponds necessarily to a deterioration in the Foreign country’s terms of trade. An analogous interpretation holds for the Foreign government at the Nash equilibrium, as the bottom equation of (6) highlights.

In the special case where governments are assumed to maximize national income with their unilateral tariff choices, (6) defines the standard (Johnson, 1953-54) “optimal tariff” for each country, which is simply the inverse of the trading partner’s export supply elasticity. As Johnson demonstrated, when governments seek to maximize national income, setting a tariff at this level is the optimal way for a country to exploit its monopoly power on world markets. For this case, on the margin the tariff creates costly distortions in the local market (the local-market-price movements in the first terms in (6)) but a portion of the costs of these distortions are borne by the trading partner (via the world-price movements in the second terms in (6)).\(^{24}\) For the more general cases of government preferences included in (6), the local-market-price movements carry additional welfare implications for the governments and this leads to Nash tariffs that will in general differ from the Johnson “optimal tariff” formula, but the logic of the trade-offs faced by each government in setting their unilaterally optimal tariffs is otherwise exactly the same.

Efficiency frontier If a trade agreement is to be useful to governments, there must be an inefficiency associated with the Nash tariff choices of the governments when evaluated with reference to their objectives. A trade agreement can then provide value to governments by correcting this inefficiency and increasing the welfare of both governments. Absent such an inefficiency, it would not be possible for a trade agreement to yield Pareto benefits for the governments involved. We therefore next characterize the efficiency frontier, so that we may assess whether Nash policies deviate from the frontier.

The efficiency frontier is defined by the set of tariffs that satisfy the familiar tangency condition

\(^{24}\) As is well known, for a national income maximizing government and for a positive tariff close to zero, the cost of the distortions in the local market is second order while the international cost-shifting associated with the induced terms-of-trade movement is first-order, which is why the Johnson optimal tariff is always strictly positive for a country that enjoys any market power on world markets (i.e., faces a finite foreign export supply elasticity).
Making use of the Domestic and Foreign government welfare functions $W(p, \bar{p}^w)$ and $W^*(p^*, \bar{p}^w)$, this tangency condition can be written as

$$\frac{\partial W_p}{\partial \tau} + \frac{\partial W_{\bar{p}^w}}{\partial \tau} = \frac{\partial W^*_p}{\partial \tau} + \frac{\partial W^*_{\bar{p}^w}}{\partial \tau}.$$  (7)

The characterization of the efficiency frontier provided by (7) is a generalization of the more familiar Mayer (1981) locus of efficient tariffs for the case of national-income maximizing governments. As is well-known, when governments maximize national income, reciprocal free trade ($\tau = 1 = \tau^*$) is efficient; yet as Mayer pointed out, this is but one point on the efficiency frontier defined by the locus of points $\tau = 1/\tau^*$ which ensure equality of the Domestic and Foreign local-market prices $p$ and $p^*$. Under the assumption that governments maximize national income and with the particular forms for $W(p, \bar{p}^w)$ and $W^*(p^*, \bar{p}^w)$ that this assumption implies, (7) reduces to the Mayer locus $\tau = 1/\tau^*$. With our more general government welfare functions, however, efficiency need no longer imply equality between $p$ and $p^*$, and (7) describes the required condition that must hold for efficiency in this more general environment.

With the efficiency frontier defined, it is now a simple matter to confirm that Nash policies are indeed inefficient in this model. Specifically, inspection of (6) confirms that (7) is not satisfied at Nash tariffs: by the top equation of (6), the denominator of the left-hand side of (7) is 0 and the numerator is strictly negative, implying that the left-hand side of (7) is $-\infty$ when evaluated at Nash policies; and by the bottom equation of (6), the numerator of the right-hand side of (7) is 0 and the denominator is strictly negative, implying that the right-hand side of (7) is 0 when evaluated at Nash policies.

Hence, according to the terms-of-trade theory, an inefficiency exists when governments set their tariffs non-cooperatively, and therefore there is a role for a trade agreement to address this inefficiency and improve the welfare of each government.

Interpreting the purpose of a trade agreement That Nash policies are inefficient according to the terms-of-trade theory is not surprising, especially in light of the way we have written the Nash first-order conditions in (6). After all, the cost-shifting motives that are embodied in the second term of each reaction curve in (6) impose a negative externality on the trading partner, pointing to an obvious source of inefficiency and suggesting that Nash tariffs might be too high relative to internationally efficient tariffs as a result of this motive. In fact, it can be confirmed that Nash tariffs are indeed higher than efficient tariffs according to the terms-of-trade theory.\(^{25}\) But recalling that our reduced form government preferences are specified in a way that is sufficiently general to capture all of the leading models of trade policy determination, it seems reasonable to expect that additional sources of inefficiency would also arise depending on which model of trade policy determination is relevant; and in this light it would be surprising if the cost-shifting externality were the only source of inefficiency that a trade agreement can correct in our model. Yet this is

---

\(^{25}\)See Bagwell and Staiger (1999a, 2002).
what the terms-of-trade theory implies.

To see this, let us suppose that the Domestic and Foreign governments were replaced by hypothetical governments that were not motivated by the terms-of-trade implications of their unilateral trade-policy choices; that is, let us consider a hypothetical Domestic government for which \( W_{p^*} = 0 \) and a hypothetical Foreign government for which \( W_{p^w} = 0 \). If these hypothetical governments were to select their tariffs unilaterally and non-cooperatively, then according to (6) their tariff choices would satisfy

\[
W_p(p, \tilde{p}^w) = 0 = W_{p^*}(p^*, \tilde{p}^w). 
\]

Following Bagwell and Staiger (1999a, 2002), we will refer to tariffs that satisfy (8) as politically optimal tariffs. If politically optimal tariffs are efficient when evaluated with reference to the actual Domestic and Foreign government preferences, then we may conclude that the terms-of-trade externality is the sole rationale for a trade agreement in this model. But it is immediate that (7) is satisfied when evaluated at tariffs that satisfy (8), and hence politically optimal tariffs are efficient when evaluated with reference to the actual Domestic and Foreign government preferences. We may therefore conclude that the terms-of-trade externality is the sole rationale for a trade agreement in this model.

Of course, as (7) suggests, there will in general be an entire locus of tariff combinations that satisfy the condition for efficiency, and the politically optimal tariffs represent only one point on this locus. But the politically optimal tariffs can be viewed as focal, as they remedy the terms-of-trade inefficiency in a direct way. For example, in the case where governments maximize national welfare we have already observed that efficient tariffs lie on the locus defined by \( \tau = 1/\tau^* \) as Mayer (1981) showed, and here the politically optimal tariffs correspond to the reciprocal free trade point \( \tau = \tau^* = 1 \), which seems focal in this case because it conforms to the trade policies that national income maximizing governments would have adopted in the first place if they were not motivated by terms-of-trade/”optimal tariff” considerations.

Having determined that the cost-shifting externality is the only source of inefficiency of the Nash policies according to the terms-of-trade theory, it is worth emphasizing now the role that the large-country assumption plays in establishing a purpose for a trade agreement in this theory. In a world of small countries where no country can impact its terms of trade with its tariff choices, no

---

26 An alternative way to interpret this thought experiment is the following: if each government agreed to delegate its tariff choices to an agent that shared its preferences over local-market prices but who did not value the terms-of-trade implications of its tariff choices, and if the agents then chose tariffs unilaterally and non-cooperatively, the politically optimal tariffs would be implemented and, as we have argued in the text, the governments would therefore achieve the efficiency frontier. In this sense, if a trade agreement has eliminated the terms-of-trade motives from unilateral tariff choices, (in our alternative thought experiment here by simply committing each government to delegate its tariff choices to an agent with its preferences over local prices but no concern for world prices), there is nothing further for the trade agreement to do.

27 Our claim that politically optimal tariffs are focal does not translate into a claim that a trade agreement will necessarily deliver the politically optimal tariffs. That depends, in part, on whether politically optimal tariffs lie on the contract curve, which in turn requires that countries not be too asymmetric. For example, in the special case of national income maximizing governments, it is well known that a sufficiently big country can “win the tariff war,” meaning that it is better off in the Nash equilibrium than at reciprocal free trade (see Johnson, 1953-54, and Kennan and Reizman, 1988).
country can engage in international cost-shifting and Nash policies will therefore be efficient, and there is nothing for a trade agreement to do. According to the terms-of-trade theory, then, the inefficiencies that a trade agreement can address are associated with the policies of governments that exercise market power on world markets. This is of course the same conclusion that Johnson (1953-54) drew, but that is in fact the striking point: the introduction of government preferences that can capture the wide diversity of government motives that we see in the real world does not qualify, complicate or change this conclusion.

When it comes to the purpose of a trade agreement, therefore, the terms-of-trade theory embodies a very simple idea. A trade agreement can be valuable to governments, but only if in the absence of an agreement governments would attempt to shift costs onto one another and as a consequence adopt inefficient unilateral policies. The terms-of-trade externality is simply the mechanism by which this cost-shifting occurs. But while simple and intuitive at one level, the practical relevance of this idea has traditionally met with deep skepticism among many economists. Some of this skepticism reflects a lack of empirical evidence relating to the central tenets of the theory, such as the degree and prevalence of market power that real-world governments can wield in international markets and whether their unilateral tariff choices reflect this market power when they possess it, and in a later section we survey the recent empirical work that is beginning to fill this gap. But some of the skepticism reflects a more visceral objection to the plausibility of the theory. Here we briefly consider three of the main objections of this kind.

A first objection is that the terms-of-trade theory unrealistically posits that governments seek to maximize national income with their tariff choices. But as we have just illustrated, while the terms-of-trade theory was originally posed by Johnson (1953-54) under this assumption, it holds equally well when governments have political motivations. A second objection is that there is a disconnect between the theory and the way real-world governments think. This objection is often expressed by the view that real-world governments almost never mention the “terms of trade” in their policy discussions, and instead the language of real-world trade agreements and the negotiators that craft such agreements emphasizes “market access.” But this disconnect is more apparent than real, because while we have developed the terms-of-trade theory as it is traditionally developed in the economics literature, it can be easily translated into the language of market access. Specifically, when a government raises its import tariff it shifts in its import demand curve, and 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. Using this link between price and volume effects, the terms-of-trade theory can then easily be recast using the market-access language that trade-policy negotiators favor.28 A third objection relates to the role played by tariff revenue in the theory’s account of the terms-of-trade motives of governments, and the apparent unimportance of tariff revenue to real-world governments. To this objection we offer two responses. First, as we describe in the next

28 Bagwell and Staiger (2002) provide a formal definition of market access and further develop the relationship between the terms-of-trade theory and the language of market access.
section, the cost-shifting motives at the center of the terms-of-trade theory do not hinge on government pursuit of tariff revenue. And second, it is not clear that tariff revenue should be seen as unimportant to real world governments. For example, as Kim (2013) points out, the tariff revenue collected by the United States for FY 2012 amounted to $31 billion, which is comparable to the amount that the United States spent in FY 2012 on foreign aid ($23 billion) and foreign military assistance ($14 billion) combined.

**Behind-the-border measures** We now briefly discuss an extension of the benchmark model in which governments also have behind-the-border policy instruments at their disposal, and consider whether the purpose of a trade agreement is changed in this richer policy setting. Our extended model broadly mirrors that of Bagwell and Staiger (2001).

Specifically, we introduce into the trade model of the previous section a Domestic standard, $\sigma$, and a Foreign standard, $\sigma^*$. To fix ideas we interpret these standards as a labor regulation (e.g., maximum legal work-hours per week) in each country that impacts that country’s production possibilities: for a given local-market price in a country, we assume that changes in its standard will shift its production of $x$ and $y$, and hence ultimately its import demand and export supply functions and therefore the market clearing world price. Proceeding to derive the market clearing world price as before, we therefore now have $p^w = p^w(\sigma, \sigma^*, \tau, \tau^*)$; and in addition to the assumed responses of $p^w$ to $\tau$ and $\tau^*$ contained in (4) above we now also assume that each country can also improve its terms of trade with an increase in its standard (i.e., we assume $\partial p^w / \partial \sigma < 0$ and $\partial p^w / \partial \sigma^* > 0$).

We assume that each government cares directly about its standard but does not care directly about the standard imposed in the other country. Hence, the choice of standards can impose international pecuniary externalities (through $p^w$) but by assumption standards do not impose non-pecuniary externalities at the international level. With this we can now express the government welfare functions in this extended model by $W(\sigma, p, p^w)$ and $W^*(\sigma^*, p^*, p^w)$. We continue to impose the structure in (5) on these extended government welfare functions, and we continue to leave unrestricted how governments care about changes in their own local-market prices; and we also leave unrestricted how governments care about changes in their own standards.

There are now four first-order conditions that define the Nash tariff and standards choices of the Domestic and Foreign governments:

$$
\frac{dW(\sigma, p, p^w)}{d\tau} = W_p \frac{dp}{d\tau} + W_{p^w} \frac{\partial p^w}{\partial \tau} = 0, \quad (9)
$$

$$
\frac{dW(\sigma, p, p^w)}{d\sigma} = W_\sigma + [\tau W_p + W_{p^w}] \frac{\partial p^w}{\partial \sigma} = 0,
$$

$$
\frac{dW^*(\sigma^*, p^*, p^w)}{d\tau^*} = W_{p^*} \frac{dp^*}{d\tau^*} + W_{p^w} \frac{\partial p^w}{\partial \tau^*} = 0, \text{ and}
$$

$$
\frac{dW^*(\sigma^*, p^*, p^w)}{d\sigma^*} = W_{p^*} + \left[\frac{1}{\tau^*} W_{p^*} + W_{p^w} \right] \frac{\partial p^w}{\partial \sigma^*} = 0.
$$

In addition to the two conditions defining the Domestic and Foreign best-response tariffs, which
are unchanged from our earlier discussion and which therefore have the same interpretation, there
is an additional condition for each government in (9) that defines its best-response standard. Here
each government weighs the direct impact on its welfare of its standards choice against the impact
of this choice on its welfare that runs through the induced local-market- and world-price effects.

Turning to the efficient policies in this extended setting, there are now three tangency conditions
that must be satisfied for efficiency:

\[
\frac{[\tau W_p + W_{p}^w]}{W_p} \frac{\partial p^w}{\partial \tau} + \frac{W_p^*}{\partial \tau} + \frac{W_{p}^w}{\partial \tau} = \frac{1}{\tau} \frac{W_p^*}{\partial \tau} + \frac{W_{p}^w}{\partial \tau},
\]

The top condition in (10) is the same tangency condition as in (7) that defined the efficiency frontier
in our earlier benchmark model. This condition can be interpreted as ensuring that the Domestic
and Foreign tariffs are set so that the volume of trade between the two countries is at an efficient
level.\textsuperscript{29} The remaining two conditions can then be interpreted as ensuring that each country’s mix
of border and behind-the-border policies delivers in an efficient way this efficient level of trade
volume.

Finally, we may again define politically optimal policies and make use of these hypothetical
constructs to investigate the purpose of a trade agreement in this extended setting. We therefore
consider a hypothetical Domestic government for which \( W_{p} = 0 \) and a hypothetical Foreign
government for which \( W_{p}^* = 0 \). If these hypothetical governments were to select their tariffs
and standards unilaterally and non-cooperatively, then according to (9) their tariff and standards
choices would satisfy

\[
W_p(\sigma, p, \tilde{p}^w) = 0 = W_\sigma(\sigma, p, \tilde{p}^w), \quad \text{and}
\]

\[
W_{p}^*(\sigma^*, p^*, \tilde{p}^w) = 0 = W_{\sigma}^*(\sigma^*, p^*, \tilde{p}^w).
\]

As before, if these politically optimal policies are efficient when evaluated with reference to the
actual Domestic and Foreign government preferences, then we may conclude that the terms-of-trade
externality is the sole rationale for a trade agreement in this extended model. But it is immediate
that (10) is satisfied when evaluated at tariffs and standards that satisfy (11). Hence, politically
optimal tariffs and standards are efficient when evaluated with reference to the actual Domestic
and Foreign government preferences, and we may conclude that the terms-of-trade externality is
the sole rationale for a trade agreement in this extended model.

Therefore, regardless of the extent of behind-the-border measures that governments may have at
their disposal, according to the terms-of-trade theory the purpose of a trade agreement remains the

\textsuperscript{29}See Bagwell and Staiger (2001) for elaboration on this interpretation and the remaining interpretations of the
efficiency conditions that we describe in the text.
same. Governments use trade agreements to undo the policy inefficiencies that arise with unilateral policy choices when those choices can shift the costs of intervention onto trading partners through movements in foreign exporter prices (terms of trade).

But we may now also emphasize a further insight offered by the terms-of-trade theory that follows from its stance on the purpose of a trade agreement. Specifically, as the only “problem” for a trade agreement to “solve” is terms-of-trade manipulation, and as the tariff is the first-best policy for terms-of-trade manipulation, the terms-of-trade theory implies that the tariff is in fact the only policy that is distorted in the Nash equilibrium: behind-the-border measures are set efficiently under Nash choices. This insight is lurking in the conditions presented above, and it can be confirmed by noting that the top two conditions of (9) imply the middle condition of (10) while the bottom two conditions of (9) imply the bottom condition of (10). Hence, the only efficiency condition that is not met in the Nash equilibrium according to the terms-of-trade theory is the top condition of (10) that determines the efficient level of tariffs (which are lower than their Nash levels) and the efficient level of trade volume (which is higher than the Nash level). In this sense, the terms-of-trade theory provides a foundation for trade agreements that adopt a shallow approach to integration.30

Our formal discussion here has focused on a particular form of behind-the-border measure that is best thought of as a production standard, such as a workplace regulation or a regulation on the use of an open access resource in the production process. So it is important to note that the points we have emphasized apply more broadly to behind-the-border measures of various kinds, including tax and subsidy policies and various forms of standards beyond production standards.31 Of particular relevance to the world trading system are product standards, such as minimum burn-through rates for doors or prohibitions on lead additives to paint, that can raise the costs of supplying a market but, unlike import tariffs, do not raise revenue. In light of the prominent role played by tariff revenue in our account of the terms-of-trade motives of governments described above, it might be thought that those motives do not apply to such standards. But our discussion applies equally well to these kinds of behind-the-border measures, once it is understood that international cost-shifting occurs when such product standards are imposed as long as foreign exporters do not pass the full cost of meeting the product standards on to consumers in the country where the standard applies.32

3.2 Evidence

We now survey the evidence related to the central tenets of the terms-of-trade theory. We focus on three basic questions. First, how significant and widespread is the market power that countries

30 More specifically, the terms-of-trade theory supports an approach to behind-the-border measures that revolves around reliance on a “market access preservation” rule. As we explain further below, this insight survives well beyond the simple setting that we consider here. See for example, Bajona and Ederington (2011), who explore these issues in a setting where trade agreements must be self-enforcing and where behind-the-border measures are privately observed, and who show that the optimal agreement then takes the form of a market access preservation rule combined with a binding tariff ceiling.

31 For an extension of these points to the case of domestic tax/subsidy policies, see Bagwell and Staiger (2006).

32 See Staiger and Sykes (2011) for a treatment of product standards in this context.
possess in world markets? Second, do the unilateral tariff choices of countries reflect the market power that they possess? And third, does the pattern of negotiated tariff liberalization that we observe correlate with the pattern of observed market power in the way that the theory suggests it should? Answers to these questions seem central to all of the theory’s predictions, and so we focus on them here, postponing until later points in our survey a discussion of the empirical work relating to various other predictions of the terms-of-trade theory.

**Market power** Do countries routinely possess a degree of market power on world markets that would allow them to manipulate their terms of trade? To answer this first question, we begin by observing that there is a large body of indirect evidence that suggests an affirmative answer. We are referring here to the literature on exchange rate pass-through.

To establish the connection between the exchange rate pass-through literature and an answer to our first question, we proceed in two steps. First, we note that Feenstra (1989) shows theoretically that the pass-through to domestic prices associated with exchange rate shocks can be thought of as comparable in magnitude to the pass-through associated with tariff changes, and he offers econometric evidence supporting this hypothesis of symmetric pass-through between tariffs and exchange rates in the data. Second, we note that, while we have adopted above a general equilibrium setting to present the terms-of-trade theory, the theory can also be developed in a partial-equilibrium model, where cost shifting then occurs through changes in the terms of trade provided that foreign exporters bear some of the incidence of the import tariff so that it is not fully passed through to domestic prices. That is, incomplete pass-through of the tariff to domestic prices is synonymous with cost-shifting and the exercise of importer market power. Hence, by the second step, the terms-of-trade effects of a tariff arise whenever the incidence of the tariff is not fully passed through to domestic prices; and by the first step, we may look to the vast empirical literature on exchange rate pass-through for indirect evidence about the degree of tariff pass-through.

What does the exchange rate pass-through literature imply for the answer to our first question? First, in light of the conclusion from Goldberg and Knetter’s (1997) survey of this literature that pass-through rates average about 60%, it would appear that the existence of countries that possess substantial market power on world markets is routine. And second, after surveying this literature in his Handbook chapter, Feenstra (1995, p. 1569) concludes that, while there is strong evidence that pass-through is less than complete, the magnitude of pass-through differs substantially across industries and therefore “...we should not have any presumption about the extent of terms of trade gain due to tariffs, but must treat each industry on a case-by-case basis.”

Turning to the direct evidence on this first question, in a provocative paper Magee and Magee (2008) construct measures of world market concentration and trade elasticities to argue that even a “large” country like the United States has little market power to exert on world markets, suggesting that market power can safely be ignored when considering the effects of tariffs. But for the most part, the literature has produced results which are consistent with the conclusions from the indirect

---

33See Bagwell and Staiger (2001b) for a development of the terms-of-trade theory of trade agreements in the context of a partial equilibrium model.
evidence emphasized above. For example, the studies of Kreinin (1961), Winters and Chang (2000), Chang and Winters (2002), Anderson and VanWincoop (2002), and Bown and Crowley (2006), among others, all offer evidence that unilateral tariff changes can significantly affect a country’s terms of trade. On the other hand, in their study of New Zealand’s unilateral trade liberalization of the 1980’s, Winkelmann and Winkelmann (1998) find only weak evidence of incomplete pass-through of New Zealand tariffs to prices in the New Zealand economy, and only in a few industries, and conclude that New Zealand truly is an example of a small country in most products.

Taken together, these studies seem to confirm the basic conclusion of Feenstra (1995), that there is abundant evidence of substantial market power but its presence varies greatly from industry to industry and country to country. This conclusion is further reinforced by the most comprehensive study to date on the issue, that of Broda, Limão and Weinstein (2008). Focusing on the unilateral tariff choices of 15 countries prior to their membership in the GATT/WTO, Broda et al. estimate the foreign export supply elasticities faced by each of these countries, which as they note provides an inverse measure of the degree of market power that each of these countries was able to exert on the foreign export (world) prices. With these estimates, they confirm that most countries, even apparently “small” countries, have significant ability to alter their terms of trade on many imported products with their tariff choices.

A final piece of evidence on this first question is institutional, and relates to the widely acknowledged fact that trade agreements are driven by exporter interests. In particular, it is probably safe to say that the vast majority of tariff reductions agreed to by each GATT/WTO member government came about because exporters in some other GATT/WTO member country pushed their government to negotiate those tariff cuts. And the most direct explanation for why these exporters cared about having those particular tariffs reduced in the first place is that those tariffs were suppressing their prices: that is, these exporters were bearing some of the incidence of those tariffs, which is precisely the cost-shifting impact of market power that the terms-of-trade theory indicates should trigger a tariff negotiation.34 In this light, the widespread efforts of exporters to convince their governments to negotiate particular tariff reductions from particular trading partners is itself evidence of a world brimming with importer-government market power.

Unilateral tariffs We next turn to the second question raised above: Do the unilateral tariff choices of countries reflect the market power that they possess?

An early paper that provides evidence on this question is Olarreaga, Soloaga and Winters (1999). Examining the determinants of the common external tariff adopted by MERCOSUR in 1996 (a period when the common external tariff of MERCOSUR was essentially unconstrained

---

34 We point out here that the inference we draw in the text follows from the MFN tariff reductions agreed to by GATT/WTO members, but it would not follow from the preferential tariff cuts that governments might agree to in their PTAs even if those agreements were also exporter driven, as they surely are. The reason is that a preferential tariff cut can raise the price of foreign exporters to whom the preference is given even while having no effect on “world” prices outside the PTA, by simply permitting the preferred exporters to supply the local market at the importer’s local-market price (which will be higher than the world price if the importer is maintaining a strictly positive MFN tariff).
by WTO commitments), they conclude that terms-of-trade effects account for a substantial part (between 6 and 28 percent) of the explained variation in the structure of MERCOSUR tariffs, despite the fact that during the period of their analysis MERCOSUR had only a 1 percent market share of world trade.

This question is also addressed by the Broda, Limão and Weinstein (2008) paper discussed above. Specifically, after estimating the foreign export supply elasticities faced by each of the 15 non-GATT/WTO member countries in their sample, Broda et al relate this measure of the power to affect world prices to the unilateral (i.e., prior to WTO accession) tariff choices that each country made. They find that, prior to joining the WTO, these countries set tariffs an average of 9 percentage points higher on imports for which they could exert large effects on world prices as compared to the tariffs they set on imports where their ability to affect world prices was limited, an impact whose magnitude is roughly comparable to the size of the average tariffs in these countries. And 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 Olarreaga et al. (1999) and Broda et al. (2008) findings, governments who set their trade policies unilaterally and noncooperatively respond to terms-of-trade motives and the market power that they possess strongly and in the way that the theory predicts. These findings are reinforced by the recent paper of Dhingra (forthcoming). Dhingra shows that the median-voter model of non-cooperative tariff determination is strongly rejected by the data in its traditional small-country formulation, but that its central predictions receive strong cross-country empirical support once it is cast in a large-country setting and the existence of market power and terms-of-trade motives are taken into account when predicting the cross-country pattern of tariffs.

**Negotiated tariffs** Finally, we consider the third of our three questions: Does the pattern of negotiated tariff liberalization that we observe correlate with the pattern of observed market power in the way that the theory suggests it should? As with the literature that takes up this question, we focus below on the pattern of negotiated tariff liberalization in the GATT/WTO rather than PTAs. We consider the pattern of liberalization that has taken place within PTAs and its compatibility with the terms-of-trade theory in the context of later discussions.

In light of the literature’s finding surveyed just above that market power is a strong determinant of unilateral tariff choices, one way to pose this third question is to ask whether the same can be said for negotiated tariffs: if market power is also a strong determinant of negotiated tariffs, then this would indicate a lack of support for the view that trade agreements serve to limit the exercise of market power by their member governments; on the other hand, if negotiated tariffs are unrelated to market power, then together with the findings reported in the literature that unilateral tariffs are strongly related to market power, this would suggest that trade agreements do indeed limit the exercise of market power. This is the approach that Broda et al. (2008) take in the paper we have already discussed at several points above. Focusing on the United States, they find that US non-tariff barriers and so-called “statutory” tariff rates – neither of which has been the subject
of 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 US MFN tariffs – which have been subjected to the many rounds of GATT/WTO negotiations – exhibit no such relationship.

A related approach is taken in the paper by Nicita, Olarreaga and Silva (2014). They focus on the nature of the tariff commitments made by WTO member countries – commitments that as we have described above take the form of “bindings” defining the maximum allowable level for the tariff – and exploit the fact that countries differ in the degree to which their negotiated WTO tariff commitments actually constrain their “applied” tariffs, that is, the tariff levels that they actually set. Olarreaga et al. observe that tariffs that are unconstrained by WTO bindings should exhibit a positive correlation with market power, for the reasons associated with unilateral tariffs that we have described above. But Olarreaga also derive a new prediction: they argue that the tariffs that are constrained by WTO bindings (the “cooperative tariffs”) should exhibit a negative correlation with market power. This prediction follows intuitively under their assumptions that (a) exporters enjoy extra political economy weight in the objectives of their governments and (b) their governments lack trade instruments of their own (e.g., export subsidies) to shift surplus to these exporters. Under these assumptions, the only way for a government to help its exporters is to negotiate a tariff cut in the foreign markets served by its exporters, and the political payoff to the government from negotiating such tariff cuts will be higher the greater is the market power in the foreign market to which it focuses its negotiating efforts. On this basis, they predict that the cooperative tariff levels will be negatively correlated with importer market power. Examining the tariffs of 101 WTO members, Nicita et al. find that the sign of the correlation between tariff levels and market power indeed switches from positive to negative as the WTO tariff bindings vary from levels that are well above applied tariff rates to levels that are at the applied tariff rates. \[35\]

Beshkar, Bond and Rho (2014) derive a related prediction in an extension of the basic terms-of-trade model that emphasizes the trade-off between commitments and flexibility (we discuss their model and paper further at later points in our survey). Within this extended model, they show that applied tariffs are more likely to be set at their WTO-bound levels at any point in time the greater is the market power for that product possessed by the importing government; and with a sufficiently high level of import market power, applied tariffs will always be set at the level of the negotiated binding. Beshkar et al. find strong evidence for these predicted relationships between negotiated tariff commitments and importer market power with data on the tariffs of 109 WTO member countries.

A different approach to this question is taken by Bagwell and Staiger (2011). They use the terms-of-trade theory to derive an expression for the component of the noncooperative tariff that embodies the international cost-shifting motive. They then use this expression to derive the pattern

\[35\] Nicita, Olarreaga and Silva (2014) also report that the expected positive correlation between tariffs and market power emerges in their data only once the WTO tariff commitments are sufficiently above the applied tariff levels. They view this as something of a puzzle, and suggest that it may reflect some form of implicit cooperation among WTO member countries.
of negotiated tariff cuts that is implied by the terms-of-trade theory; intuitively, if the cost-shifting component is big, then the negotiated tariff cut implied by the terms-of-trade theory should be large. Working from this basic insight, Bagwell and Staiger show that the tariff cuts predicted by the terms-of-trade theory can be expressed as a function of pre-negotiation import volumes and prices and measures of the power to affect world prices. The predicted relationship 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.

Ludema and Mayda (2013) advance the literature on this question by allowing that free-rider issues associated with the MFN principle (which we discuss at length further below) might impede governments from fully addressing terms-of-trade manipulation in their GATT/WTO tariff negotiations. Augmenting the basic terms-of-trade model with a model of endogenous bargaining participation, they show that negotiated tariffs should be negatively related to the product of the importer’s market power and exporter concentration. Intuitively, the lower is exporter concentration, the more severe will be the free-rider problem associated with MFN, and the less effective negotiations will be in ridding the tariff of its cost-shifting component; and the larger the market power, the larger this cost-shifting component will be. Ludema and Mayda take this prediction to the tariff data for 30 WTO-member countries and find strong support, concluding (p. 1837) that “...the internalization of terms of trade effects through WTO negotiations has lowered the average tariff of these countries by 22% to 27% compared to its noncooperative level.”

Yet a different perspective on answering this question is provided by Bown and Crowley (2013a), who investigate empirically some of the predictions of the terms-of-trade theory when that theory is developed in a repeated-tariff-game setting subject to stochastic trade volume shocks and where self-enforcement constraints are binding. A basic prediction of the terms-of-trade theory in this setting is that, where the self-enforcement constraint is binding, cooperative tariffs must rise somewhat in the face of import surges in order to mitigate the increased temptation implied by the higher import volume to shift costs onto foreign exporters with an even higher tariff. Among the additional predictions of the model highlighted by Bown and Crowley is that, for a given import surge, a cooperative tariff increase is more likely on a product if the market power possessed by the importing government on that product is higher. Utilizing data on the time-varying protective measures (anti-dumping and safeguard actions) of the United States over the period 1997-2006, the authors confirm these and other predictions of the terms-of-trade theory in this setting.

We close our survey of the evidence by returning once more to an institutional perspective. Specifically, notice that the empirical relevance of the terms-of-trade theory does not require that all countries are large in all products. But it does imply that the large players in the market should

36 Consistent with Ludema and Mayda’s (2013) finding that GATT/WTO negotiations have substantially reduced but not eliminated market-power considerations from its members’ tariffs, Dhingra (forthcoming) also finds a weaker but still positive terms-of-trade component in the tariffs of WTO-member countries as compared with the tariffs of non-members.

37 Bagwell and Staiger (1990) develop this model, which we will describe further along with further discussion of the findings of Bown and Crowley (2013a) at later points in our survey.
be the most active participants in any particular negotiation. When viewed from this perspective, the broad 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.

When viewed together, the empirical papers surveyed above provide a growing body of evidence consistent with the central tenets of the terms-of-trade theory of trade agreements. And while we will discuss below evidence that lends some support as well to the other strands of the trade agreement literature, the literature we have reviewed here suggests that, at a minimum, the terms-of-trade theory is central for understanding actual trade agreements. We therefore continue to rely on the terms-of-trade theory and the literature that builds from it as we seek answers to our survey’s motivating questions, and we then turn to the remaining strands of the literature and describe the qualifications to the answers provided by the terms-of-trade theory that this literature indicates must be introduced.

3.3 Has Globalization under the WTO Gone Far Enough?

With the broad features of the terms-of-trade theory of trade agreements described, it is now possible to pose a first and basic question. Viewed from the lens of this theory, might it be possible to conclude that globalization under the GATT/WTO has already allowed its member governments to reach the international efficiency frontier? This is a question that we will need to confront in the ensuing pages in the process of answering the motivating questions of our survey. But we pause here to dispel one natural misconception, namely, that an immediately plausible answer to this question might be “No, as long as trade is not truly free.” Here the discipline imposed by the terms-of-trade theory requires a more nuanced answer. After all, according to this theory, the purpose of a trade agreement is not to secure free trade, but to remove the internationally inefficient cost-shifting component from the unilateral tariff choices of its member governments. As we have observed, this outcome could be compatible with free trade, but only under certain strong assumptions about the preferences of governments, assumptions which most economists would agree are not likely to be met in the real world. Therefore, according to the terms-of-trade theory, the continued existence of tariffs is not by itself evidence of further work to be done.

A stark example of the terms-of-trade theory’s implications in this regard is the case where the Nash protection level is prohibitive and there is hence no trade volume in the Nash equilibrium. In this case of extreme protection, and with its trade volume driven to zero, the importing government does not enjoy a cost-shifting benefit from the reduction in world prices that its protection has caused (because it imports no volume at these lower world prices), and so according to the terms-of-trade theory this Nash tariff choice must in fact reflect underlying features of the government’s preferences rather than inefficient cost-shifting motives. And for this reason, the terms-of-trade theory implies that autarkic Nash trade policy choices are internationally efficient, and as a conse-

31
quence implies that there is nothing for a trade agreement to accomplish in the presence of such choices.

The takeaway point is that when interpreted through the lens of the terms-of-trade theory of trade agreements, our observation above, that the MFN tariffs of WTO members have for the most part not been eliminated and in certain industries in most countries and most industries in certain countries significant tariff peaks still exist, does not mean it is obvious that the GATT/WTO has left important efficiency-enhancing tariff liberalization unfinished. Put slightly differently, from the lens of the terms-of-trade theory, it is not self-evident that the attack on remaining tariffs being waged by PTAs has moved the world closer to the international efficiency frontier, even setting aside the preferential nature of the PTA tariff liberalization. To know whether globalization under the WTO has gone far enough requires, according to the terms-of-trade theory, a more nuanced assessment of the kinds of theoretical and empirical issues addressed by the literature we survey below.

Much as the continued existence of tariffs cannot by itself be interpreted as a sign of lingering inefficiencies in the GATT/WTO system, neither can one interpret the lack of negotiated GATT/WTO commitments on behind-the-border measures as a sure sign of GATT/WTO failings according to the terms-of-trade theory. Indeed, as is suggested by our discussion of behind-the-border measures above and as we will highlight further below, the GATT/WTO “shallow integration” approach is compatible with international efficiency in the presence of certain kinds of accompanying rules, rules that we will suggest at a broad level find representation in the GATT/WTO. Again with regard to behind-the-border measures, to know whether globalization under the WTO has gone far enough requires, according to the terms-of-trade theory, a more nuanced assessment.

Finally, we emphasize two important points that are implicit in our discussion here and should be kept in mind as we proceed through the survey. First, for the most part the terms-of-trade theory of trade agreements treats government preferences – where as we have described above these preferences can reflect in part the differing political systems in each country – as fixed and sovereign, much as consumer preferences are taken as sovereign in standard consumer theory, and seeks to understand trade agreements as agreements that lead to Pareto improvements for the member governments when gauged from the perspective of their own preferences. This is not the only possible assumption, and below we will discuss the few papers in this strand of the literature that entertain alternative assumptions. But the assumption does resonate well with the “member driven” nature of real-world trade agreements, especially the GATT/WTO. As we will later discuss, the commitment theory strand of the trade agreements literature can be interpreted as adopting a major departure from this assumption, and so we will consider these alternatives in more depth when we survey that strand of the literature.

A second point is that the terms-of-trade theory (and indeed the other strands of the trade agreements literature as well) takes the instruments of intervention that a government has at its disposal as fixed and given. Hence, for example, while it is well-known that a tariff is a second-best instrument for virtually every domestic policy goal that could be imagined, the trade agreements...
literature starts from the view that there are often constraints (usually unmodeled) that prevent governments from using the first-best instruments for their policy goals and lead them to use tariffs to achieve these goals instead. And given that these governments are constrained to use tariffs for, e.g., distributional reasons, the trade agreements literature then explores the role of trade agreements in eliminating international inefficiencies from their tariff choices. This feature is important to keep in mind, as without it the notion that the use of tariffs could ever be compatible with a position on the international efficiency frontier (beyond tariffs that reside on the Mayer locus) would seem strange.

4 Evaluating the GATT/WTO approach to trade liberalization

We now turn to an evaluation of the GATT/WTO approach to trade liberalization from the perspective of the terms-of-trade theory of trade agreements literature. This strand of the literature seeks to evaluate the design and performance of the GATT/WTO as an institution that could plausibly help governments in their attempts to solve the terms-of-trade externality problem that we described in Section 3. We use our survey of this literature to ask: Is the GATT/WTO well-designed at a fundamental level to help governments address the terms-of-trade problem? And can the successes and failures of the GATT/WTO record be broadly understood to reflect the strengths and weaknesses of these design features when interpreted within the context of the terms-of-trade problem? If the answers to these questions are affirmative, then we tentatively conclude that the GATT/WTO is an institution worthy of the continued support of economists. This is not to say that PTAs might not also be worthy of support – we take up that question in the next section – but rather to say that the GATT/WTO appears fundamentally to be well-designed to solve the terms-of-trade problem and is therefore an institution that should not be allowed to whither.

Notice that our discussion just above presupposes that in the context of solving problems of international trade policy cooperation institutions matter in the first place, and as is well-known this implies in turn that we are at least implicitly assuming that important frictions are present. This is an assumption shared by the literature that we survey, though the literature is not always explicit about the underlying frictions that would imply the relevance of institutional features for outcomes, a fact that Maggi (forthcoming) emphasizes as a weakness of the existing literature and an important direction for the literature’s further development. We are sympathetic to Maggi’s critique of the literature on this point, and yet we also believe (as does Maggi) that valuable insights can be learned from the literature by simply taking on faith that institutions do matter for solutions to this problem. That institutions matter in this context seems abundantly evident from the historical experience with international trade policy cooperation, and the many failed attempts at international cooperation that preceded the creation of GATT.39

38 We discuss some partial exceptions in the literature at later points in our survey, such as Limão and Tovar (2011), but even there the availability of first-best instruments are ruled out by assumption.
39 For more on the historical antecedents of GATT and the many failures of international cooperation that helped to inform GATT’s design, see Bagwell and Staiger (2010a) and the literature cited therein.
We organize this section into six subsections. Section 4.1 considers the GATT pillars of reciprocity and the most-favored nation (MFN) clause, and the general self-enforcing nature of GATT/WTO commitments. Sections 4.2 through 4.5 survey the literature on more specific features of the GATT/WTO approach that are central to identifying current challenges. Throughout we highlight a number of significant challenges, some brought on by the recent wave of globalization and others long-standing, including the “latecomer” problem associated with integrating major emerging economies such as China into the trading system, and the treatment of subsidies. A final subsection summarizes by offering an initial terms-of-trade-theory-based perspective on how the GATT/WTO has worked, its current woes and possible fixes. We will revisit this perspective again at later points in our survey, first when we add in consideration of PTAs as a possible source of GATT/WTO woes and then again when we consider the non-terms-of-trade theories of trade agreements.

4.1 Reciprocity, MFN and enforcement

In this subsection, we review research based on the terms-of-trade approach that examines the three pillars of the GATT/WTO architecture: reciprocity, MFN and enforcement.

Reciprocity  Governments reach agreements to lower tariffs through GATT/WTO negotiation rounds. It is often observed that these negotiations reflect a norm of reciprocity and entail a “balance of concessions,” whereby each government makes the “concession” of lowering its import tariffs in exchange for receiving the benefit of a similar concession from a trading partner. Such a norm would be difficult to interpret in a model with small countries. The reciprocity norm, however, admits a straightforward interpretation when countries are large.

To develop this point, we return to the terms-of-trade model described in Section 2. Let us suppose that the two governments start at the Nash tariffs. We know from our earlier discussion that these tariffs are inefficiently high, due to the terms-of-trade externality. Since each government selects its best-response tariff at a Nash equilibrium, no government can gain from a unilateral tariff cut. This simple observation provides an immediate interpretation for why a government regards its own tariff cut as a concession. Governments may enjoy mutual gains, however, if they jointly reduce tariffs. Indeed, as Bagwell and Staiger (1999a, 2002) show, when governments start at the Nash tariffs, \((\tau^N, \tau^*N)\), they can mutually gain from moving to a new pair of tariffs \((\tau, \tau^*)\) only if they both offer strictly lower tariffs, so that \(\tau < \tau^N\) and \(\tau^* < \tau^*N\). A general form of reciprocity is thus necessary for mutual gains.

We may also interpret reciprocity at a more specific level. Following Bagwell and Staiger (1999a, 2002), let us associate the principle of reciprocity with mutual changes in trade policy that bring about changes in the volume of each country’s imports that are of equal value to changes in the volume of its exports. Formally, for given initial and subsequent tariff pairs, \((\tau^0, \tau^0)\) and \((\tau^1, \tau^1)\), respectively, a set of tariff changes \(\Delta\tau \equiv \tau^1 - \tau^0\) and \(\Delta\tau^* \equiv \tau^1 - \tau^0\) satisfies the principle of
reciprocity if

\[ \tilde{p}^{w0}[M(p^1, \tilde{p}^{w1}) - M(p^0, p^{w0})] = [E(p^1, \tilde{p}^{w1}) - E(p^0, p^{w0})], \tag{12} \]

where \( \tilde{p}^{w0} \equiv \tilde{p}^{w}(\tau^0, \tau^{w0}), \tilde{p}^{w1} \equiv \tilde{p}^{w}(\tau^1, \tau^{w1}), p^0 = p(\tau^0, p^{w0}), p^1 = p(\tau^1, \tilde{p}^{w1}) \) and trade-volume changes are valued at the initial world price, \( \tilde{p}^{w0} \). In this two-good model, as Bagwell and Staiger (1999a, 2002) establish, a set of tariff changes satisfies the principle of reciprocity if and only if the world price is unchanged so that \( \tilde{p}^{w0} = \tilde{p}^{w1} \). This finding may be easily confirmed by applying (1) to (12), where (1) is applied at both the initial and subsequent tariff pairs.

Is liberalization according to the principle of reciprocity sufficient for mutual gains? To address this question, we start at the Nash tariffs as characterized in (6) and consider for simplicity the preferences of the home government. Given (4) and (5), it is straightforward to see that \( W_p < 0 \) when tariffs are at their Nash levels. Thus, at Nash tariffs, the home government would prefer more trade (a lower relative price of imports) if it could achieve a greater trade volume without inducing a loss in the home country’s terms of trade. A unilateral tariff cut delivers greater trade volume for the home country, but it does so at the cost of a deterioration in the home country’s terms of trade. By contrast, mutual changes in trade policy that satisfy the principle of reciprocity enable the home country (and similarly the foreign country) to enjoy greater trade volume without suffering a terms-of-trade loss. Thus, and as Bagwell and Staiger (1999a, 2002) show, starting at the Nash equilibrium, trade liberalization that satisfies the principle of reciprocity is sure to raise each government’s welfare, at least initially.\(^{40}\)

This finding offers an interpretation for the reciprocity norm that guides trade-liberalization negotiations in GATT/WTO.

While more evidence is needed, recent empirical work offers some support for the important role played by reciprocity in GATT/WTO tariff negotiations. In an early effort, Shirono (2004) considers the Uruguay Round and finds that the negotiated tariff reductions conform well with the reciprocity norm. Limão (2006, 2007) considers US tariff cuts in the Uruguay Round and also reports evidence consistent with the application of reciprocity. After constructing a measure of market-access concessions and identifying instruments that address possible endogeneity concerns, he presents evidence of reciprocity for US products that were not subject to non-tariff barriers. Among such goods, US import tariff reductions embody a general form of reciprocity, being greater for goods exported from countries whose own tariff cuts provided greater market access to US exports. Karacaoglu and Limão (2008) provide similar support for reciprocity in a related exercise for the EU. Finally, we also note that evidence of reciprocity may be stronger in some product groups than in others. Gulotty (2014) focuses on tariff liberalization by the US in the Uruguay Round and reports evidence that sectors with high contract-intensive products, characteristic of global production, exhibit the least reciprocity.

The principle of reciprocity is a key pillar of the GATT/WTO approach to trade liberalization.\(^{40}\)In a symmetric setting, if governments were to liberalize according to the principle of reciprocity and move thereby from the Nash equilibrium to the politically optimal tariffs, then they would enjoy mutual gains all along this path. In an asymmetric setting, if governments were to liberalize according to the principle of reciprocity, then mutual benefits may disappear before the efficiency frontier is reached. In both settings, mutual gains are enjoyed provided that the liberalization is not too deep.
Our discussion to this point considers the principle of reciprocity as a bargaining norm that characterizes the manner in which tariffs are reduced in GATT/WTO negotiations. The principle of reciprocity, however, also explicitly arises in GATT/WTO rules that govern the manner in which tariffs may be raised as part of renegotiation or dispute resolution procedures. In Section 7, we provide further discussion of reciprocity in the context of renegotiation and retaliation procedures.

**Most-Favored Nation (MFN) treatment** A second pillar of the GATT/WTO architecture is the principle of non-discrimination. For member countries, this principle requires that a country apply the same tariff on a given import good, regardless of which country exports the good. This principle is embodied in the most-favored nation (MFN) rule, under which no exporter of a given good can be treated any less favorably than any other exporter.

To interpret and evaluate the principle of non-discrimination, we extend the model in a simple fashion to allow for three countries. As before, the home country imports good \( x \) and exports good \( y \). We now assume, however, that there are two foreign countries. Foreign countries 1 and 2 each trade with the home country, importing good \( y \) and exporting good \( x \), and we assume for simplicity that the two foreign countries do not trade with each other. The local price in the home country is again represented as \( p = \frac{p_x}{p_y} \), while the local price in foreign country \( i \), where \( i = 1, 2 \), is given by \( p^i = \frac{p_{xi}}{p_{yi}} \). The world price for trade between the home country and foreign country \( i \) is denoted as \( p^{wi} = \frac{p_{xi}}{p_y} \) and represents the terms of trade for foreign country \( i \). Let \( t^i \) represent the home-country ad valorem tariff applied to imports from foreign country \( i \) and let \( t^{si} \) denote the ad valorem tariff of foreign country \( i \) applied to imports from the home country. Defining \( \tau^i = 1 + t^i \) and \( \tau^{si} = 1 + t^{si} \), we then have from arbitrage conditions that \( p^i = \frac{p^{wi}}{\tau^{si}} \) and \( p = \tau^1 p^{w1} = \tau^2 p^{w2} \).

From here, we see that if the home country adopts discriminatory tariffs, defined by \( \tau^1 \neq \tau^2 \), then different world prices, \( p^{w1} \neq p^{w2} \), obtain for its two foreign trading partners. The foreign country that receives the lower import tariff enjoys a better terms of trade. If instead the home country satisfies the principle of non-discrimination (i.e., the MFN rule), defined by \( \tau^1 = \tau^2 \equiv \tau \), then \( p^{w1} = p^{w2} = p^w \) follows, and so the two foreign countries enjoy the same terms of trade, \( p^w \). Under the MFN rule, the home country’s terms of trade are given as \( 1/p^w \).

Our next steps are to determine market-clearing prices and represent each government’s welfare function. We begin with the simplest case, where the home country’s tariffs satisfy the MFN rule. In that case, as noted, a single world price obtains. The market-clearing world price, \( \bar{p}^w(\tau, \tau^1, \tau^2) \), may then be determined similarly to how we do so above for the two-country model. With the market-clearing local prices then given as \( p(\tau, \bar{p}^w) = \tau \bar{p}^w \) and \( p^i(\tau^{si}, \bar{p}^w) = \frac{\bar{p}^w}{\tau^{si}} \), we may represent the respective government welfare functions as \( W(p, \bar{p}^w), W^1(p^1, \bar{p}^w) \) and \( W^2(p^2, \bar{p}^w) \). Following Bagwell and Staiger (1999a, 2002), we may now again impose two key assumptions: each country can improve its terms of trade by raising its import tariff (i.e., \( \partial \bar{p}^w(\tau, \tau^1, \tau^2)/\partial \tau^{si} > 0 > \partial \bar{p}^w(\tau, \tau^1, \tau^2)/\partial \tau \)); and each government prefers an improvement in its terms of trade, holding fixed its local price (i.e., \( W_{\bar{p}^w}(p, \bar{p}^w) < 0 < W_{p^{si}}^i(p^i, \bar{p}^w) \)).

Analogous steps apply as well for the case in which the home country utilizes discriminatory
tariffs, although in that case the home country’s terms of trade is more complex to define. As Bagwell and Staiger (1999a, 2002) discuss, when the home country’s tariffs are discriminatory, the home country’s multilateral terms of trade is defined as a weighted average of its bilateral terms of trade, $1/p^{w1}$ and $1/p^{w2}$, where the weights are export shares and thus depend on foreign local prices. It follows that international externalities associated with foreign import tariffs then travel to home’s multilateral terms of trade via foreign local-price channels as well as bilateral terms-of-trade channels. Intuitively, when the home government employs discriminatory tariffs, it cares not just about the bilateral terms of trade but also about the share of exports that comes from each partner, as it enjoys greater tariff revenue when a higher share of a given import volume comes from the partner on which the import tariff is highest. The formal counterpart to this intuition is that the home country enjoys an improved multilateral terms of trade when a greater share of its imports comes from the partner on which it places the highest import tariff.

With the three-country model described, we now consider the value of the principle of non-discrimination. A simple observation is already at hand: when the MFN rule is used, international externalities associated with tariff choices are channeled through the world price alone, just as in the two-country model. A key implication is that the main findings presented above for the two-country model carry over as well to the multi-country model when the MFN rule is imposed. Specifically, as Bagwell and Staiger (1999a, 2002) show, if home-country tariffs are non-discriminatory, then Nash tariffs are inefficient, politically optimal tariffs are efficient, and a small liberalization between the home country and a foreign trading partner that starts at the Nash tariffs and that satisfies the principle of reciprocity is sure to generate mutual gains for the two negotiating governments. The latter finding suggests a broad complementarity between the principles of reciprocity and non-discrimination: the latter principle ensures that international externalities travel through the world price, which is a channel that the former principle is well-equipped to manage.

The multi-country model also allows consideration of sequential negotiations between trading partners. The possibility of sequential negotiations is easily motivated in the trade-agreement context. First, within a given negotiation round, some bilateral negotiations may precede others, suggesting that strategic considerations associated with sequential negotiations may come into play. Second, liberalization efforts in the GATT/WTO system have occurred over more than 65 years in the context of eight different negotiation rounds, and so negotiations among a given set of countries naturally occur through time across different negotiation rounds. Finally, accession dates differ across GATT/WTO members, so that some countries participating in a given negotiation round may not have participated in earlier rounds. Indeed, the first GATT negotiation round involved 23 countries, whereas the WTO now has nearly 160 members. To capture new insights associated with the sequential nature of negotiations, we assume that the home government initially negotiates with

---

41 For analyses of the MFN rule in other modeling frameworks, see Bagwell and Staiger (2001b), Choi (1995), McCalman (2002) and Saggi (2004), for example. Bagwell and Staiger (2010a) and Horn and Mavroidis (2001) offer further discussion of research on the legal and economic aspects of the non-discrimination principle.

42 By contrast, due to the presence of local-price externalities, politically optimal tariffs are not efficient when discriminatory tariffs are used.
the government of foreign country 1 and subsequently negotiates with the government of foreign country 2. A key question is whether the principle of non-discrimination may be interpreted as having efficiency-enhancing implications in this sequential context.

We begin by putting this question in broader context and highlighting potential inefficiencies that the MFN rule may introduce. As Caplin and Krishna (1988) emphasize, the MFN rule is a restriction on the set of instruments and thus shifts in the bargaining frontier; hence, an efficiency-enhancing role for non-discrimination is available only if the bargaining process in the absence of the MFN rule delivers inefficient outcomes. Furthermore, in the context of sequential negotiations, it is commonly argued that the MFN rule may lead to a “free-rider” problem, whereby a country refrains from offering significant concessions since it expects anyhow to enjoy MFN tariff cuts from trading partners undertaking their own negotiations. As Caplin and Krishna (1988) argue, the MFN free-rider concern suggests that bargaining under the MFN rule may fail to deliver efficient outcomes even relative to the MFN-constrained efficiency frontier. Ludema (1991) shows, however, that an MFN-efficient bargaining outcome is possible, if participating countries have the option to reject an agreement and continue bargaining when another country free rides and refuses to cut its own tariffs. For bargaining in a given round, Ludema’s finding suggests that the threat of delayed agreement may be an important defense against free riding. His model, however, may be less well suited for bargaining that occurs over time across negotiating rounds, particularly in the presence of acceding countries, as then negotiations in early rounds may be undertaken with a view toward endogenously affecting the outside options for acceding countries in later rounds.

We now consider sequential bargaining in the three-country model, where the first-stage bargain between the governments of the home country and foreign country 1 determines \((\tau^1_1, \tau^*1)\) and the second-stage bargain between the governments of the home country and foreign country 2 determines \((\tau^2_2, \tau^*2)\). Our first step is to identify a bargaining inefficiency that arises in the absence of the MFN rule. As Bagwell and Staiger (2005a) show, under a slight strengthening of the assumptions given above, any proposed efficient tariff vector is vulnerable to bilateral opportunism: by appropriately lowering the second-stage tariffs that they apply to one another, the governments of the home country and foreign country 2 can enjoy mutual welfare gains that come about at the expense of the government of foreign country 1. Intuitively, such second-stage tariff reductions lead to a terms-of-trade loss for foreign country 1, and for two reasons. First, the discriminatory tariff cut that the home government offers stimulates export supply in foreign country 2 and thus generates downward pressure on the world price of foreign country 1’s export good. Second, the tariff cut offered by the government of foreign country 2 induces greater demand for the home export good and thus generates upward pressure on the world price of foreign country 1’s import good.

In the absence of the MFN rule, bilateral opportunism in the second-stage bargain thus ensures an inefficient outcome. The prospect of second-stage bilateral opportunism may feed back and limit the scope of the first-stage bargain as well. If the government of foreign country 1 were to foresee that the value of any first-stage concession obtained from the home government might be eroded by an even greater concession extended to foreign country 2, then the government of foreign country 1
might be cautious in extending its own first-stage concession. In light of the inevitable bargaining inefficiencies that arise in the absence of the MFN rule, we move now to the second step of our discussion and examine whether the MFN rule addresses the bilateral opportunism problem and thereby promotes efficiency. As Schwartz and Sykes (1997) argue, it is natural to expect that the MFN rule could be helpful in this regard. After all, the concession-erosion concern is addressed under the MFN rule, since foreign country 1 is then assured of receiving any home tariff cut that is subsequently offered to foreign country 2. In terms of our discussion in the preceding paragraph, the MFN rule addresses the first reason for the terms-of-trade loss of foreign country 1. Unfortunately, however, the MFN rule does not address the second reason. Even if the home tariff satisfies the MFN rule, foreign country 1 may suffer a terms-of-trade loss due to the tariff cut extended by foreign country 2. Building on this point, Bagwell and Staiger (2005a) show that a subset of the tariffs that are efficient within the MFN class are also vulnerable to bilateral opportunism, even when the MFN rule is imposed. The principle of non-discrimination is thus helpful but not completely effective in eliminating the scope for bilaterally opportunism.

Is there a simple rule which, in combination with the MFN rule, fully eliminates the scope for bilateral opportunism? In fact, as Bagwell and Staiger (2005a) show, bilateral opportunism is impossible if any bilateral negotiation must satisfy both the principle of non-discrimination and the principle of reciprocity. As in the two-country model, under the principle of reciprocity, a negotiation between the governments of the home country and foreign country 2 maintains the world price between these countries. In addition, as argued above, the principle of non-discrimination implies that foreign countries 1 and 2 have a common terms of trade, \( \bar{w} \). It follows that a bilateral negotiation between the governments of the home country and foreign country 2 that satisfies the principles of non-discrimination and reciprocity leaves unaltered foreign country 1’s terms of trade. Since foreign country 1’s tariff is not altered, foreign country 1’s local price, \( p^1 = \bar{w}/\tau^1 \), also remains unchanged. We thus arrive at the following welfare-preservation result: if the governments of the home country and foreign country 2 engage in a bilateral negotiation that satisfies the principles of non-discrimination and reciprocity, then government welfare in foreign country 1, \( W^*(p^1, \bar{w}) \), is preserved. Interestingly, this result implies as well that the MFN rule does not generate a free-riding incentive, when bilateral negotiations also abide by the principle of reciprocity.

To see the intuition behind with the welfare-preservation result, let us suppose that the governments of the home country and foreign country 2 enter into a bilateral negotiation in which they lower their respective tariffs. Under the MFN rule, exporters from foreign country 1 then face a lower home import tariff, which in isolation provides improved access to the home market. But the impact of foreign country 2’s tariff cut also must be considered. In foreign country 2, this tariff cut lowers the local price of the import good relative to the export good, and thus both stimulates consumption of the import good and production of the export good. For both of these reasons, foreign country 2’s export volume increases. Hence, and as the Lerner symmetry theorem would suggest, foreign country 2’s import tariff cut has the same effect as would an export subsidy increase. Thus, while exporters in foreign country 1 enjoy a lower home import tariff, they
also face, in effect, subsidized competing exporters from foreign country 2. Under the principles of reciprocity and non-discrimination, these effects exactly balance out, with the increase in home import demand exactly fulfilled by the expanded export volume from foreign country 2. Exporters in foreign country 1 thus ultimately do not gain additional access to the home market, which is to say that the bilateral negotiation has no impact on foreign country 1’s terms of trade.

Bagwell and Staiger (2010b) provide further analysis of sequential bargaining in the three-country model. They assume that the home government negotiates sequentially with the two foreign governments, where the home government makes take-it-or-leave-it offers and negotiations are over MFN tariffs as well as lump-sum international transfers. The assumption that lump-sum international transfers are feasible is extreme but ensures analytical tractability. As they show, the MFN rule alone is then completely ineffective in addressing bilateral opportunism: for any proposed MFN-efficient outcome, the governments of the home country and foreign country 2 can adjust the tariffs and transfers under their control so as to enjoy mutual gains that come at the expense of the government of foreign country 1.

In addition to this “backward-stealing problem,” they identify a “forward-manipulation problem”: the home government may keep its MFN tariff inefficiently high in its initial negotiation, and thus engage in “foot-dragging,” in order to endogenously generate a less attractive outside option (i.e., disagreement point) for its subsequent negotiation partner. Hence, governments generally would be unable to achieve an efficient outcome, even if the backward-stealing problem were addressed. Bagwell and Staiger (2010b) argue, however, that efficient outcomes may be reached when the MFN rule is combined with other bargaining rules. The forward-manipulation problem, in particular, may be addressed if the opportunities for renegotiation are so “sweeping” as to disconnect the initial negotiation outcome from the outside option in the subsequent negotiation. At the same time, we note that other problems may arise if the renegotiation option is too readily available, as then the significance of any GATT/WTO tariff commitment would be put in question.

We turn now to the empirical evidence that concerns the non-discrimination principle. A first form of evidence considers the trade-volume impacts associated with GATT/WTO membership and relates these impacts to the theoretical implications developed above. Subramanian and Wei (2007) find that GATT/WTO membership is associated with large and significant trade-volume impacts for developed countries but that the trade-volume impacts of membership are small or nonexistent for most developing countries. Since developed countries have been the most active participants in GATT/WTO rounds, one interpretation of this finding draws from the welfare-preservation result presented above. In particular, if developed countries negotiate tariff reductions that broadly adhere

---

43 Limão (2007) explores a related forward-manipulation strategy. In his model, a government engages in foot dragging in order to influence a future bargain involving a preferential trade agreement with non-trade objectives. We provide further discussion of preferential trade agreements in Section 5.

44 The uneven trade effects of GATT/WTO membership found by Subramanian and Wei (2007) overturn the findings of an earlier paper by Rose (2004) that constrained the trade effects of GATT/WTO membership to be the same for all countries and found no membership effect at all. Evidence that the trade effects of GATT/WTO membership are restricted primarily to developed countries has subsequently been confirmed by a number of papers (see, for example, Chang and Lee, 2011, and Eicher and Henn, 2011, though Eicher and Henn attribute to PTAs most of the trade effects that Subramanian and Wei attributed to WTO membership).
to the principles of reciprocity and non-discrimination, then the trade-volume impacts on third-party countries should be limited. As Bagwell and Staiger (2014) discuss, an implication of this interpretation of relevance for the Doha Round is that substantial trade-volume gains for developing countries from negotiated tariff reductions may be achieved most effectively if, in markets where they are large, developing countries negotiate reciprocally with each other and with developed countries. This implication runs counter to the non-reciprocal approach for developing countries in the Doha Round, as codified under “special and differential treatment” clauses. A second interpretation of the Subramanian-Wei (2007) finding is that developed countries have managed to circumvent the MFN rule and discriminate against non-participating GATT/WTO members. Further empirical analysis of this interpretation is certainly warranted. We note, however, that in the specific context of GATT/WTO bilateral dispute settlement negotiations, Bown (2004a) finds evidence that countries comply with the MFN rule.

A second form of evidence relates GATT/WTO negotiated import tariff reductions to industry and country characteristics. Ludema and Mayda (2009, 2013) relate negotiated bilateral tariff reductions to measures that capture foreign exporter concentration and importer market power. In particular, Ludema and Mayda (2013) find evidence that the level of the importer’s tariff resulting from negotiations is negatively related to the product of two terms: exporter concentration (as measured by the Herfindahl-Hirschman index) and the importer’s market power (as measured by the inverse elasticity of export supply, on a product-by-product basis). The positive effect of market power on a country’s MFN tariff thus diminishes as exporter concentration increases, which is consistent with the view that negotiations are especially effective in “undoing” terms-of-trade-driven inefficiencies when exporter concentration is large. The important role played by exporter concentration supports the existence of an MFN free-rider effect. They also provide evidence that the free-rider effect is quantitatively important: they estimate that between 12 and 25% of potential liberalization on average goes unrealized, with most of this gap falling on the exports of developing countries due to their low-concentration product mix. Their findings thus also suggest a novel concentration-based interpretation of the Subramanian-Wei (2007) finding.

**Enforcement and repeated-game models** Our preceding discussion assumes that a negotiated trade agreement can be enforced. While this abstraction is helpful for interpreting and evaluating the principles of reciprocity and non-discrimination, the manner in which a trade agreement is enforced is also essential to understand. According to the terms-of-trade theory of trade agreements, the strategic environment confronting governments setting trade policy corresponds to a Prisoners’ Dilemma setting. In a symmetric, two-country and two-good model, for example, if governments were to attempt to cooperate with a common tariff strictly below the Nash level, then each government would have an incentive to “cheat” by raising its tariff to the best-response level. What deters governments from cheating in this way? Since there is no GATT/WTO jail or other external enforcement device, a government cooperates in a trade agreement if and only if the government perceives that such behavior is in its self interest. In other words, and as argued by
McMillan (1986, 1989), Dixit (1987) and Bagwell and Staiger (1990), the theory of repeated games suggests that a trade agreement must be self-enforcing so that for each government the short-term gain from cheating is smaller than the long-term discounted cost of any consequent breakdown in cooperation.

We highlight here three implications of this repeated-game perspective. First, this perspective suggests an interpretation of the decision by initial GATT contracting parties to concentrate protective measures, with certain exceptions, into tariffs. As Bagwell and Sykes (2004) argue, by “tariffying” quantitative restrictions such as quotas, governments facilitate mutually beneficial and reciprocal trade liberalization. Specifically, by imposing tariffs rather than allocating quotas across foreign exporters, governments make market-access gains easier to assess and thereby reduce negotiation transactions costs, reduce the uncertainty facing perhaps risk-averse exporters and thus enhance the value of market-access concessions, and increase the transparency of trade-policy conduct so that cheating is less tempting. The latter point, in particular, corresponds to the familiar notion from repeated-game theory that cooperation is typically easier to achieve when players’ actions are publicly observable or transparent.  

As Maggi (1999) shows, a second and related implication of the repeated-game perspective is that a trade agreement may facilitate self-enforcing cooperation by enhancing third-party transparency. Maggi distinguishes between bilateral and multilateral enforcement mechanisms. Under a bilateral enforcement mechanism, if country A deviates with respect to the tariff that it applies to country B, then countries A and B revert to a Nash trade war. Countries A and B continue to cooperate with country C, however. By contrast, under a multilateral enforcement mechanism, if country A deviates with respect to the tariff that it applies to country B, then country A reverts to a Nash trade war with both countries B and C. Maggi shows that a multilateral enforcement mechanism can achieve greater cooperation than is possible under bilateral enforcement mechanism. Correspondingly, when a trade institution ensures that any deviation would be observed by all member countries, the future cost of cheating could be quite severe indeed, which in turn implies that more efficient tariffs can be achieved without violating the self-enforcement constraint.

A third implication of the repeated-game perspective is that novel predictions may be generated when the self-enforcement constraint binds and political or economic shocks occur. Intuitively, when a change in the environment upsets the balance between the short-term incentive to cheat and the long-term discounted value of cooperation, an adjustment in trade policy may be required to bring the self-enforcement constraint back into balance. We develop this point further in the next subsection when we interpret GATT/WTO safeguard rules.

A final point is that our discussion here concerns retaliation that occurs off the equilibrium path in the repeated game. The role of such retaliation is to induce compliance; however, off-equilibrium-

45 An additional and important advantage that tariffs have over quantitative restrictions on trade is that the latter may be more difficult to apply and enforce on a non-discriminatory basis. See, e.g., Curzon (1965, p. 130) for an early discussion of this issue.

46 As Furusawa (1999) shows, the repeated-game approach also suggests that greater patience is not always beneficial for a country in trade-agreement negotiations, when the negotiation outcome also must be self-enforcing.
path retaliation is, by definition, not predicted by the model and thus should be distinguished from retaliation that actually appears in WTO rules and practice. At various points below, we note that on-equilibrium-path retaliation consistent with WTO rules and practice emerges naturally in the model once the repeated-game model is extended to include privately observed shocks.

**Implications of basic WTO principles and rules** This subsection has summarized research on reciprocity, non-discrimination and enforcement, all from the perspective of the terms-of-trade theory. On the whole, our discussion suggests that the principles of reciprocity and non-discrimination are well-designed for addressing the inefficiencies associated with the terms-of-trade externality. The theory also provides a natural means for interpreting the self-enforcing nature of trade agreements. At the same time, our review directs attention to significant challenges that confront the GATT/WTO approach. The possibility of third-party externalities warrants particular attention, whether such externalities are positive (and raise free-riding concerns) or negative (and raise bilateral-opportunism concerns). The theory reviewed above suggests that third-party externalities are eliminated when tariff policies adhere to the principles of reciprocity and non-discrimination. The extent to which negotiated tariff cuts satisfy the principle of reciprocity may differ somewhat across market settings, however, and GATT/WTO rules allow for important exceptions to the MFN rule, with the most important exception being the provisions for preferential trading agreements given in GATT Article XXIV. Addressing the potential for third-party externalities thus remains an on-going challenge for the GATT/WTO.

In this context, our review encourages consideration of the interaction between potential free-rider benefits and the benefits of GATT/WTO membership for developing countries. As we discuss, recent empirical work indicates that many developing countries have not experienced significant trade-volume benefits from GATT/WTO membership. The research reviewed above delivers two possible interpretations for this pattern. First, if free-rider benefits are significant, then countries may perceive modest gains from pursuing trade-liberalization negotiations in general but perhaps especially with developing countries whose exporters often operate in unconcentrated industries. Second, if free-rider benefits are largely eliminated, then developing countries may benefit little from pursuing a non-reciprocal approach to trade liberalization. Together, these interpretations suggest that the benefits of GATT/WTO membership for developing countries may be enhanced if negotiated tariff reductions adhere closely to the principles of reciprocity and non-discrimination and if, in markets where they are large, developing countries negotiate reciprocally with each other and with developed countries.

### 4.2 Bindings, Binding Overhang and Safeguards/Contingent Protection

A fundamental design feature of the GATT/WTO is that governments negotiate “tariff bindings” or “bound tariff levels” rather than exact tariff levels. For a given country and good, a bound tariff, which is also referred to as a “tariff cap,” identifies the maximal import tariff that can be applied. A tariff cap thus permits “downward flexibility,” in that a government can apply a tariff
that is strictly below the tariff cap. In that event, “binding overhang” is said to occur. But a
tariff cap constrains “upward flexibility,” since under normal circumstances the applied tariff cannot
exceed the cap. Some potential for upward flexibility does exist within the GATT/WTO agreement,
however, when certain contingencies arise. In this subsection, we review economic research that
interprets tariff caps, binding overhang, and contingent protection or “safeguards.”

We develop our discussion in the context of a standard partial-equilibrium model with two
countries and three goods, where the two countries are symmetric. One of the goods is the
import good for the home country, while another good is the import good for the foreign country.
The third good is a standard numeraire good, which is produced in both countries under constant
returns to scale where labor is the only factor in the model. For the two non-numeraire goods,
production occurs in each country under conditions of perfect competition and with diminishing
marginal productivity. The resulting supply functions are upward sloping. We assume further
that the consumers in both countries have a common utility function, which is additively separable
and takes a quasi-linear form. The consumption of the numeraire good exhibits constant marginal
utility while the consumption of each of the non-numeraire goods exhibits diminishing marginal
utility. The latter property delivers downward sloping demand functions for the non-numeraire
goods. Finally, each country has available a tariff for its import good, and we assume that trade
in the numeraire good is untaxed.

**Tariff bindings and tariff caps** With this standard model in mind, we now consider the in-
terpretation of tariff caps. At the time of negotiation, we imagine that governments are in an “ex
ante” stage, with each government being uncertain about the importance that it will place in the
future on profits in the import-competing sector relative to consumer welfare. After government
preference shocks are realized, tariffs are applied in a manner consistent with the negotiated trade
agreement. In this “ex post” stage, a government’s optimal unilateral tariff and also the efficient
tariff are higher when the government’s “type” (i.e., the welfare weight that the government at-
taches to import-competing profits) is larger. Due to the terms-of-trade externality, the optimal
unilateral tariff is higher than the efficient tariff for any government type, provided only that the
efficient tariff does not prohibit all trade. In this context, the challenge is to design a trade agree-
ment that permits some flexibility, so that applied tariffs may respond to preference shocks and
thereby facilitate greater efficiency, without opening the door to opportunistic tariff hikes.

To fix ideas, it is helpful to consider a couple of extreme scenarios in which governments in fact
can design a trade agreement in the presence of preference shocks that achieves full efficiency (i.e.,
an efficient tariff in each state). First, if governments’ realized types were publicly observed and
verifiable, and if there were no contracting costs, then governments could write a “state-contingent”
contract that delivers full efficiency. Second, even if each government’s type were privately observed,

---

47 The phrase “water” or “water in the binding” is sometimes used interchangeably with the phrase “binding
overhang.”

48 Symmetry here means that the two countries are “mirror images” of one another. This model is frequently used
in trade-policy research. See, for example, Bagwell and Staiger (2001a).
standard arguments establish that governments could again achieve full efficiency if a lump-sum transfer instrument were available.⁴⁹ These scenarios are extreme, however. Governments are likely to have some private (or at least unverifiable) information about their preferences, and explicit monetary transfers are not required in GATT/WTO rules and are rarely used in WTO dispute resolution.⁵⁰

Motivated by these considerations, we therefore turn now to a scenario in which governments negotiate a trade agreement under uncertainty, have private information about their respective preferences at the time that tariffs are applied, and do not have available an instrument with which to effect contingent transfers. Allowing for a continuum of possible government types that are distributed uniformly, Bagwell and Staiger (2005b) consider this scenario in a linear-quadratic model and compare two possible trade agreements.⁵¹ In the first agreement, governments adopt a rigid tariff rule, under which each government commits to an exact tariff level for all types. Since a government’s type enters its welfare function in a linear fashion, the most efficient agreement of this kind places the rigid tariff at the level that is efficient for the average type. They then compare this agreement with one in which each government adopts a weak-binding tariff rule, consistent with GATT/WTO rules, under which it commits to a tariff cap. They find that the most efficient weak-binding yields strictly higher expected joint welfare than does the most efficient rigid tariff, and that the most efficient weak binding is strictly higher than the most efficient rigid tariff. Intuitively, under a weak binding, the applied tariff exhibits binding overhang when a government draws a low type, and so the binding only constrains the applied tariff when the government’s type is high and the efficient tariff is thus also high.

Amador and Bagwell (2013) generalize this analysis in several directions. Allowing for more general payoff and distribution functions, they derive conditions under which a trade agreement with a tariff cap (i.e., a weak-binding tariff rule) maximizes expected joint welfare among all incentive-compatible trade agreements. They thus provide a first theoretical explanation for the use of tariff caps in an optimal trade agreement. Their approach is to represent the problem of finding an optimal trade agreement for the import good of the home country as a delegation problem, in which the principal’s expected welfare corresponds to the associated expected joint welfare of the two governments and the agent is the government of the home country.⁵² In this context, a trade agreement identifies a set of permissible tariffs. In the ex ante stage of trade-agreement design, the two governments thus choose the set of permissible tariffs for the home-country import good that maximizes ex ante joint welfare while satisfying the incentive-compatibility constraint that the home government will choose its preferred tariff from this set after observing its type. The tariffs induced by rigid-tariff rules and weak-binding tariff rules are of course candidate solutions, but

---

⁴⁹See Bagwell and Staiger (2005b) for details in the trade-agreement context.
⁵⁰Further, even monetary transfers may entail inefficiencies, due to the distortions associated with raising funds through taxation.
⁵¹In a linear-quadratic model, economic agents have quadratic payoffs and the corresponding demand and supply functions are linear.
⁵²A delegation game is a principal-agent game in which the agent has private information and transfers are infeasible. The delegation game was first introduced and analyzed by Holmstrom (1977).
so, too, are many discontinuous tariff functions. Amador and Bagwell (2013) further allow that the trade agreement may specify that tariff choices are bundled with money-burning expenditures perhaps corresponding to administrative procedures. They characterize a rich set of environments, which includes the linear-quadratic model with a uniform distribution as a special case, in which the optimal trade agreement takes the simple form of a tariff cap. The optimal trade agreement then also exhibits binding overhang and does not employ money burning.

Related results arise in other settings as well. Amador and Bagwell (2013) move beyond the partial-equilibrium model with perfect competition and use their main propositions to establish conditions under which a tariff cap is also optimal for a monopolistic-competition model of trade with a fixed number of firms. Amador and Bagwell (2012) similarly employ these propositions to provide conditions for the optimality of a tariff cap in a linear-quadratic model with a uniform distribution when private information concerns the weight that tariff revenue receives in the government welfare function. Finally, the models discussed here assume a continuum of possible types. While this specification seems natural, it is not innocuous. As Bagwell (2009) confirms, the optimal trade agreement does not take the form of a tariff cap in the linear-quadratic model when government preferences concerning the relative importance of import-competing firms are drawn from two possible types.

Beshkar, Bond and Rho (2014) extend the linear-quadratic model to a setting with asymmetric country sizes. They restrict attention to tariff caps and provide theoretical and empirical support for the prediction that the optimal tariff cap is higher, and thus the likelihood of binding overhang is greater, for smaller countries. To gain some intuition for their theoretical findings, it is helpful to consider the extreme case of a small country. Since the tariff policy of a small country imposes no terms-of-trade externality on its trading partner, the optimal trade agreement would permit such a country to impose its unilaterally optimal trade policy for whatever preference shock it experiences. A high (i.e., unrestrictive) tariff cap achieves this goal. More generally, when a country is larger, its tariff policy imposes a greater externality on its trading partner, and so a lower tariff cap, with an associated reduced frequency of binding overhang, is optimal. Using data on applied and bound tariffs for WTO member countries, they then provide strong empirical support for this prediction.

Tariff caps and binding overhang have also received attention in other modeling frameworks. In a model with contracting costs, Horn, Maggi and Staiger (2010) compare a weak binding rule and a rigid tariff rule. The framework is different from that considered by Bagwell and Staiger (2005b) but interestingly points to a related set of insights as regards bindings and overhang. In particular, Horn, Maggi and Staiger (2010) show that a weak binding rule is preferred to a rigid tariff rule, since the former permits efficiency-enhancing downward flexibility, and they also note that the weak binding rule is characterized by binding overhang. A different approach is pursued by Maggi and Rodriguez-Clare (2007). As discussed above, they analyze a model in which trade agreements address both commitment and terms-of-trade problems. In their model, binding overhang does not

\footnote{Ossa (2012) considers a similar model, although he focuses on other issues and does not address the optimality of tariff caps.}

\footnote{See Bacchetta and Piermartini (2011) for additional empirical evidence regarding tariff caps and binding overhang.}
occur in equilibrium; however, the potential to apply a tariff below the bound level induces ex post lobbying that has the beneficial effect of diminishing an ex ante problem of excess investment.

The tariff-cap theory described above establishes a rationale for tariff caps when governments have private information and contingent transfers are infeasible. As we argue next, in addition to providing an interpretation of tariff caps and binding overhang, this theory provides a foundation from which to understand contingent protection. A key idea is that contingent protection potentially may enhance efficiency by “linking” tariff choices through time for a given government or across governments, and creating thereby some scope for imperfect contingent transfers. More generally, contingent protection provides a form of upward flexibility that in some cases may enhance efficiency when certain shocks occur. At the same time, it also must be emphasized that tariff caps are a valuable means of “stabilizing” tariff commitments and diminishing the potential for “unwinding” tariff commitments.\textsuperscript{55} The optimal rules for contingent protection thus reflect a delicate balance between maintaining reduced tariffs in response to the terms-of-trade externality and providing some upward flexibility in light of various shocks that may confront governments.

**Contingent protection such as safeguards and antidumping** Bagwell and Staiger (2005b) explore a role for contingent protection when governments experience preference shocks that are privately observed. Formally, they consider a repeated-game model in which government’s privately observed types are iid through time, and in this context they show that expected government welfare may be improved when the tariff choices of any given government are appropriately linked through time.\textsuperscript{56} Their analysis is motivated in terms of the “escape clause” defined in the WTO Agreement on Safeguards. This agreement describes contingencies under which a country can set a tariff above the tariff cap and thus achieve some degree of upward flexibility.\textsuperscript{57} As Bagwell and Staiger (2005b) note, an interesting feature of the WTO safeguards agreement is that it embodies a *dynamic-use constraint*: if a government imposes escape-clause protection in an industry for $X$ years, then it is not allowed to reimpose escape-clause protection in that industry for the next $X$ years. This constraint introduces an opportunity cost to a government from selecting an escape-clause tariff in the current period, so that incentive compatibility is achieved when only a government with a sufficiently high type in the current period imposes an escape-clause tariff. The dynamic-use constraint may thus promote further efficiency by facilitating more efficient tariffs when a government draws a sufficiently high type. The prospect of improved efficiency emerges because a government that goes above the cap today effectively makes a transfer to the other government in the form of an improved continuation value.

Martin and Vergote (2008) develop a related set of insights for a situation in which the tariff choices of one government are linked to those of the other government. They consider anti-dumping

\textsuperscript{55}For further discussion of the importance of tariff stabilization, see Curzon (1965, Chapter 4). Focusing on India, Bown and Tovar (2011) provide empirical evidence that countries use antidumping and safeguard exceptions to unwind commitments to lower tariffs in the presence of domestic political-economic pressure.

\textsuperscript{56}Their work builds on a literature in game theory that associates continuation values with transfers. See Athey and Bagwell (2001) and Fudenberg, Levine and Maksin (1994).

\textsuperscript{57}For an overview of the the WTO safeguards agreement, see Wauters (2010).
duties as allowed under certain contingencies in the WTO Anti-dumping Agreement. Motivated by empirical work by Blonigen and Bown (2003), Feinberg and Reynolds (2006) and Prusa and Skeath (2002) that provides evidence of a retaliatory role for anti-dumping duties, Martin and Vergote (2008) argue that such on-equilibrium-path retaliation may facilitate efficiency gains by ensuring that the home government will protect its import-competition industry with an anti-dumping duty only when the current importance of that industry to the home government is sufficiently great. In this case, when the home government imposes an anti-dumping duty, a link is forged to a future retaliatory anti-dumping duty of the foreign country; and upon raising its applied tariff via an anti-dumping duty the home government thereby again makes a transfer to the foreign government in the form of an improved continuation value.

Our discussion of contingent trade-policies above focuses on the idea that the potential for upward flexibility might improve efficiency when governments are privately informed about their preferences, if contingent transfers achieved through continuation values ensure that only governments with high types choose to exercise this potential. The WTO Safeguards Agreement and also the WTO Anti-dumping Agreement, however, explicitly indicate contingencies under which upward flexibility can be exercised. For example, the WTO Agreement on Safeguards permits the application of a safeguard tariff when the domestic industry is seriously injured as a result of increased imports. If such contingencies are interpreted as defining verifiable market conditions, then an alternative modeling approach is suggested under which a government that seeks to impose a contingent trade policy must incur the necessary costs to publicly verify that the relevant contingent state is present. Bond and Beshkar (2012) offer a first example of this modeling approach. They analyze a partial-equilibrium model with asymmetric country sizes that features costly state verification, in the specific sense that at a cost the importing government can publicly verify the welfare weight that determines its type. A novel feature of their model is that both caps and escape clauses are endogenously determined as part of an optimal trade agreement. Interestingly, they find that circumstances under which the possibilities of overhang (downward flexibility) and escape (upward flexibility) co-exist are relatively rare. Intuitively, when higher types use the escape clause, the cap can be set at a lower level; and this implies in turn that the likelihood of overhang is small.

An alternative theory of escape clauses dispenses with the possibility of private information and focuses instead on the self-enforcement constraints that underlie any cooperative trade agreement between governments. Bagwell and Staiger (1990) offer a first paper of this kind. They consider a partial-equilibrium model with two countries, in which publicly observed trade-volume shocks occur in an iid fashion over time. When a country imports a large volume, it has an increased incentive to cheat on a cooperative agreement and select its unilaterally optimal tariff. The cost

As Sykes (2003) discusses, however, the extent to which the WTO Agreement on Safeguards and subsequent legal cases serve to articulate a clear set of contingencies may be questioned.

Interesting future work might consider the possibility that the importing government can undertake costly state verification for parameters that are more directly related to the market variables highlighted in WTO rules for contingent trade policies.
of such behavior is that it may induce governments to abandon cooperation and revert to Nash trade policies in the future. If governments maximize national income and are sufficiently patient, then they can enforce free-trade policies for all possible trade-volume shocks. If governments have moderate patience, however, then they can enforce free-trade policies only in periods with low trade volumes. When trade volumes take higher values, the incentive to cheat is acute at free-trade policies and the gains from defection overwhelm the discounted future cost of initiating a Nash trade war. Governments with moderate patience, however, can still cooperate in the presence of large trade volume shocks, but they do so by setting positive (and below-Nash) tariffs. Intuitively, by cooperating with a positive tariff, governments reduce the incentive to cheat and bring it back in line with the future discounted cost of a reversion to Nash trade policies. The positive tariffs that accompany high trade-volume shocks can be interpreted in terms of an escape clause, and Bagwell and Staiger (1990) show that an escape clause of this kind is a feature of an optimal self-enforcing trade agreement for governments with moderate patience.60

Bown and Crowley (2013a) provide a first empirical investigation of the cross-sectional and intertemporal predictions of the Bagwell-Staiger (1990) model. As discussed, that model predicts that an import tariff increase is more likely when the import volume increases. In addition, conditioning on a positive import surge, the model predicts that the probability of a tariff increase is positively associated with less elastic import demand and export supply functions and also with less variable import demand volumes. Since the escape clause featured in the Bagwell-Staiger (1990) model could correspond to any instrument of special protection, such as a safeguard tariff or an anti-dumping duty, Bown and Crowley (2013a) analyze the model’s predictions using data on US import tariff increases arising under the US antidumping and safeguard laws and find strong support for the predictions of the Bagwell-Staiger (1990) model.

A special feature of the Bagwell-Staiger (1990) model is that trade-volume shocks are iid through time. A trade-volume shock thus affects the incentive to cheat in the current period but has no direct impact on the discounted future value of cooperation. Bagwell and Staiger (2003) extend the model to allow for both iid shocks to the trade-volume level and stochastic but persistent trends in the growth rate for trade volume.61 In particular, at any given point in time, governments observe the current shock to the trade-volume level and observe as well whether the trading relationship is in a fast- or slow-growth phase, where the relationship moves between the two phases according to a Markov process. They find that the most-cooperative trade agreement for governments with moderate patience is countercyclical: all else equal, the most-cooperative tariffs are lower in a fast-growth phase. The key intuition is that governments have more to lose from initiating a trade war in a (persistent) fast-growth phase; as a consequence, they can then withstand that heightened

---

60 Their work builds on game-theoretic methods developed by Rotemberg and Saloner (1986) for the study of collusion.

61 Hochman and Segev (2010) extend the Bagwell-Staiger (1990) model in a different direction by allowing that governments may imperfectly observe the trade-volume shock before applying tariffs. Interestingly, they identify conditions under which governments can enjoy higher expected welfare when trade-volume shocks are imperfectly observed. Another interesting extension is considered by Tabakis (2010), who examines the use of special protection when countries are transitioning into preferential trading agreements.
incentive to cheat that a lower cooperative tariff implies.\textsuperscript{62} Bown and Crowley (2013b) provide an empirical investigation of the macroeconomic determinants of time-varying trade policy. Using quarterly data for the United States, European Union, Australia, Canada and South Korea, they find evidence of a countercyclical trade policy response in the pre-Great Recession period (namely, the first quarter of 1988 through the third quarter of 2008).\textsuperscript{63} Their findings are broadly consistent with the theoretical predictions offered by Bagwell and Staiger (2003) as well as those in Crowley’s (2011) reciprocal dumping model, which predicts an increase in import restrictions in response to macroeconomic weakness abroad.

A different perspective on the use of safeguards is offered by Sykes (1991, 2003). He observes that instruments of special protection are often applied to assist declining domestic industries. Firms in such an industry retain a large share of the benefits of price-increasing protection, since temporary protection is unlikely to encourage entry in a declining industry. Thus, firms in a declining industry may lobby hard for protection and, therefore, figure prominently in the domestic government’s welfare function. Foreign exporters are naturally harmed by the domestic tariff; however, if the foreign export industry is growing, then they may complain relatively little about facing protection, since their profits eventually would be lost to entry anyhow. These firms therefore may not figure prominently in the foreign government’s welfare function. Consequently, governments in an ex ante state may be attracted to a safeguard rule for declining industries, since the welfare benefit to the government that uses the safeguard may exceed the welfare cost to the government whose exporters face the safeguard.

In summary, our selective survey highlights that tariff caps and binding overhang occur in an optimal trade agreement when governments are privately informed about their preferences and contingent transfers are infeasible. With this foundation for tariff caps in place, we then consider three complementary theories for the use of special instruments of protection. First, when governments experience privately observed preference shocks, upward flexibility might promote further efficiency, provided that higher tariffs are only used by governments facing high political pressures. This incentive compatibility constraint, in turn, can be met, when current tariffs are linked to future tariffs so that contingent transfers can be achieved, at least to some degree, through changes in continuation values. The incentive-compatibility constraint also can be met when costly state verification is feasible. Second, optimal cooperation when trade volumes are volatile entails a low baseline tariff coupled with an escape clause that allows for higher tariffs when trade-volume shocks are high. Safeguards emerge in an optimal trade agreement in this setting and indeed complement tariff liberalization, as the baseline tariff would be higher were safeguards not allowed. Finally, safeguards may enable governments to achieve greater welfare by rewarding (penalizing) industries that figure more (less) prominently in governments’ welfare functions. This theory, in particular, associates the use of safeguards with declining industries. While more empirical work remains to

\textsuperscript{62}This work builds on methods used by Bagwell and Staiger (1997c) in their analysis of collusion over the business cycle.

\textsuperscript{63}Bown and Crowley (2014) provide additional empirical support from a sample of 13 emerging economies and use of annual data covering the period 1995-2010.
be done, we have also identified several studies that provide empirical support for themes emerging from the theoretical analyses.

4.3 Subsidies

The appropriate treatment of subsidies in an international trade agreement is subtle. On the one hand, a domestic production subsidy can be a “first-best” instrument with which to address a market failure that results in an inefficiently low level of output. A production subsidy may also be a valuable instrument to a government with political-economic objectives that wishes to redistribute surplus to producers in a given industry. On the other hand, some restrictions on the use of domestic production subsidies are clearly necessary, since otherwise a government could always use a domestic production subsidy to an import-competing industry so as to undermine the benefits offered to other countries through negotiated tariff cuts. The appropriate treatment of export subsidies is also not obvious. An export subsidy lowers the world price for the export good and thus generates a terms-of-trade improvement for importing countries, but it may also displace exports from other countries and alter entry and exit patterns across countries. In this subsection, we briefly review the GATT/WTO legal treatment of subsidies and then discuss economic research that interprets and evaluates this treatment.\footnote{For closely related discussions, see also Bagwell (2008) and Bagwell, Staiger and Sykes (2013).}

**Subsidy rules under the GATT versus WTO** The treatment of subsidies in GATT was relatively tolerant in nature. Under GATT rules, a foreign trading partner could respond to the subsidies of the domestic country through two possible means.\footnote{Sykes (2005) offers a detailed discussion of the evolution of subsidy rules under GATT and the WTO.} First, the foreign government could unilaterally impose a countervailing duty (CVD) if its import-competing industry experienced material injury as a consequence of an export subsidy given to producers in the domestic country. Second, the foreign government could file a non-violation complaint, if it had previously negotiated a tariff reduction from the domestic government on a given product and the domestic government later offered a subsidy to its import-competing producers of that product. To succeed with such a complaint, the foreign government would have to show that a new or increased domestic subsidy program emerged that had the effect of nullifying or impairing the market-access benefits that the foreign government had reasonably expected at the time of the tariff negotiation.\footnote{See Petersmann (1997, pp. 151-4) for discussion of the conditions under which a subsidy could be determined to upset market-access expectations.} In this case, the domestic government would be under no obligation to remove the subsidy; however, it would be expected to make policy adjustments that restored the foreign country’s negotiated market access. In addition, as part of GATT negotiations, several countries agreed to restrict the use of export subsidies, especially for non-agricultural goods.

The treatment of subsidies in the WTO’s *Agreement on Subsidies and Countervailing Measures* (the SCM Agreement) is much more restrictive. First, except as otherwise allowed for in the *Agreement on Agriculture*, export subsidies (and also local-content subsidies) are prohibited. Second,
“specific” subsidies that have “adverse effects” on other members are actionable, where adverse effects could take the form of an injury to an industry in another member country, the nullification or impairment of benefits expected by another WTO Member, and “serious prejudice” to the interests of another WTO Member. The first two forms of adverse effects are broadly reflected in the GATT treatment of subsidies and may be associated with the use of CVDs and non-violation complaints by the adversely impacted member country. The more novel ingredient is serious prejudice, which may occur if the effect of a subsidy offered by the domestic country is to cause a loss of exports by a WTO member into the domestic market or into a third-country market. Importantly, a domestic production subsidy can be actionable under the SCM Agreement independently of whether the subsidy nullifies or impairs the market-access expectations associated with some prior tariff negotiation.\(^{67}\)

WTO rules thus treat subsidies in a fairly severe manner. In contrast to import tariffs, for which caps are negotiated, exports subsidies are banned outright. As well, under WTO rules, a country that uses a domestic production subsidy must withdraw it, or at least remove its adverse effects, even if the subsidy itself does not erode any negotiated market-access expectation. We consider next research that interprets and evaluates the WTO’s treatment of subsidies.

**Production subsidies** We begin with the treatment of domestic production subsidies. On the one hand, and as noted above, a domestic production subsidy is a first-best instrument with which to address a market failure that leads to an inefficiently low level of production.\(^{68}\) A domestic production subsidy also may be an attractive instrument for a government with political-economic preferences that seeks to redistribute surplus to the import-competing industry. On the other hand, if the use of domestic production subsidies were completely unregulated, then governments would be unable to achieve efficient outcomes through reciprocal tariff negotiations alone. Intuitively, in the absence of any restrictions on the use of such subsidies, a government that exchanged reciprocal tariff cuts with a trading partner could subsequently “undo” the market-access consequences of its own tariff cut by providing a production subsidy to its domestic import-competing industry. This discussion suggests that domestic production subsidies have a potential efficiency-enhancing role to play but that their use must be regulated in some manner. In light of these considerations, we may ask: how should domestic production subsidies be treated in a trade agreement?

The trade-offs just described suggest a potential answer to this question: grant each government flexibility when choosing its domestic policies, provided that the overall effect of its chosen domestic policy does not erode the market-access commitments achieved through its preceding tar-

---

\(^{67}\)As regards countermeasures, if the domestic government refuses to remove an export subsidy, then the complaining Member government may take “appropriate countermeasures.” For an actionable subsidy, in the absence of an agreement on compensation, and if steps to remove the adverse effects are not undertaken or if the subsidy itself is not withdrawn, then the complaining Member may be granted authority to impose countermeasures that are “commensurate” with the adverse effects attributable to the subsidy. See Lawrence (2003, pp. 54-60), Mavroidis (2000), Spamann (2006), and Bown and Ruta (2010), for further discussion.

\(^{68}\)An import tariff is another instrument that might be used to expand domestic production levels; however, an import tariff is a second-best instrument. An import tariff is equivalent to a consumption tax and a production subsidy, and thus affects both consumer and producer margins.
iff negotiations. Under this approach, following the completion of a tariff negotiation, a government would be allowed to alter its domestic policies in any way, provided that the overall effect of any changes does not result in a terms-of-trade loss for its trading partner. To formally examine this approach, Bagwell and Staiger (2006) augment the two-country, general-equilibrium model of trade considered above to include domestic policies as well. Under the assumption that governments have available a set of domestic instruments that is sufficiently rich to create a degree of policy redundancy, they find that GATT rules enable governments to achieve an efficient outcome using tariff negotiations. A key feature of GATT rules in this context is that a government can file a non-violation complaint if it suffers a terms-of-trade loss as a consequence of a change in the domestic policies of its trading partner. By contrast, the WTO’s SCM Agreement places further restrictions on the use of domestic production subsidies. These additional restrictions limit policy redundancy and may thereby prevent governments from achieving efficient outcomes through tariff negotiations.

Bagwell and Staiger (2006) also consider a setting with limited instruments. In this setting, GATT rules are no longer sufficient for achieving efficiency through tariff negotiations. Intuitively, in a limited-instrument setting, a government may be unable to reposition its subsidy to an efficient level without imposing a terms-of-trade loss on its trading partner and thus encouraging a non-violation complaint or the application of a CVD. Indeed, when the set of instruments is limited, the SCM Agreement could represent an improvement over GATT rules. For example, if governments respectively maximize national income and no market failure exists that creates a corrective role for domestic production subsidies, then subsidies would not be used in an efficient outcome. More broadly, though, market failures and/or redistributive goals suggest a potential role for domestic production subsidies in an efficient outcome. As Bagwell and Staiger (2006) argue, tight restrictions on subsidies could then have a “chilling effect” on tariff negotiations. To the extent that the SCM Agreement imposes tight restrictions on the use of domestic production subsidies, GATT rules on subsidies are then preferred to those in the WTO.

GATT non-violation rules on domestic policies identify an attractive approach in granting flexibility to governments up to the point where an externality is opposed on trading partners. At a practical level, however, it also must be acknowledged that non-violation rules have their own limitations. As two illustrations, we mention that it may not be obvious what a government should reasonably expect at the time of negotiation, and it also may not be clear where to draw the line in terms of which sorts of domestic policy changes are appropriately disciplined using non-violation nullification-and-impairment complaints. While these points certainly raise valid considerations, we nevertheless find much to be admired in the GATT approach.

---

69 Sufficient policy redundancy is satisfied if each government has available an import tax, a domestic production subsidy and a domestic consumption tax.

70 The idea is that governments may be hesitant to negotiate tariff bindings if subsidies are disciplined heavily, since tariffs may then be the best remaining means of providing assistance to domestic import-competing industries.

71 Another important caveat is that a market-access-preservation rule may fail to be optimal in settings with private information, as Lee (2014) argues. We discuss his paper in greater detail in Section 5.
Export subsidies in two-country models  We turn now to the treatment of export subsidies. We begin with a simple observation: in the simple two-country, two-good general equilibrium model described above, an increase in the domestic country’s export subsidy would have the same effect on prices and thus government welfare functions as would a decrease in its import tariff. This observation, known as the Lerner symmetry theorem, follows since either policy change would lower (raise) the relative price of the domestic import good in the domestic (world) market. One implication is that both policy changes would generate a terms-of-trade loss (gain) for the domestic (foreign) country. Since each government’s welfare is expressed as a function of the relative price in its country and its country’s terms of trade, it follows that the two policy changes affect government welfare functions in the same fashion as well. Hence, our arguments above carry over immediately when governments select export policies instead of import policies. In particular, Nash export policies are inefficient and result in too little trade, and governments can mutually gain from an agreement on export policies only if they agree to make reciprocal adjustments that lead to greater trade volume.

This result is familiar when governments respectively maximize national income. In that case, free trade is efficient, but each government’s optimal unilateral policy is an export tariff. The key intuition is that a government can use an export tariff to induce its competitive export industry to restrict output as would a monopolist, where monopoly rents are now retained in the form of tariff revenue. In the resulting Nash equilibrium, both governments impose export tariffs, and the trade volume is inefficiently low. More generally, the sign of a government’s unilaterally optimal export policy depends on the government’s specific preferences. In particular, the Nash export policy is an export subsidy for a government that gives sufficient welfare weight to the interests of its export sector. The fundamental point, though, is the governments’ noncooperative export policies, whatever their sign, induce too little trade from the governments’ joint perspective.

A trade agreement can thus generate mutual gains for governments only if it facilitates reciprocal increases in export subsidies (or reciprocal decreases in export tariffs) relative to noncooperative levels. Intuitively, an increase in a country’s export subsidy generates a positive terms-of-trade externality for its trading partner, whose consumers now enjoy a lower price on their import good. While this argument makes sense within the context of the terms-of-trade approach to trade agreements, it runs completely counter to the treatment of export subsidies in the WTO. There are thus two possibilities: either the WTO’s prohibition on export subsidies is misguided, or the simple two-country, competitive-markets version of terms-of-trade theory is missing something important. To explore the latter possibility, we discuss next a sequence of enriched terms-of-trade models within which to further explore the treatment of export subsidies.

Export subsidies in third-market models  One potential cost of export subsidies not featured in the above discussion is that an export subsidy offered in one country lowers the world price and thereby imposes a negative terms-of-trade externality on other exporting countries. To explore this...
issue, we follow the “strategic-trade” literature and construct a “third-market” model. Specifically, we illustrate our points using a simple partial-equilibrium model with a single good, where all firms are located in Country’s 1 and 2 and all consumers are located in Country 3. Firms in Country’s 1 and 2 then compete for sales to consumers in Country 3. In this context, we ask two questions. First, when would a government have a unilateral incentive to offer an export subsidy? And second, when if at all should an international trade agreement discipline the use of export subsidies?

Following the pioneering model of Brander and Spencer (1985), we begin with a setting in which a finite and given number of firms engage in Cournot competition for sales in Country 3. The key issues involved can be illustrated in a simple third-market model with two firms, wherein Firm 1 is located in Country 1, Firm 2 is located in Country 2 and all consumers are located in Country 3. In the absence of any subsidy, Firms 1 and 2 have the same constant marginal cost. The governments of Countries 1 and 2 respectively maximize national income. The game has two stages. The governments simultaneously select their respective specific (i.e., per unit) export subsidies, and after observing these selections Firms 1 and 2 simultaneously select their respective output levels.

A key finding is that, starting at free trade, a government that maximizes national income now has a unilateral incentive to offer an export subsidy. As is now well known, an export subsidy lowers the cost of the exporting firm and thus shifts out this firm’s output reaction curve. The strategic advantage of such a shift is that the other exporting firm responds by reducing its own output. In this general fashion, an export subsidy “shifts profits” to the subsidizing country. After all output adjustments are made, the overall effect of a strategic export subsidy is to expand aggregate output and thus induce a fall in the world price. Hence, as in the competitive model, an export subsidy generates a positive terms-of-trade externality for consumers. A new feature of the third-market model is that an export subsidy generates as well a negative terms-of-trade externality for the firm in the other exporting country.

Of course, the other exporting country has a similar incentive to engage in strategic export subsidization, and the resulting subgame perfect Nash equilibrium of the two-stage game involves export subsidies from both exporting countries. In the end, the two exporting countries are worse off when export subsidies are legal, since their strategic efforts offset and simply result in a lower world price. Consumer and global economic welfare, however, are higher when subsidies are allowed than would be the case were subsidies banned. The described model therefore provides an interpretation for why exporting countries would seek a ban on export subsidies as a means to keep the price high, but it suggests that an international trade agreement designed to maximize the combined welfare

---

73 As we discuss in Section 6.2, some recent research evaluates whether the strategic-trade and “delocation” theories of export subsidies can be interpreted as providing a rationale for trade agreements that is distinct from that provided by the terms-of-trade externality. Bagwell and Staiger (2009, 2012a) argue, however, that the problem for a trade agreement to solve in the profit-shifting and delocation settings can be given a terms-of-trade interpretation, provided that governments have available both import and export policies. In particular, they show that politically optimal policies remain efficient in these settings when a full set of trade policies is available. We thus include some discussion of these theories here, as part of our discussion of the implications of the terms-of-trade approach for the treatment of export subsidies, while postponing further discussion of the rationale of trade agreements in these settings until Section 6.2.
of all countries in fact should encourage even greater use of export subsidies than occurs in the noncooperative equilibrium.\textsuperscript{74}

The simple strategic subsidy model just described has been extended in many directions to allow for alternative forms of imperfect competition, multiple exporting firms in exporting countries, consumers in exporting countries, endogenous cost technologies and other considerations.\textsuperscript{75} Such extensions provide important qualifications about the sign of the unilaterally optimal export policy, but a robust feature of models with a fixed number of firms is that a more expansionary export policy provides a positive terms-of-trade externality to importing consumers. In fact, as Bagwell and Staiger (2001c) argue, a related perspective on the treatment of export subsidies emerges even in third-market models with competitive industries that operate subject to rising marginal costs. Using such a model, they show that the subgame perfect equilibrium features export subsidies when government welfare functions in exporting countries place sufficient weight on export interests. Moreover, whatever the specific preferences of governments in exporting countries, an efficient export policy that maximizes the combined welfare of all three governments recognizes the positive terms-of-trade externality to consumers and calls for more expansionary export policies than would occur in the noncooperative equilibrium.

Returning now to the two questions identified above, we see first that the third-market models reviewed above provide an interpretation for the unilateral appeal of export subsidies, at least in some settings. The third-market models in which export subsidies would be used, however, suggest that export subsidies may be under-supplied. More generally, the models reviewed above suggest that noncooperative export policies are less expansionary than would be efficient, when efficiency is assessed using government welfare functions. The third-market models thus do not provide a foundation from which to understand a cap on export subsidies, much less the WTO’s prohibition on export subsidies. The models also do not provide an interpretation for the fact that the WTO treats export subsidies more severely than import tariffs.

What are some other considerations not captured in the models above that potentially might explain stronger disciplines on export subsidies?\textsuperscript{76} First, an export subsidy might hurt government welfare in the importing country, despite the consequent terms-of-trade improvement, if the importing country has an import-competing sector that is harmed. This consideration is certainly valid but also has important limitations. Notably, in competitive markets, the government of the importing country must gain if its tariff is near the unilaterally optimal (i.e., best-response) level or if it can stabilize the local price by applying a CVD.\textsuperscript{77} Second, the models developed above do not capture all of the welfare costs that are associated with export subsidies: such subsidies may

\textsuperscript{74} The market is initially distorted with too little production as a consequence of oligopolistic competition. The equilibrium with strategic export subsidies expands output closer to the competitive level, and a trade agreement could further improve global welfare by increasing subsidy levels so that the competitive output is produced.

\textsuperscript{75} Eaton and Grossman (1986) and Maggi (1996) provide the key studies examining the dependence of the sign of optimal export policies on the form of imperfect competition in the product market. Bagwell and Staiger (1994) argue that the sign of the optimal strategic R&D policy is less sensitive to the form of imperfect competition. See Brander (1995) for a review of the strategic-trade literature.

\textsuperscript{76} For further discussion of these considerations, see Bagwell, Staiger and Sykes (2013, pp. 186-9).

\textsuperscript{77} The first point is established in Bagwell and Staiger (2002, p. 192).
create or exacerbate distortions in production, and may also encourage rent-seeking behavior. We do not intend to minimize the importance of such considerations; however, we also point out that similar welfare costs are associated with the use of import tariffs, and we further note that many of these costs are domestic in nature and are not obviously best addressed in an international trade agreement. Third, an export subsidy offered by a country with competing exporters could undermine the market-access benefits that another exporting country obtained in a previous negotiation. This consideration suggests that export subsidies could impede the success of reciprocal tariff negotiations more generally. We agree that this is an important - and under-studied - consideration, and we simply remark here that it is not obvious that a prohibition on export subsidies represents the optimal legal response. Finally, it might be argued that, on fairness grounds, export subsidies should be disciplined more severely than import tariffs, since some countries face higher costs of raising revenue and funding export-subsidy programs. We acknowledge this point, too. At the same time, we note that these same countries may also benefit when export subsidies are applied to the goods that they import, and so it is again not obvious that a prohibition on export subsidies is the optimal legal response.

**Export subsidies and industrial policy**  The models described so far do not focus on the long-run implications of export policies for industrial structure. To explore this issue, we now follow Venables (1985) and consider the “delocation” effects of trade policies in a two-country model with two-way trade in a homogeneous good. The game has three stages. In the first stage, governments simultaneously select (specific) import and export tariffs. Each government seeks to maximize national income in its country. In the second stage, after observing the trade-policy selections, potential entrants decide whether to locate in the domestic or foreign market, where entry entails a positive fixed cost. In any country, entry occurs in this stage until expected profits (including the fixed cost) are driven to zero. Finally, in the third stage, after observing trade policies and the numbers of firms located in each of the two countries, the entering firms simultaneously choose output levels in Cournot fashion, where an individual firm selects both an output level for the market in which it is located as well as a separate output level for exportation into the other market. The two markets are segmented, and a positive (per-unit) transport cost must be incurred for exported units.

A key feature of this model is that it exhibits a Metzler paradox: if a government raises its import tariff (or raises its export subsidy), the price of the good within its country falls. To see the intuition, suppose that we start at global free trade with levels of entry in each country that generate zero profit for each firm, and suppose that the domestic government then imposes a slight import tariff (or a slight export subsidy). Holding fixed the numbers of firms in each country, domestic firms would then enjoy positive profits while foreign firms would experience negative profits. Some adjustment in the patterns of entry across the two countries is thus necessary to restore zero profits. Due to positive transport costs, each firm sells greater output in its local market than in its export market; hence, the only way to reduce the profit of a domestic firm while increasing the profit of
a foreign firm is to adjust entry patterns until the domestic price falls and the foreign price rises. Consequently, the domestic policy change must induce (reduce) entry into the domestic (foreign) market to such an extent that the domestic price falls (foreign price rises). In this sense, when a government imposes a higher import tariff (or export subsidy), it “delocates” firms from the other country to its own country.

For this model, as Venables (1985) shows, if all policies are initially set at free trade, then the domestic government can gain by imposing a small import tariff. Producer surplus is unaffected by the change, since free-entry conditions ensure that firms earn zero profit. But a small import tariff generates positive tariff revenue and also leads to a lower price and higher consumer surplus in the domestic country, due to the Metzler paradox. A small export subsidy likewise leads to a lower price and higher consumer surplus in the domestic country; however, the small export subsidy imposes a cost in the form of subsidy expenses. Venables (1985) shows that, when demand and costs take linear forms, a small export subsidy also generates a net gain for the domestic government. Importantly, both policy changes result in a higher foreign price, lower foreign consumer surplus and lower foreign government welfare. Starting at free trade, therefore, Venables’ (1985) analysis shows that export subsidies are unilaterally attractive and impose a negative terms-of-trade externality on the trading partner.

Bagwell and Staiger (2012b) generalize the analysis of the linear Cournot delocation model in several respects. They show that the Nash policies for governments in fact are characterized by the use of import tariffs and export tariffs. Thus, while an export subsidy is unilaterally attractive for a government when its import policy is free trade, the government prefers an export tariff when its import tariff is optimally set at a positive level. The new consideration that underlies this finding is that, when a positive import tariff is in place, an export tariff generates additional tariff revenue on imports by encouraging foreign entry and thus exports. They also find that free trade in import and export policies is efficient in the linear Cournot delocation model.

Together, these findings suggest a possible interpretation of the treatment of export subsidies in GATT/WTO. The linear Cournot delocation model suggests that governments would perceive a unilateral gain from using export subsidies only once import tariffs were negotiated to levels sufficiently close to free trade. From this perspective, it is not surprising that GATT rules did not impose strong restrictions on the use of export subsidies. Over time, however, as import tariffs were negotiated through GATT rounds to lower levels, governments may have perceived a unilateral gain from imposing export subsidies. Furthermore, since free trade is an efficient outcome in the linear Cournot delocation model, governments could achieve mutual gains given low import tariffs if they were to cap or even prohibit export subsidies. In this way, the model offers a potential efficiency-enhancing interpretation of the prohibition of export subsidies in the WTO SCM Agreement.

Among the models reviewed above, the linear Cournot delocation model offers the most successful interpretation of the treatment of export subsidies in the WTO. At the same time, we note that the model is not completely successful. In particular, the linear Cournot delocation model also predicts that governments would benefit from a prohibition of import tariffs, and so it does not deliver
Implications for export subsidy rules under the WTO

In sum, the appropriate treatment of subsidies in a trade agreement is a subtle issue. Our review of the literature in this subsection focuses on models for which trade policies generate terms-of-trade externalities for trade partners. The review reinforces the subtle implications of subsidies: domestic productive subsidies can play both efficiency-enhancing and opportunistic roles, export subsidies generate positive externalities to foreign consumers and negative externalities to foreign firms in models with fixed industrial structures, and export subsidies may generate negative externalities to foreign consumers in long-run settings with endogenous entry and exit. On the whole, our review does not provide strong support for the specific treatment of subsidies in the WTO. We describe work suggesting that WTO rules on domestic production subsidies may be a step backwards relative to GATT rules, and we also summarize a range of models under which export subsidies are actually under-supplied relative to the efficient level for governments. While the linear Cournot delocation model provides a potential interpretation for an agreement to limit or even prohibit the use of export subsidies, the existing formal models that we review do not identify a reason for treating export subsidies more severely than import tariffs. We also identify several directions for future research that might provide further insights regarding the interpretation and evaluation of WTO subsidy rules.

4.4 Non-Violation Complaints, Shallow Integration and National Treatment

The central implication of the terms-of-trade theory of trade agreements is that governments set unilateral tariffs at levels that are inefficiently high, since they fail to internalize the terms-of-trade implications of their tariff policies for each other. A trade agreement can “undo” this inefficiency by facilitating mutually advantageous and reciprocal tariff reductions that expand the volume of trade to more efficient levels. Tariffs are not the only instruments, however, that impact the terms of trade. For large countries, domestic taxes, subsidies and standards may also affect the terms of trade and lead thereby to possible inefficiencies. At the same time, domestic policies may well have legitimate and even first-best roles to play as instruments with which to address various market failures or distributional concerns within a country. Attempts to regulate domestic policies through an international trade agreement therefore must balance the possible opportunistic use of such policies against the efficiency-enhancing roles that such policies may play. We thus arrive at the following question: how should domestic policies be treated in a trade agreement? In this
subsection, we describe research that responds to this question while utilizing the terms-of-trade approach to trade agreements.

In fact, we have already encountered this question above in the specific context of our discussion of the treatment of domestic production subsidies. As we note there, a natural answer to this question is that a trade agreement should grant each government flexibility when choosing its policies, provided that the overall effect of any policy changes does not erode the market-access commitments achieved through its preceding tariff negotiations. Bagwell and Staiger (2001b) formally explore this answer in the context of a two-good general equilibrium model of trade in which governments have available domestic policies (e.g., labor or environmental standards) as well as tariff policies. Working with a model in which all international externalities flow through the terms of trade, their main finding is that efficiency can be achieved through tariff negotiations alone, provided that each government is free to make subsequent adjustments in its domestic and trade policies that leave its negotiated market access commitment (i.e., the terms of trade) unaltered. As discussed in Section 4.3, Bagwell and Staiger (2006) obtain a related finding in the specific context of domestic production subsidies.

As noted in Section 3.1, a key assumption of the Bagwell-Staiger (2001b) model is that each government only has a direct interest in the domestic policy adopted by its country, where this interest in turn may reflect various national considerations that impact the government’s welfare (e.g., a government may have a direct interest in the health and safety of its citizens, or in the environmental quality within its country’s borders). The lack of any direct interest by any one government in the domestic policy selected by another government indicates that the model does not allow for non-pecuniary international externalities (e.g., global pollution). For this family of models, the domestic policy choices of one government therefore impact the welfare of another government only indirectly, through the terms of trade. Notice, though, that “race to the bottom” concerns are about the pecuniary international externalities (trade effects) associated with a choice of weak standards, and so this family of models is quite capable of capturing those concerns.

When international externalities travel only through the terms of trade, the main finding of Bagwell and Staiger (2001b) suggests that a “shallow integration” approach to trade agreements may suffice. Governments need not negotiate directly over domestic policies; instead, they may negotiate over tariffs alone, provided that the market-access gains so achieved are secure. The important task for an international agreement in this context is then to ensure that negotiated market-access concessions are secure against opportunistic policy adjustments. Bagwell and Staiger (2001b) and Bagwell, Mavroidis and Staiger (2002) argue that current GATT/WTO rules, which focus on market access, can with some strengthening strike the right balance, so that governments can set efficient domestic policies while pursuing international negotiations over tariffs alone.

**Nonviolation complaints and the preservation of market access** In particular, GATT rules that permit non-violation complaints are an attractive means of securing market-access concessions. A government may file a non-violation complaint when a trading partner undertakes a policy change...
that nullifies or impairs the market access gains that a government reasonably expected as part of an earlier negotiation. For example, following a tariff negotiation, the possibility of facing a non-violation complaint might deter a government from an opportunistic (terms-of-trade improving) relaxation in the production standards that it requires for an import-competing industry. Existing rules are insufficient, however, to enable a government to adjust its policy mix following a tariff negotiation by raising its standards in the import-competing industry while also raising its import tariff so as to maintain its negotiated market-access commitment. Bagwell, Mavroidis and Staiger (2002) propose that a modification to GATT rules of renegotiation, whereby a government could use a higher standard in an import-competing industry as compensation for a higher import tariff, could in principle provide the needed flexibility.

Ederington (2001) explores related themes in a model in which all international externalities travel through the terms of trade and any agreement on tariffs and domestic policies must be self-enforcing. In his model, each government has two instruments – an import tariff and a domestic production tax – and the latter instrument has a legitimate role since domestic production generates a non-pecuniary externality that resides entirely within the country in which production occurs. Both policies affect the terms of trade, and the challenge is to ensure that the efficient policy mix is selected. In Ederington’s (2001) model, the efficient domestic policy is a Pigouvian tax that offsets the domestic distortion, and the efficient level of the domestic tax in fact is independent of the import tariff and thus the level of market access. Consistent with the themes developed above, Ederington (2001) shows that the most-cooperative solution is achieved when domestic policies are set at the efficient (Pigouvian) level and import tariffs are lowered so as to expand market access to the level that is as close to efficient as possible before the self-enforcement constraint of the repeated game binds. The key intuition is that an efficient domestic policy raises the discounted future value of cooperation, which in turn enables governments to enforce lower tariff levels.78

The principle of national treatment Our discussion of the terms-of-trade implications of domestic policies to this point emphasizes the benefit that the domestic country may enjoy when domestic production standards in an import-competing industry are relaxed in an opportunistic manner. As Staiger and Sykes (2011) argue, however, the bulk of WTO disputes concern instead cases in which foreign suppliers complain about standards that apply to their own products. Following a similar line of reasoning, we may anticipate potential incentive for domestic government to set standards on foreign products in an opportunistic fashion that could undermine the security of negotiated market-access gains for the foreign exporters. A key design feature of GATT/WTO rules that guards against such opportunism is the “national treatment” principle. This principle, which is embodied in GATT Article III, The WTO Agreement on Technical Barriers to Trade (TBT), and the WTO Agreement on Sanitary and Phytosanitary Measures (SPS), restricts the ability of

78Lee (2007) considers a related model but with the important difference that each government is privately informed about the magnitude of the domestic production externality in its country. As Lee (2007) shows, in this case it may be optimal to distort the tariff in order to limit the potential for disguised protectionism, which in his model occurs when a government misrepresents its information by selecting a low production tax even though the externality cost is high.
member governments to impose regulations on foreign suppliers that are more stringent than those imposed on domestic suppliers.

To formally explore the effectiveness of the national treatment principle, Staiger and Sykes (2011) adapt and extend the general insights of Bagwell and Staiger (2001b) to a product-standards setting. In the Staiger-Sykes (2011) model, the domestic government chooses trade policy as well as domestic tax and regulatory policy with respect to a product that is domestically produced and also imported. Regulatory policy has a legitimate role to play in their model, since domestic consumption generates a negative and non-pecuniary consumption externality that resides entirely within the domestic country. Staiger and Sykes (2011) show that governments of large countries indeed have incentive to impose discriminatory tax and regulatory policies against foreign imported products once import tariffs are bound. The model thus provides an interpretation of the national treatment principle as a guard against such opportunistic behavior. When product-specific consumption taxes are infeasible, they further show that the domestic government has an incentive to impose inefficiently stringent non-discriminatory product standards, since foreign exporters bear some of the cost of achieving higher product standards that benefit domestic consumers. In light of their findings, they conclude that WTO legal framework does a reasonably good job of policing regulatory discrimination but does relatively little to address nondiscriminatory regulations that are excessively stringent. At the same time, it is not obvious how to craft alternative and superior rules directed toward limiting the use of excessively stringent non-discriminatory regulations.

The national treatment principle has been interpreted and evaluated in other studies as well. In particular, Horn (2006) and Horn, Maggi and Staiger (2010) examine the national treatment principle with a focus on domestic taxes rather than regulatory standards. Grossman, Horn and Mavroidis (2013) also provide an extensive study of the GATT national treatment provision and argue that case law, economic theory and the negotiation record all suggest that the purpose of the national treatment provision is to outlaw protectionist use of domestic policies.\textsuperscript{79}

\textbf{Moving beyond shallow integration?} The case for shallow integration described above rests on the assumption that all international externalities are pecuniary and, in particular, travel through the terms of trade. As Antràs and Staiger (2012a, 2012b) argue, more complex forms of integration may be required in the presence of offshoring, which alters the means through which prices are determined and complicates the nature of international pecuniary externalities. We discuss their work further in Section 6.3. Similarly, “deeper” forms of integration may be needed if the trade agreement is created with the goal of also addressing non-pecuniary international externalities (e.g., global pollution).\textsuperscript{80} Limão’s (2005) model, which we discuss next, offers one illustration of this point.

Limão (2005) explores a model of self-enforcing cooperation among governments, with the new

\textsuperscript{79}Other related studies include Battiglini and Maggi (2003), Copeland (1990) and Costinot (2008).

\textsuperscript{80}For discussion of the extent to which GATT/WTO rules can be used to address non-pecuniary international externalities, see, e.g., Bagwell, Mavroidis and Staiger (2002) and Trebilcock, Howse and Eliason (2013, Chapters 17 and 18) and the references cited therein.
feature that production in the import-competing sector generates a negative non-pecuniary externality that travels (at least to some degree) across national borders. Each government selects an import tariff and a domestic production tax, and both policies affect the terms of trade. In Limão’s (2005) model, therefore, international externalities travel through the terms of trade as well as through a non-pecuniary channel. His findings illustrate that a form of “deep integration” is attractive to governments in such a setting as a means of relaxing enforcement constraints, where deep integration in this context refers to a policy linkage whereby the potential for retaliation in both policies deters deviations in any one policy alone. In particular, Limão (2005) finds that governments can achieve higher welfare in a self-enforcing agreement when the policies are linked; however, there is no guarantee that linkage raises the level of cooperation in each policy.

In total, our survey of research in this subsection provides support for the “shallow integration” approach of the GATT/WTO when externalities are pecuniary and travel through the terms of trade. GATT/WTO rules concerning nonviolation complaints and national treatment can be interpreted from this perspective and rest on a solid economic foundation. At the same time, we also note that nondiscriminatory product regulations may be inefficiently stringent under such rules. Anticipating discussion in later sections, we also note that arguments for deeper integration may emerge in settings where international externalities travel through other channels as well.

4.5 Investment and Services

The creation of the WTO in 1995 includes new agreements on investment and services. In this subsection, we consider these new agreements and discuss economic research utilizing the terms-of-trade theory that interprets and evaluates the provisions contained therein.

Foreign direct investment, local content, and international cross-ownership We start with the treatment of investment in the GATT/WTO, with an initial focus on foreign direct investment (FDI). The past two decades have witnessed a significant growth in FDI activity. This growth in turn encourages consideration of the investment measures that host governments may impose and the appropriate treatment of those measures in the GATT/WTO. Investment measures interact with GATT rules when they have direct effects on trade; in particular, local content, export and trade-balancing requirements may distort investment decisions and generate tension with basic GATT rules concerning national treatment and quantitative restrictions. As

81 Retaliation in Limão’s (2005) repeated-game model occurs off the equilibrium path and thus carries the interpretation of a breakdown in cooperation with respect to the relevant policies.

82 See also Spagnola (1999a,b), who develops related themes for a distinct class of interdependent payoffs.

83 FDI may be attractive to firms for a variety of reasons. For example, FDI may provide access to cheap inputs, reduce trade costs and facilitate “tariff jumping.” FDI also may be advantageous relative to outsourcing as a means of maintaining tighter control over technology. At the same time, firms considering FDI confront a variety of possible costs, including the possibility of rent expropriation via government policy changes after sunk costs are incurred. For a survey of research on multinational firms, see Antrás and Yeaple (2014).

84 Local content requirements concern measures that require foreign-owned firms to discriminate between domestic and imported goods that are used as inputs for production in the host country; export requirements concern measures that require exportation of a certain percentage of the foreign-owned firm’s output; and trade-balancing requirements
Trebilcock, Howse and Eliason (2013, Chapter 15) discuss in further detail, the WTO Trade-Related Investment Measures Agreement (TRIMs) builds on GATT rules to subject some measures with direct effects on trade, such as local content requirements and quantitative restrictions, to explicit scrutiny under GATT norms. The appeal of additional restrictions on investment measures, however, is controversial, with some countries expressing concerns about the broader implications of extensive investor-protection provisions in trade agreements. Investment has now been removed as a topic for further discussion in the WTO Doha Round. A number of investment agreements have arisen, however, via Bilateral Investment Treaties (BITs) and as part of investor-protection provisions in PTAs.

The purpose of restrictions on investment measures in a trade agreement can be developed at several levels. We mention two here. First, in the absence of a trade agreement that imposes restrictions on investment measures, a government might be tempted to impose measures favoring local input suppliers as a means of generating advantageous price changes for its country. Second, a trade agreement that appropriately restricts investment measures may also encourage efficiency-enhancing FDI when the host government is otherwise unable to commit not to expropriate foreign rents once the foreign firm has sunk costs. We briefly discuss the first purpose next and note that the second purpose is more directly associated with the “commitment theory” of trade agreements developed in Section 6.

Bagwell and Sykes (2005) examine conditions under which a local content requirement generates advantageous price changes for the domestic (i.e. host) country. As they emphasize, a local content requirement is logically distinct from import tariffs and quotas, since a local content requirement does not generate government revenue. They consider a simple two-country partial-equilibrium model in which a single homogenous input supplied in both the domestic and foreign country is used by foreign-owned firms to manufacture a final good for sale in the domestic market. When markets are competitive and the domestic country is small in that its local content requirement does not affect the world (i.e., foreign) input price, a domestic local content requirement raises the domestic input price, and thus redistributes domestic surplus and creates deadweight loss, but does not generate an international externality. If instead a foreign monopolist supplies the final...
good, then a domestic local content policy may redistribute surplus from the foreign monopolist to domestic input suppliers. Such a policy becomes more appealing to the domestic government when the foreign monopolist does not respond by significantly reducing output. In turn, a significant output reduction is less likely if the domestic country has market power (i.e., is large) so that its local content policy induces a fall in the foreign input price that offsets to some degree the rise in the domestic input price. The end result is that, for settings in which market power is present, local content policies may be unilaterally appealing to the domestic government and harmful to the foreign trading partner. From this perspective, restrictions on the use of local content requirements rest upon a solid economic foundation when market power is present.

Blanchard (2010) explores a different aspect of the relationship between foreign investment and trade agreements. She does not focus on investment measures and rules that restrict such measures; instead, she considers the broad implications of general cross-border equity holdings for optimal tariffs and the role of the GATT/WTO. Augmenting the two-country, two-good general equilibrium model of trade described above to include exogenous international cross-ownership, she identifies the channels through which cross-border ownership impacts the optimal tariff. The internal effect is that a country has less incentive to maintain a high tariff in the presence of a larger degree of foreign ownership in the local import-competing industry, and the external effect is that a government likewise has less incentive to raise its tariff for a terms-of-trade gain when its constituents hold a greater stake in the foreign export industry. Finally, the compositional effect is that industry bias in ownership patterns may encourage a government to manipulate local prices to benefit industries with a relatively higher proportion of national ownership.

As Blanchard (2010) argues, consideration of cross-border ownership leads to several interesting policy implications. An implication of the external effect is that a country may welcome foreign investment into its export sector as a means of encouraging a unilateral tariff reduction from its trading partner. Similarly, an implication of the internal effect is that foreign investment into an import-competing sector may encourage the host country to lower its import tariff as a means of extracting rent from foreign investors. Perhaps the most provocative implication of her analysis is that international ownership, by encouraging governments to liberalize their tariffs unilaterally, may substitute partially (or even in some cases completely) for negotiated tariff reductions. Indeed, with a sufficient degree of international cross-border ownership, unilateral tariffs are lower than

---

87 See Brander and Spencer (1981) for analysis of a related model, in which an import tariff is used to extract surplus from a foreign monopolist.
88 As Blanchard (2010) notes, her analysis of optimal tariffs under exogenous international cross-ownership generalizes and combines themes from previous theoretical work.
89 See also Blanchard (2007) and Krishna and Mitra (2005). Blanchard and Matschke (forthcoming) provide some evidence to this effect through an assessment of the impact of US multinational firms’ affiliate offshoring behavior on US tariff preferences.
90 While international ownership may lead in this way to lower tariffs, Gulotty (2014) argues that greater international ownership does not similarly lead to reductions in regulatory barriers. He argues that regulatory barriers raise fixed costs, and that the associated reduction in entry may lead to net gains for efficient, globalized firms. See also Ethier (1998) for a different perspective under which regional agreements and associated foreign direct investment activity arise endogenously in response to multilateral liberalization.
efficient, and the role of an international agreement is then to facilitate reciprocal and efficiency-enhancing restrictions in market access. Finally, Blanchard (2010) argues that the principle of reciprocity continues to serve as an important guide to efficient outcomes, once the definition of market access reflects ownership positions. More generally, the implications of cross-border ownership for the optimal design of GATT/WTO rules is an important subject that warrants further attention.

**Services** We briefly consider next the treatment of services in GATT/WTO. A variety of evidence confirms that services play an increasingly important role in modern economies. In their survey of research on services trade and policy, Francois and Hoekman (2010) indicate that the theoretical literature on services trade highlights the complementarity between international services trade and FDI, the implications of different market structures and national regulatory policies for the trade in services, and the way in which international service firms are organized. They also describe increasing evidence that services liberalization is a major potential source of economic performance gains.

Government policies that affect international service firms are disciplined in the WTO General Agreement on Trade in Services (GATS). As Trebilcock, Howse and Eliason (2013, Chapter 13, p. 480) indicate, the GATS is a “highly complex accord.” This agreement recognizes that services may be supplied through various modes and provides a framework for market-access negotiations across these modes. While GATS contains a general MFN provision (subject to some exceptions), national treatment and market access commitments apply only where WTO members make specific commitments to such coverage in their schedules. The impact of GATS on services reform is challenging to estimate. Francois and Hoekman (2010, p. 678) review the evidence to date and conclude that “the available, limited, evidence suggests that, with the exception of the European Union, most services policy reform has been unilateral. The contribution of the GATS to services reform has been negligible.”

Given the significant potential gains from services liberalization, it is natural to ask why the reciprocity mechanism that underlies trade agreements has not played a greater role in achieving policy reforms in services. Francois and Hoekman (2010) review the literature relating to this question and identify a number of potential factors. Among these factors, we mention here one that is related to our discussion above: given that FDI is a significant mode for supplying non-tradable services, Blanchard’s (2007, 2010) arguments imply that unilateral liberalization initiatives may substitute to some degree for reciprocal trade liberalization through trade-agreement negotiations. More generally, as Antràs and Staiger (2012a, 2012b) argue, the rise of offshoring may have changed

---

91 In relation to the Bagwell-Staiger (1999a, 2002) model described above, a key difference here is that international ownership operates via the external effect to diminish the absolute value of $W_{p}^{e}$ and $W_{p}^{f}$ and may even reverse the sign of these terms.

92 For further discussion, see Francois and Hoekman (2010) and Jensen (2011).

93 As Francois and Hoekman (2010, p. 678) note, countries that acceded to the WTO after 1995 tended to make more GATS commitments and represent a further exception, although care is required in assuming that GATS commitments are actually implemented. See also Eschenbach and Hoekman (2006) and Hoekman (2008).
the nature of the international externality that a trade agreement must address, which suggests in turn that trade agreements may require additional restrictions for services policies that are associated with offshoring. The purpose and design of trade agreements for market settings with offshoring is an important direction for research, which we discuss in further detail in Section 6.3.

Investment and services are of increasing importance in the international economy. The WTO includes agreements that place restrictions on measures that affect investment and services, but the appropriate nature of such restrictions is controversial. Drawing on research that employs the terms-of-trade approach to trade agreements, we argue that restrictions on the use of local content requirements rest upon a solid economic foundation when market power is present. We also summarize research that indicates that international ownership may substitute to some degree for negotiated tariff reductions. Finally, as we note, recent research suggests a significant potential for gains from services policy liberalization, even though the liberalization achieved to date through multilateral negotiations appears modest. To our minds, all of this points to a valuable role for future research directed toward understanding the impact of trade-related investment measures and services policies and, correspondingly, the appropriate design of WTO restrictions in this context.

4.6 The story line thus far...

As interpreted through the lens of the terms-of-trade theory, the original 1947 GATT was created to solve the central problem of the day: governments were setting trade policy noncooperatively under the “law of the jungle,” and the US imposition of the Smoot-Hawley tariffs in 1930 and the international retaliatory response that followed had led to a terms-of-trade driven Prisoners’ Dilemma which resulted in excessively high trade barriers. The task confronting governments was not to achieve multilateral free trade, but rather to set up an institution that could work well to internalize the international pecuniary (terms-of-trade) externality that was at the root of the high-tariff problem, and thereby induce governments to make the tariff choices they would have made had they not succumbed to the temptation of international cost-shifting (terms-of-trade manipulation) in the first place. If successful, GATT would lead necessarily to lower tariffs and expanded market access from those countries and in those industries where significant market power was present. But with significant market power not universal in all countries and all industries, GATT would not be expected to lead to universally lower tariffs from all countries and in all industries; and with the evident desire of governments to use trade policy for goals beyond that of simple national income maximization, GATT would certainly not be expected to lead to universal free trade. And finally, while the needed restraints on tariffs would necessitate some constraints on domestic policies to ensure that subsequent adjustments in those policies were not used to undo negotiated market-access commitments, GATT’s lack of any deeper integration beyond such constraints would not constitute a weakness or limitation of the GATT system.

The literature we have surveyed thus far lends broad support to the view that, at a fundamental level, governments succeeded with the GATT/WTO in creating an institution that is well-designed to solve the terms-of-trade problem. Many of the GATT/WTO’s core features appear sensible
when interpreted in the context of the terms-of-trade problem, and many of the outcomes that have been negotiated within the GATT/WTO seem consistent at a broad level with what might be expected from an institution that worked well to internalize terms-of-trade externalities.

The literature does, however, point out some potential difficulties with the GATT/WTO approach, and these difficulties may account for some of the central challenges that the WTO confronts today. The evolution in the treatment of subsidies from GATT to the WTO is especially puzzling from the perspective of the terms-of-trade theory, both with regard to domestic subsidies and even more so with regard to export subsidies. The theory does not provide support for treating export subsidies more severely than import tariffs.

Beyond the puzzling GATT/WTO treatment of subsidies, the literature emphasizes the possibility of a serious free-rider problem under the MFN principle, and there is some evidence in the literature that this problem is significant. And to the extent that the dual principles of MFN and reciprocity have together allowed countries that negotiate reciprocal cuts in their tariffs to appropriate the gains from their bargains and thereby keep free-riding to a minimum, the exemption from reciprocity granted to developing countries may ironically have kept these countries from enjoying to their full potential the benefits of GATT/WTO membership, again something that the evidence seems to bear out. This feature may in turn be contributing to a significant “latecomer” problem for the Doha Round, as the Round grapples with how to better integrate developing and emerging economies into the world trading system when the major developed countries have already achieved low average tariffs through decades of reciprocal MFN tariff bargaining among themselves.

Moreover, to the extent that there are significant deviations from the MFN principle, such as can arise with the formation of PTAs, the literature points to these deviations as complicating international externalities beyond the simple terms-of-trade problem that the GATT/WTO seems well-designed to solve, suggesting in turn that the rise of PTAs could be creating difficulties for the GATT/WTO approach and feeding a vicious cycle for the GATT/WTO of further PTA formation and further decline in the performance of the GATT/WTO. And finally, while services trade and international investment flows are of increasing importance to the global economy, the literature has developed only a nascent understanding of the nature of the international externalities associated with them, and so the ability of the GATT/WTO approach to function well in their presence is still an open question.

5 Evaluating the PTA Approach to Trade Liberalization

As we have remarked in earlier sections, impressive and sustained trade liberalization under the GATT/WTO during the last half of the 20th century has given way to PTAs as the new face of trade liberalization in the 21st century. We have described in the previous section how the GATT/WTO approach to liberalization derives broad support from the theoretical and empirical terms-of-trade literature. Does the terms-of-trade theory also support the view that PTA-driven liberalization can be seen as contributing to a solution to the terms-of-trade problem? In this section we review the
relevant theoretical and empirical literature and provide an answer to this question.

5.1 PTAs, External Tariffs and Multilateral Bargaining

We begin with a focus on the impact of PTAs on the external (MFN) tariffs of PTA-member countries, and ask: Might PTAs be seen to work in tandem with the tariff liberalization efforts of the GATT/WTO, or should PTAs be seen rather to work against these efforts? As we have observed, according to the terms-of-trade theory, non-cooperative Nash tariffs are set inefficiently high on products where the countries possess market power. Hence, one way to shed light on this question is to assess the impact of PTAs on the non-cooperative external MFN tariffs of the member countries on such products. If the formation of PTAs lowers these tariffs, then we could say that PTAs work in tandem with the GATT/WTO’s own efforts to reduce these tariffs, and that PTAs are hence building blocks for the needed multilateral liberalization that the GATT/WTO is also orchestrating. On the other hand, if the formation of PTAs raises these tariffs, then PTAs would appear to increase the degree of multilateral tariff liberalization that is needed to reach the efficiency frontier, and hence we could say that PTAs pose stumbling blocks to multilateral liberalization in this case.

There is a large literature that evaluates the impact of PTA formation on the external non-cooperative MFN tariffs of member countries. This literature has identified several important effects of PTAs on external tariffs, and it draws a distinction between PTAs that take the form of free-trade agreements (FTAs), and PTAs that take the form of customs unions (CUs) whose members not only eliminate tariffs on trade among themselves but also adopt a common external tariff policy toward the trade of nonmembers.

Two of the effects identified by this literature operate to reduce the external tariffs of PTA members: a “tariff complementarity effect” that can take two distinct forms and applies to both FTAs and CUs, and a “rent destruction effect” that applies to FTAs. Richardson (1995) identified a first tariff complementarity effect: when an FTA is formed between countries that are competing importers of a common product from third countries, each FTA partner has an incentive to lower its external tariff on this product slightly below that of its FTA partners so as to increase its share of the tariff revenue collected on imports from outside the FTA, and this competition for tariff revenue between FTA partners can lead to a downward spiral in their external tariffs. Bagwell and Staiger (1999b) identify a second tariff complementarity effect: they focus on a world where countries are competing exporters of a common product to third countries, and show that when FTA or CU partners reduce their tariffs to zero on imports of a product from one another, these partner countries each find it attractive to lower as well the tariff they apply to the imports of this product from third countries, in order to reduce the domestic distortions associated with discriminatory tariffs in their pursuit of terms-of-trade gains against these third countries.\footnote{Richardson (1993) describes a related effect that can arise in a small-country setting. We do not emphasize this result here, because according to the terms-of-trade theory the tariffs of a small country are internationally efficient, so the external-tariff-lowering impact of PTA formation identified by Richardson (1993) would not enhance international efficiency from a terms-of-trade theory perspective.}
effect is highlighted by Ornelas (2005a,b,c 2008), and is also a force for lower external tariffs among FTA members. In a setting where special interest lobbies push for tariff protection, Ornelas shows that the rents from the external tariffs of one FTA member country will spill over to producers in FTA partner countries, creating a free-rider problem for national lobbies within the FTA which interferes with their ability to obtain high external tariffs from their governments.

Two further effects operate to increase the external tariffs of PTA members, but operate only in the presence of CUs. A “market power effect” (see Kennan and Reizman, 1990, Krugman, 1991, Bond and Syropoulos, 1996a,b, Bagwell and Staiger, 1997b and 1999b, and Cadot et al, 1999) arises when CU members are competing importers of a common product, and can collectively exert more market power on the world price of that product with their common external tariff than they could individually. And a separate “coordination effect” (see Kennan and Reizman, 1990) operates to raise the external tariffs of CU members as well, even when countries are competing exporters: when one CU member raises its external tariff on a product that it imports from third countries, other CU members that export that product to the first CU member will gain as they receive higher prices for their exports into the first CU member’s market, and this is a positive externality of higher external tariffs that can be internalized among members of a CU when they set their common external tariff policy.

Of course, while the impact of PTA formation on non-cooperative Nash external MFN tariffs is suggestive of the nature of the relationship between PTAs and the GATT/WTO, it provides at best an incomplete picture of this relationship, for several reasons. First, it is not clear that the impact of PTAs on non-cooperative external MFN tariffs translates – even with the same sign – to the impact on cooperative MFN tariffs. For example, Limão (2007) shows that an FTA that pursues non-trade objectives can result in higher cooperative external tariffs, in circumstances where FTA partners agree to provide non-trade concessions to a country who values those concessions in exchange for preferential market access rents created and maintained by the high external tariffs of the country. In effect, Limão demonstrates that the FTA partners can become a force against negotiated MFN tariff cuts of the country valuing the non-trade objectives. And Bagwell and Staiger (1997a, 1997b, 1999b) show that when self-enforcement constraints bind in a multilateral agreement over external tariffs, the formation of FTAs and CUs can have impacts on the most-cooperative MFN tariffs achievable that are the opposite of the impacts on the Nash external tariffs, and that vary through time if the PTAs are implemented in stages.95 Second, recall that the terms-of-trade theory directs attention to the question whether PTAs reduce the degree of inefficient terms-of-trade manipulation embodied in the external tariffs of PTA partners, and the results surveyed above are not always presented with this question in mind.96 And third, while a reduction in the external MFN tariffs

95 Other papers that consider the impact of PTAs on the prospects for the cooperative setting of MFN external tariffs include Levy (1997) and Krishna (1998), both of whom argue that FTAs can erode the political support for further agreements to reduce MFN external tariffs and thereby act as stumbling blocks to multilateral trade liberalization. A literature also exists that examines the impact of multilateral liberalization on the formation of PTAs. Using different frameworks, Ethier (1998) and Freund (2000a) argue that PTAs may be a response to successful multilateral liberalization.

96 For example, the rent destruction effect identified by Ornelas (2005a,b,c 2008) does not have a clear prediction
of PTA members triggered by the formation of the PTA might be viewed as partially solving the terms-of-trade problem and thereby making the remaining task easier for the GATT/WTO, Bagwell and Staiger (1999a, 2001a) show that the introduction of PTAs and the violation of MFN that this implies can change (does change for FTAs, can but need not change for CUs) the nature of the problem that a trade agreement must solve, from a simple terms-of-trade problem to a more complicated problem in which international externalities travel through local-market prices in addition to the terms-of-trade. In effect, for this reason Bagwell and Staiger argue that PTAs are inherently at odds with the GATT/WTO’s approach to multilateral trade liberalization, which as we have described above seems best-suited to address simple terms-of-trade problems.

In any case, with these various effects identified in the literature and pointing in different directions, it is clear that theory alone cannot resolve the issue of the impact of PTA formation on the external MFN tariffs of member countries. Freund and Ornelas (2010) offer this summary of the literature on this point:

Putting all these arguments together, one reaches at least three conclusions. First, it is a safe bet that external tariffs will change after the formation of a PTA. There are just too many plausible arguments indicating that governments have incentives to do so, in one direction or the other, regardless of their motivations. Second, the changes are likely to be different in FTAs and CUs. Although in general the incentives point to lower external tariffs in FTAs than in CUs, it is possible to write down models in which the reverse is true. Third, theoretical work alone cannot determine the direction of the change, and even less so the magnitudes. Therefore, it falls to empirical work to establish which forces prevail, how important they are, and how CUs differ from FTAs in that respect. (Freund and Ornelas, 2010)

We therefore turn to the empirical literature on this question.

What is the impact of PTA tariff liberalization on subsequent efforts toward multilateral tariff liberalization? Limão (2006) provides a first product-level investigation into whether PTAs are stumbling blocks or building blocks for subsequent multilateral liberalization. His approach involves a comparison of two different types of products – those for which a country has positive imports from PTA partners and those for which it only imports from PTA non-partners. An examination of subsequent US multilateral tariff changes made as a result of the Uruguay Round documents evidence that the US granted smaller MFN tariff reductions in products with positive US imports from PTA partners. Importantly, the evidence applies not only to products imported from “large” PTA partners, such as the countries in NAFTA, but also to imports from smaller PTA partners. Given that even those smaller US PTA partners export in nearly 15 percent of product lines, and that these products also have positive levels of imports from non-PTA partners, a further implication is that even small US PTAs were creating a stumbling block effect on potential multilateral

---

97 See Freund and Ornelas (2010) for a more extensive survey of this empirical literature.
liberalization taking place under the Uruguay Round.\textsuperscript{98}

In addition to the United States, another important environment to conduct such an exercise is the European Union. Karacaovali and Limão (2008) first confirm the evidence found for the United States by determining that the EU cut MFN tariffs on products not imported from PTA partners by nearly twice as much as it cut tariffs on products imported from partners during the Uruguay Round. Furthermore, the size of the EU stumbling block effect is larger for the products that are exported by more PTA partners. Second, they exploit additional margins of the data on the EU PTAs by splitting PTA partners based on whether the country eventually acceded to the European Union between the ends of the Tokyo Round and Uruguay Round of negotiations. They also find evidence consistent with theoretical predictions that accession countries should not trigger stumbling block effects,\textsuperscript{99} and that the stumbling block effect is only associated with the products from countries with which the EU had PTAs in place at that time.

Changes to multilateral tariffs need not only take place in the context of GATT/WTO negotiating rounds. It is natural, however, to analyze tariff changes achieved through negotiation rounds when considering the US and EU during the recent period in which detailed tariff data are available. This is true for two reasons. First, for most products, US and EU applied MFN tariffs are relatively close to their legal tariff bindings so that the tariffs cannot be increased without violating multilateral commitments. Second, US and EU tariffs were also relatively low to begin with during this period, and so there is also not much scope for downward variation in the form of additional unilateral reductions. However, these two conditions do not apply to a number of developing countries in the international trading system which had both i) sufficiently high applied MFN tariffs at the time of PTA implementation to allow for the possibility of meaningful unilateral reductions, and ii) legal binding commitments sufficiently above their applied MFN rates to allow applied rates to legally increase as well without violating these commitments.

Estevadeordal, Freund and Ornelas (2008) exploit these features of the data for a sample of 10 Latin America countries over 1991-2000. They assess patterns of applied MFN tariff changes following the implementation of PTA tariff reductions, under agreements like MERCOSUR, the Andean Community, and other intra-Latin American PTAs formed during the decade. In particular, these countries exhibit, on average, a positive relationship between changes in preferential tariffs and subsequent changes to applied MFN tariffs – evidence that Latin America’s preferential agreements worked as a building block toward unilateral MFN liberalization during this period. Higher shares of intra-PTA imports are also associated with reductions in applied MFN tariffs even for relatively small preference margins. However, the results are limited to the free trade areas and do not hold for Latin America’s trade agreements that were formulated as customs unions during that period.

\textsuperscript{98}Limão (2007) provides a theory to explain this – i.e., that the United States may offer and sustain preferential tariffs to certain countries in order to obtain non-trade objectives such as upholding higher environmental or labor standards, intellectual property rights protection, fighting the war on drugs – and confirming evidence from US PTAs through the Andean Trade Promotion and Drug Eradication Act and Caribbean Basin Initiatives.

\textsuperscript{99}In their theoretical model, accession countries are eligible to receive a transfer or revenue collected under the common external tariff which offsets their potential loss in intra-PTA trade that would be associated with additional multilateral tariff reduction.
such as MERCOSUR.

In a related approach, Calvo-Pardo, Freund, and Ornelas (2011) assess the evolution of multilateral tariff changes following the preferential tariff reductions associated with the creation of the ASEAN (Association of Southeast Asian Nations) Free Trade Area in 1992. ASEAN is another setting in which preferential liberalization led to applied MFN tariff cuts, with MFN tariff cuts found to be larger in products with larger preference margins and thus a greater scope for trade diversion.\textsuperscript{100}

What are the potential explanations for the differences across settings? We speculate that one potential contributor is government policy responsiveness to the threat of trade diversion. The Estevadeordal, Freund and Ornelas (2008) and Calvo-Pardo, Freund, and Ornelas (2011) settings resulted in original PTA liberalization that led to large preference margins (relative to pre-PTA applied MFN rates), and economically costly trade diversion could have arisen if governments did not subsequently also lower their applied MFN tariffs. The US and EU environments, on the other hand, subject to the Limão (2006) and Karacaovali and Limão (2008) studies, resulted in preference margins that were much smaller with potentially less scope for trade diversion. Second, the US and EU preferences were more unilateral in nature. The theory in Limão (2007) stresses the non-reciprocal nature of US and EU preferences and that they were offered as compensation for countries that took up non-trade obligations in areas such as environmental or labor standards, intellectual property rights protection, and supporting the war on drugs. An open question for research is whether building block effects may be more likely to dominate in reciprocal PTAs. The rising importance of “WTO-extra” provisions in PTAs (Horn, Mavroidis, and Sapir, 2010) suggests that this should be a priority area for additional research.

It should also be noted that preferential tariffs and MFN tariffs are certainly not the end of the line when it comes to trade policy, as there are a number of other potential non-tariff barriers to trade. The GATT/WTO provides a number of exceptions that countries can invoke to implement higher levels of protection for legitimate environmental, health, or other public safety concerns, for example under Article XX. Furthermore, most of the same major economies involved in multilateral and preferential trade liberalization since the late 1980s are also major users of antidumping and safeguards (Bown, 2011a) that are another major form of GATT/WTO exceptions allowing countries to temporarily implement higher levels of import protection under certain conditions.\textsuperscript{101}

\textsuperscript{100}Tovar (2012) is a third study examining how developing countries make unilateral changes to applied MFN tariffs after formation of a PTA through an examination of four countries after the formation of the CAFTA-DR in 2004. The results for CAFTA-DR are somewhat different than the earlier studies as they suggest at least an initial stumbling block effect. Countries increase (or decreased by less) the MFN tariffs in product categories that had previously been subject to larger preferential tariff reductions in the first two years after PTA implementation, though this is somewhat offset in the second and subsequent two year period during which the countries adjusted their MFN tariffs in a downward direction.

\textsuperscript{101}Bown, Karacaovali, and Tovar (forthcoming) provide a more general survey on relationships between PTAs and the temporary trade barriers of antidumping and safeguards. They also argue that one of the reasons behind lack of empirical progress in this particular research area is because it requires details on how countries are applying the TTB policies in practice, and such details are often lost when the details are aggregated up to the industry or country level. To illustrate, they provide a set of case studies describing how PTA members have applied the global safeguard policy in order to document the range of outcomes that have occurred. One extreme involves PTA members imposing
While the general relationship between PTA liberalization and non-tariff barriers (NTBs) use is not yet well understood, here we highlight a recent paper that has initiated the investigation of this relationship.

Antidumping is the most frequently applied temporary trade barrier policy in use across countries and over time since the 1980s. Prusa and Teh (2010) use a cross-country sample involving 80 PTAs and antidumping use dating back to 1980. While there is only a modest impact of PTA formation on the overall use of antidumping, after controlling for other aggregate-level determinants, there is evidence of important differences in policy treatment between PTA partners and non-partners. PTA implementation is associated with both a reduction in antidumping against new PTA partners and an increase against non-partners. They attribute some of this post-PTA change in behavior to PTA variation in the significance of legal provisions affecting antidumping use - i.e., the sort of “WTO-plus” provisions described in more detail in section 2.2. In any event, the Prusa and Teh findings suggest that reorientation of temporary trade barriers toward the imports of non-PTA partners could be an important avenue by which PTAs lead to rising external tariffs.

Finally, in addition to the literature we have surveyed just above, there is an active related literature that formally models the choice countries face between PTA formation and multilateral trade bargaining. That literature is concerned with the question whether global free trade is more or less likely to be achieved when PTAs are an option that is available to governments as an alternative to multilateral tariff negotiations. Guided by the terms-of-trade theory, our focus here is on a related but distinct question, namely, whether PTAs can contribute to (building blocks) or rather interfere with (stumbling blocks) the ability of multilateral negotiations to achieve globally efficient policy outcomes when judged against the governments’ own preferences. With regard to this statement of the stumbling block/building block question, the findings of this related literature can be summarized as follows. First, when there are no bargaining frictions, as is assumed by most of the literature (a point emphasized by Maggi, forthcoming, and which we have noted above), global efficiency would always be achieved under multilateral negotiations if PTAs were banned, and so PTAs can’t possibly be building blocks in the sense we are interested in here. In this case PTAs can facilitate the attainment of global free trade, but only when global free trade does not mark a Pareto improvement over the outcome that would be delivered if PTAs were banned (see, for example, Aghion, Antràs and Helpman, 2007, and Saggi and Yildiz, 2010). Nevertheless, in this no-bargaining-frictions case PTAs have been shown to be stumbling blocks to Pareto efficient

\footnote{We note, however, two qualifications. First, as Freund (2000b) argues in an oligopoly context, the path by which global free trade is achieved may matter. In her model, world welfare is higher when global free trade is achieved through expanding preferential agreements rather than through multilateralism. Second, and as we discuss briefly in the next section (see note 110), a possible commitment theory interpretation of these findings could suggest that PTAs enhance efficiency when viewed from an ex-ante perspective.}
outcomes under certain conditions (see, for example, Aghion, Antràs and Helpman, 2007, Seidman, 2009, and Saggi, Yildiz and Woodland, 2013). Second, when there are bargaining frictions (as in McCalman, 2002, and McLaren, 2002) and where a building block role for PTAs is therefore possible, no such building block role has been found, but a stumbling block role has again been shown to be possible. Hence, while these papers have shown that PTAs can serve as a building block for the attainment of free trade, if anything this branch of the building-block/stumbling-block literature reinforces the view that PTAs should be viewed with some skepticism from the perspective of global efficiency when judged against the governments’ own preferences.

5.2 PTAs and Third-Country Externalities

According to the terms-of-trade theory, the purpose of a trade agreement is not to secure free trade, but to allow governments to internalize the terms-of-trade externalities associated with their tariff choices, and in so doing to remove the internationally inefficient cost-shifting component from their unilateral tariff choices. As we have described in earlier sections of our survey, the GATT/WTO appears well-equipped to help governments internalize terms-of-trade externalities, in part through its norms of reciprocity and MFN which we have argued can help to keep the terms-of-trade consequences for third countries to a minimum when subsets of countries negotiate tariff cuts. PTAs, by definition, deviate from the MFN norm, raising the possibility that, rather than contributing to a solution to the terms-of-trade problem, PTAs are a surviving vehicle for imposing terms-of-trade externalities on third countries within the GATT/WTO system. Under this possibility, some PTAs may be viable, in the sense that their member governments support their formation, only because they have been able to use discriminatory tariff cuts between them to impose negative terms-of-trade externalities on third countries and convert those third-country loses into their own gains. This possibility is emphasized by Bagwell and Staiger (2005a), and it provides one reason according to the terms-of-trade theory why the proliferation of PTAs could reflect a development which is inefficient from a global perspective.

A necessary feature for PTAs to impose negative terms-of-trade externalities on third countries is that the discriminatory market access granted to PTA partners diverts trade volumes that would otherwise have occurred between PTA member countries and third countries. This “trade diversion” effect of a PTA is the trade volume reduction that can lead to changes in trade prices with third countries, and hence to third-country terms-of-trade impacts. A number of papers have emphasized the likelihood that PTAs rely on substantial trade diversion in order to keep them viable (see, for example, Grossman and Helpman, 1995b, and Krishna, 1998; and see Ornelas, 2005a,b,c for qualifications to this claim). The possibility described above is a particular version of this claim, in which the third-country terms-of-trade externality associated with trade diversion is the mechanism by which an otherwise nonviable PTA is kept viable.

What is the evidence regarding the importance of third-country externalities imposed by PTAs? As we next describe, the evidence is mixed, with some studies finding substantial trade diversion and terms-of-trade impacts of PTA formation on third countries and other studies finding only
insignificant effects.

Chang and Winters (2002) take up the international externality question by investigating the experience of third-country export prices to Brazil in light of MERCOSUR. Relying on product-level unit values data to proxy for export prices, intra-PTA tariff reductions are empirically associated with the price declines of third-country (Chile, Korea, Japan, US) exports to Brazil, relative to the prices of these third-country’s exports of the same products to the rest of the world. Furthermore, welfare calculations arising from the model’s estimates indicate PTA non-partner countries such as the US and Germany experienced sizeable welfare losses due to the price declines, even after taking into account the effects of Brazil’s subsequent MFN tariff reductions in many of the same products. Winters and Chang (2000) present a similar approach by examining the impact of Spain’s 1986 EC accession on US and Japanese exports to Spain. They argue that these earlier results are not as strong due to methodological and data issues, including the reliance on data at higher levels of aggregation. Nevertheless, results from this study are consistent with the Chang and Winters (2002) evidence from MERCOSUR. In particular they find that each 1% preferential Spanish tariff cut toward new PTA members is associated with roughly a 0.5 percent export price decline for Japanese and US exporters to Spain, relative to these new PTA partners’ (France, Germany, Italy, UK) export prices of the same good.

That important negative terms-of-trade externalities arise after PTA formation is consistent with some, though certainly not all, of the evidence arising from other studies. Romalis (2007), for example, found that the EU’s trade with the US, Canada, and Mexico was negatively impacted by the implementation of NAFTA, confirming significant trade diversion effects, but found little in the way of third-country price impacts associated with these trade volume reductions. Both Romalis (2007) and Clausing (2001) found only insignificant trade diversion effects from the Canada-US FTA, a result consistent with the analysis of Trefler (2004), who found that Canada’s trade creation associated with CUSFTA dominated the welfare effects of any trade diversion. Using data on the manufacturing trade and FTAs for 64 countries over the period 1990-2002, Dai, Yotov and Zylkin (2014) report large trade diversion impacts of FTA formation, with the largest third-country impacts suffered by existing FTA members when a country joins a new FTA from which its other FTA partners are excluded. On the other hand, Frazer and Van Biesebroeck (2010) did not find significant evidence that US preferences under AGOA drew African exporters trade in those products away from the EU in ways that may have affected EU consumer prices. Furthermore, there is also mixed evidence on international externalities associated with the application of other discriminatory trade policy, such as antidumping. Bown and Crowley (2006, 2007) find international externalities associated with US antidumping imposed on Japanese exports via the trade volume (“trade deflection”) and price (third-market terms-of-trade) effects on its exports of those products to third markets. However, Bown and Crowley (2010) investigate similar trade restrictions imposed on a developing country exporter (China, during 1992-2001) and do not find evidence of trade deflection to third countries in that setting.
Finally, in recent work, Caliendo and Parro (forthcoming) and Tintelnot (2014) build on the Eaton and Kortum (2002) framework to quantify the trade impacts of NAFTA (Caliendo and Parro) and the trade and production impacts of the proposed Canada-EU trade and investment agreement (Tintelnot). Caliendo and Parro find small price and quantity impacts of NAFTA on the rest of the world, while Tintelnot finds larger potential third-country impacts of the proposed Canada-EU agreement. Tintelnot’s findings may be especially relevant for assessing the impact of the new wave of PTAs to impose third-country externalities, as he focuses on the deep-integration features of the proposed agreement and their implications for the location of multinational production facilities for purposes of export platform.

Taken together, we interpret the literature as indicating that the potential for PTAs to impose important third-country externalities is real and has at times probably been exercised, but that it does not appear to be a pronounced and consistent feature of existing PTAs. Of course, it is possible that PTAs have mostly not imposed significant third-country effects at least in part because of the role of the GATT/WTO dispute procedures in policing such effects, a possibility that is given some credence in light of the literature we survey in section 7.3. In any event, the literature surveyed above establishes that the potential for third-country externalities is there, and is suggestive of the possibility that the problem could become more substantial with the increasing focus of new PTAs on deep integration. For this reason, maintaining a healthy skepticism regarding PTAs because of this potential seems warranted.

5.3 PTAs and Deep Integration

We now turn our focus to the growth in deep-integration PTAs. Does the terms-of-trade theory provide support for this development? We have already described how the theory can provide support for shallow integration; this is the basic message of Bagwell and Staiger (2001b)’s claim that a well-working market-access-preservation rule can allow country’s to negotiate to efficient outcomes without directly negotiating over domestic policies (i.e., without engaging in negotiations over deep integration); and as we have described, that message survives a variety of generalizations of the original model in which it was made. But one environment where the terms-of-trade literature points to qualifications of this message is when governments approach trade negotiations with private information. In Bajona and Ederington (2011), the private information is over the degree of a government’s intervention with domestic policies, and there the case for shallow integration survives largely intact, albeit with a rigid tariff binding being an important component of the optimal agreement in combination with a market access preservation rule. But if private information instead concerns a government’s type (e.g., the magnitude of a domestic distortion associated with an externality), as in the model of Lee (2014), and if the type is uncertain at the time the agreement is written, then the optimal agreement with a market access preservation rule would by necessity require a state-contingent market access level, which is not possible when the state-realization is privately observed. In this setting, Lee shows that a form of deep integration is needed to construct the optimal agreement.
Hence, the terms-of-trade literature suggests that the case for shallow integration can be weakened to the extent that private information is important in trade negotiations, depending on the source of the private information. This raises the possibility that if the WTO is struggling with deep integration, PTAs might be helping to provide the deep integration that governments need but can’t achieve at the WTO. What is missing from this argument is a reason why the WTO would struggle with deep integration while PTAs would not; after all, the underlying problem is still the terms-of-trade problem in these models, and so Nash domestic policies are set efficiently whether or not there exists private information across governments, with the inefficiency amounting simply to a level of market access that is too low. Why, then, would PTAs be better than the WTO at providing the degree of deep integration necessary to achieve efficient market access levels in the presence of private information? And if PTAs do not have such an advantage over the WTO, then there is an alternative possibility that gains prominence as a possible explanation for why PTAs seem to succeed at deep integration where the WTO has struggled: the deep integration being achieved in the web of emerging PTAs is simply inefficient from a global perspective, and hence such an agreement would not be expected in a multilateral institution such as the GATT/WTO. To date, the terms-of-trade literature has not gone far enough to sort through these possibilities and provide answers. We therefore see this as an important open question for the literature, but until it is resolved more definitively we view the broad affinity between the GATT/WTO design and solutions to the terms-of-trade problem as providing a presumption that any deep integration that would be required to achieve efficiency (say, due to the particular sources of private information) is likely better provided within the GATT/WTO than by PTAs.

5.4 The story line continued...

As WTO liberalization efforts seem to have stalled, PTA liberalization has taken off. Has the explosion of PTAs interfered with the WTO’s ability to deliver countries to the global efficiency frontier? Or are PTAs carrying countries to the global efficiency frontier in ways that the WTO could not? Or, are PTAs succeeding where the WTO could not because PTAs can impose costs on third countries that WTO rules successfully internalize, in which case PTAs are likely moving the world away from the global efficiency frontier?

Our survey of the terms-of-trade literature suggests a cautious interpretation of the benefits of PTAs to the world trading system. As we have emphasized, this literature provides broad support for the view that the GATT/WTO is fundamentally well-designed to minimize the influence of terms-of-trade externalities on the policy choices of member governments and thereby solve the terms-of-trade problem. The literature provides a more mixed view of PTAs in this regard, with theory pointing out many avenues through which PTAs could permit terms-of-trade externalities to re-enter the calculus of trade-policy making, and empirical evidence providing only partial assurance that these avenues have not been exercised. In this sense, the terms-of-trade literature supports a skeptical view of the wisdom of entrusting PTAs with the rules of globalization.

The terms-of-trade literature also provides ample reasons for caution concerning the position
that PTAs are complementary to the GATT/WTO. While both theory and empirical evidence are mixed, it would seem difficult based on this literature to make a strong case that PTAs have not to some extent limited the ability of the GATT/WTO approach to deliver and maintain an efficient outcome.

In short, according to the terms-of-trade theory, the GATT/WTO looks to be at its core a well-designed trade policy institution for guiding governments to the global efficiency frontier, and the explosion of PTA activity should be viewed with caution and some skepticism. We next survey the three remaining branches of the trade agreements literature and assess the qualifications to this position that are implied by those literatures.

6 Beyond the Terms-of-Trade Theory

Thus far we have relied on the terms-of-trade theory, and the strand of the literature that builds on this theory, to evaluate the relative merits of the GATT/WTO and PTA approaches to liberalization, and to provide answers to the questions posed at the beginning of this article. And we have reviewed an empirical literature that lends support to the relevance of this theory. We now survey the other strands of the trade agreements literature, namely the literature on the commitment, delocation/profit-shifting and offshoring theories of trade agreements. Our purpose is to identify insights from each of these strands of the literature that suggest qualifications to the answers provided by the terms-of-trade theory.

6.1 The Commitment Theory

The most established alternative to the terms-of-trade theory of trade agreements is the commitment theory. According to the commitment theory, governments value trade agreements as a way to tie their hands against their own lobbies and citizens. Of course there is no reason why trade agreements couldn’t serve multiple purposes, which is to say the commitment and terms-of-trade theories need not be mutually exclusive and could be viewed as complements. The question for us here is whether the commitment theory offers a more supportive interpretation of PTAs relative to the GATT/WTO than does the terms-of-trade theory and, if so, whether there is sufficient empirical support for the commitment theory more generally to qualify or alter the initial conclusions we have drawn concerning the relative merits of PTAs and the GATT/WTO from our survey of the terms-of-trade theory literature.

While expressions of the commitment theory of trade agreements can be found in a variety of early papers (see, for example, Carmichael, 1987, Staiger and Tabellini, 1987, Lapan, 1988, Matsuyama, 1990, Tornell, 1991 and Brainard, 1994), a particularly elegant treatment that has become the workhorse model of this idea is provided by Maggi and Rodriguez-Clare (1998). Their model is one of a small open economy, where the terms-of-trade argument for trade agreements is absent. The focus of the model is on the idea that an anticipated trade-policy-lobbying relationship between a government and producers in some sector is likely to distort the equilibrium allocation
of resources in the economy toward that sector, and on the possibility that the lobby might then not fully compensate the government for this distortion. To formalize this possibility, Maggi and Rodriguez-Clare extend the lobbying model of Grossman and Helpman (1994) to include a prior stage in which resources in the economy are allocated across sectors. They confirm that the government will be compensated by the lobby for the ex-post distortions its trade-policy choice imposes on the economy given the sectoral allocation of the economy’s resources that are sunk at the time the trade policy choice is made; this finding is the same as in the original Grossman and Helpman model. But Maggi and Rodriguez-Clare establish that 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. This second finding is novel, and as Maggi and Rodriguez-Clare show, it provides a reason that the government might (under certain conditions which they explore) wish to tie its hands ex ante against influence by the lobby ex post. And in this way, a possible commitment role for a trade agreement is thereby identified.\footnote{See also Mitra (2002), for a similar commitment story where the avoidance of wasteful lobbying resources rather than distorted sectoral allocations is the driving factor that motivates governments to use trade agreements as a commitment device. As Maggi (forthcoming) observes, commitment arguments can also serve as reasons against joining a trade agreement, as the paper by McLaren (1997) elegantly illustrates.}

The commitment theory has been used to offer interpretations of some of the features of the GATT/WTO which appear puzzling when viewed through the lens of the terms-of-trade theory. One example is Potipiti (2012), who uses the commitment theory to explain why, in the WTO, tariffs are the subject of negotiated limits while export subsidies are banned outright. From the perspective of the terms-of-trade theory, this feature is puzzling on two levels, as we have observed above: first, if anything the standard terms-of-trade theory would suggest encouraging export subsidies, not discouraging them; and second, why the asymmetric treatment across instruments, with the allowable levels of one instrument left to negotiations and the other banned outright? Potipiti shows that these puzzles can be resolved in the context of the commitment theory, once an asymmetry between the rents earned by import-competing and exporting interests is introduced.

Potipiti (2012) builds on the small-open economy model of Maggi and Rodriguez-Clare (1998), where the anticipation of protection generates inefficient investment ex ante for which the government is not compensated in its (ex-post) political relationship with the industry. In Potipiti’s version of this model, a government can join an agreement that bans tariffs and/or an agreement that bans export subsidies, and doing so will eliminate the anticipation of protection by the private sector and the associated investment distortion, and thereby generate a social welfare gain. On the other hand, commitment to such an agreement also requires the government to forfeit the political contributions it would otherwise collect for the protection it offers. The government therefore faces a trade-off, and it commits to a trade agreement covering a particular policy (import tariff and/or export subsidy) only if the social welfare gain from banning the use of that policy is greater than the government’s valuation of the associated loss in political contributions. Applied to export policies, the underlying Maggi and Rodriguez-Clare model can therefore immediately account for an agreement that discourages (bans) export subsidies. But how can the asymmetry in treatment
across import tariffs and export subsidies be understood? Potipiti argues that this asymmetric
treatment can arise from an underlying asymmetry in growth prospects of the two sectors that he
shows occurs when trade costs are decreasing through time, and from the differences in the rent-
generating capacity of protection in (expanding) export and (declining) import-competing sectors
that this implies. As Potipiti demonstrates, because of the relative inability for protection to create
rents in expanding as opposed to contracting sectors, for a range of model parameters it is optimal
for a government to agree to a ban on export subsidies and thereby give up the (smaller) political
rents in favor of the social welfare gain, while at the same time not banning import tariffs and
instead opting to retain the (larger) political rents that their use generates.\footnote{Potipiti's (2012) model can explain why export subsidies might be banned while import tariffs are not banned, but it doesn't actually explain why some limits on tariffs might still be negotiated. However, it is not hard to see that introducing a small amount of terms-of-trade motive into the model (by relaxing slightly the small country assumption) could provide a reason for negotiating tariff bindings while not altering the other results of the model.}

The commitment theory has also been used to interpret the evolution of rules on domestic
subsidies from GATT to the WTO, an evolution that as Bagwell and Staiger (2006) demonstrate
does not find easy support from the perspective of the terms-of-trade theory. Here the relevant
paper is Brou and Ruta (2013), who augment the Maggi and Rodriguez-Clare (1998) model by
allowing the domestic government to use both a production subsidy and an import tariff in its
relationship with the lobby in an import-competing sector. Taxation to raise revenue is assumed to
be distortionary, so that a production subsidy does not dominate a tariff for redistributive purposes;
rather, as Brou and Ruta show, in this setting optimal intervention will typically include a mix of
tariffs and production subsidies.

In the model of Brou and Ruta (2013), the fundamental reason for signing a trade agreement
that commits a government to free trade is the same as that in Maggi and Rodriguez-Clare (1998)
and in Potipiti (2012). But there are two novel twists in the model of Brou and Ruta. First, the
lobby's anticipation of both a tariff and a domestic subsidy serves to create ex-ante distortions that
the government will not be compensated for ex post, and so the government may have a direct
reason to sign agreements which constrain both tariffs and domestic subsidies, rather than simply
an agreement that constrains the tariff. And second, if a government does sign on to an agreement
that constrains its tariff only, this commitment will induce the government to simply turn more
intensively to production subsidies in its political relationship with the import-competing lobby –
what Brou and Ruta term “the policy substitution problem” – and the resulting distortions
are themselves welfare-reducing. What Brou and Ruta show is that in the presence of a tariff-only
commitment the new subsidies associated with the policy substitution problem can be handled with
a “nullification-or-impairment” rule, offering support for the GATT shallow-integration approach
to domestic subsidies in much the same way that the terms-of-trade theory supports GATT's
approach to domestic subsidies (with what we have termed above a “market access preservation
rule”). But Brou and Ruta show as well that according to the commitment theory, there is a
remaining distortion associated with the original subsidy level (just as with the tariff) that a tariff-
only agreement in combination with the GATT nullification-or-impairment rule cannot address.
It is with this second finding that Brou and Ruta demonstrate that the commitment theory can provide support for the WTO’s new disciplines imposed directly on domestic subsidies, and in this sense provide support for (at least some) movement toward deep integration.\(^{106}\)

The Brou and Ruta (2013) finding that a desire for commitment can help explain the “deep integration” features of the WTO treatment of domestic subsidies – and by extension, perhaps the value of deep integration more generally – raises the possibility that the ongoing pursuit of deep integration through PTAs might be seen in a more positive light than that suggested by the terms-of-trade theory. This is especially true when viewed along side results of Maggi and Rodriguez-Clare (2007), a large-country version of Maggi and Rodriguez-Clare (1998) which combines the commitment and terms-of-trade theories. Among their results, Maggi and Rodriguez-Clare find in a dynamic version of their model that it is optimal for governments to separate their liberalization into two phases, a first (and in their model instantaneous) phase in which liberalization reflecting the elimination of terms-of-trade motives occurs, and then a second (and in their model gradual) phase in which further liberalization to handle the domestic commitment motive occurs. In the Maggi-Rodriguez-Clare model, only tariffs are considered. But in light of Brou and Ruta’s demonstration that commitment motives can give rise to the need for deep integration, it is tempting to conjecture that, if behind-the-border policies such as the domestic subsidies considered by Brou and Ruta were added to the Maggi-Rodriguez-Clare model, the resulting model might yield predictions that could support, as an optimal development, the gradual spread of deep integration that PTAs appear to be delivering, but only after terms-of-trade considerations had been removed from tariff choices; that is, a period of GATT-like shallow integration to address terms-of-trade motives, followed by the spread of deep integration to aid governments in their desire to make domestic commitments and possibly led by PTAs. Notice, though, that even if this conjecture were to pan out, it would not support the view that the WTO is passé, but rather would at most give additional weight to the view that the WTO and PTAs are complementary: there would still be a continued important role for an institution that is well-designed to handle the terms-of-trade problem, as we have argued is the case for the GATT/WTO.

Finally, we note that Ethier (1998) employs the commitment theory to take on directly the question whether the emergence of PTAs following a period of multilateral liberalization might be viewed as a positive development for the world trading system. Actually, like Maggi and Rodriguez-Clare (2007), Ethier combines two reasons for a trade agreement into one model. One of these reasons can be interpreted as the commitment motive, and in Ethier’s model this is the motive that best describes why governments might be interested in PTAs (see especially Ethier’s discussion

\(^{106}\)Like Brou and Ruta (2013), Limão and Tovar (2011) also study the role of trade agreements as a commitment device when governments can use both tariffs and behind-the-border policies to redistribute to favored groups. But the focus of Limão and Tovar is on whether a government might wish to constrain its use of a more efficient instrument (in their model the tariff) knowing that this would result in more reliance on a less efficient instrument (in their model reduced form behind-the-border non-tariff barriers). As they demonstrate, a government might find such a commitment desirable despite the associated efficiency costs because the commitment can improve its bargaining power relative to the lobby. Limão and Tovar do not consider the possibility that international commitments might be extended to cover behind-the-border non-tariff barriers, so unlike Brou and Ruta their model does not yield insights about the desirability of deep integration.
on pp. 1240-41). The second reason for a trade agreement in Ethier’s model is an international externality, but it is not the terms-of-trade externality; rather, it is a (Marshallian) scale economy that operates at the world-wide level and creates a positive international (non-pecuniary) externality associated with greater investment. This form of international externality is what underpins the purpose of a multilateral trade agreement in Ethier’s model.

Ethier’s (1998) model is meant to capture the forces behind the growth in numbers of PTAs beginning in the 1990s that involved large developed countries forming PTAs with small reforming developing countries. In Ethier’s model, attracting foreign investment from the developed world is by assumption a necessary condition for success of reform in a developing country, and PTAs are ways in which developing countries compete among themselves for the required foreign investment. In essence, according to Ethier, a PTA with a large developed country can provide the necessary commitment device for the developing country to credibly “lock in” its reforms with commitments to deep integration: these deep integration commitments attract foreign investors to sink capital in the developing country which in turn, by creating natural interests that will push the foreign government to enforce the developing country’s commitments, ensures the success of the reform. As Either argues, once multilateral liberalization among developed countries has occurred and makes entry into the global economy attractive for developing countries, the commitment role of the PTAs can lead to a greater level of reform and investment world-wide than would otherwise occur; and the international scale economy externality then implies that the greater investment in and scale of the successfully reformed developing countries leads to gains for everyone.\(^{107}\)

Viewed together, these papers and the commitment theory branch of the trade agreements literature that they represent lend credence to the attractiveness of deep integration in trade agreements as a way to solve commitment problems. Three important issues remain. First, from the perspective of the commitment theory, are there good reasons to think that the required deep integration is best carried out in PTAs rather than in the WTO? Second, when it comes to tariff commitments, what does the commitment theory say about the attractiveness of preferential tariff cuts? And third, is there evidence that commitment motives are in fact important for understanding real-world trade agreements?

Regarding the first question, clearly deep integration is possible in the WTO, as the WTO treatment of domestic subsidies illustrates. At the same time, the failure of the WTO’s Doha Round to gain traction on the so-called Singapore Issues (or even, it now seems, on the much-less ambitious “trade facilitation” initiative)\(^{108}\) indicates that there are severe limits to how far deep integration is likely to proceed in the WTO. So to put the first question slightly differently: Does the commitment theory literature provide reasons to believe that the WTO cannot generate enough deep integration

\(^{107}\)We have confined our discussion here to the economic benefits that commitments via a trade agreement may provide, but there may also be important political benefits, as the recent paper by Liu and Ornelas (2014) suggests. Adopting the model of Ornelas (2005b) to a domestic commitment setting and introducing the possibility of coups which lead to the failure of democracy, Liu and Ornelas show that PTAs can serve as commitment devices for future governments that lower the probability of democracy failure, and they show as well that unstable democracies are more likely to join PTAs as a result.

\(^{108}\)Hoekman (2014) provides an assessment of the trade facilitation agreement.
and that PTAs should be called upon to shoulder the load? Here the literature does not provide a direct answer, but there are elements of the literature that are suggestive of a partial answer: the WTO may be ineffective at helping small countries make commitments, and the preferential nature of PTAs may allow PTAs to be more effective for small countries in serving this role. The reason is that, as we have discussed above in the context of the terms-of-trade theory, small countries are likely to face special challenges in the WTO in enforcing their rights, especially when these rights are not of a systemic kind (such as non-discrimination); but small countries may become “large” to foreign exporters who enjoy preferential access to their markets and the possible rents that such access can create, and those foreign exporters will then have an incentive to work to ensure that the PTA partner-countries follow through on their commitments. This suggests in turn that, at least when it comes to PTAs between large developed countries and small developing countries (small because of the small-country enforcement issues associated with the WTO, developing because commitment issues are arguably most germane for developing countries), the commitment theory could provide a reason to look more favorably on PTAs than does the terms-of-trade theory.

Regarding the second question raised above, we have just pointed out one possible reason that the commitment theory could lend support to the preferential tariff cuts that define PTAs, namely to aid in the enforcement of commitments for small countries. But beyond this, the commitment theory of trade agreements does not display any particular affinity to PTAs over the GATT/WTO. To the contrary, as can be seen with reference to the Maggi and Rodríguez-Clare (1998) workhorse model, a preferential tariff cut could be completely ineffective in reducing the ex-ante distortions that the domestic government is seeking to address with its commitment to (in the model, multilateral) free trade. More generally, the domestic distortions that the government is attempting to reduce with commitments under a trade agreement are related to local-market prices in the domestic economy, and any set of local-market prices that can be achieved via preferential tariff cuts can be achieved as well with appropriate MFN tariffs.

---

109 This reasoning is not without caveats, however. For example, the rents created for the foreign exporters of a large country by preferential access to a small PTA partner market may only exist in the short run, and be dissipated by supply responses as the foreign exporter prices return to their long run MFN world-price levels. Notice also that if Ethier’s (1998) position – that it is sunk foreign investment rather than preferential trade access per se that creates the natural interests to make sure the foreign government enforces the deep integration commitments of its PTA partner – is correct, then it is not obvious why the preferential nature of a PTA should matter for enforcement, and therefore not clear why a PTA would be more effective than the WTO in this capacity.

110 A second suggestive possibility comes from a particular interpretation of the findings of Aghion, Antrás and Helpman (2007). They show that global free trade may not be achievable via multilateral negotiations but could be achievable if FTAs are permitted, but this possibility arises only when governments are not maximizing welfare (so that global free trade would not be efficient when judged by the governments’ own objectives). This has the flavor of a possible commitment story that could support the creation of an institution that permits PTAs: ex ante, governments would like to maximize overall welfare, but ex post they know they will be captured by lobbies; hence setting up an institution to maximize their ex-ante objectives might involve permitting FTAs when their ex-post objectives will be politically tainted.

111 This would be the case as long as the domestic-country imports from non-PTA countries were not driven to zero in the presence of the preferential tariff cut (because then the unchanged MFN tariff of the domestic country together with its smallness in world markets and the arbitrage condition would ensure that the local-market prices in the domestic economy are unaffected by its tariff preferences and so the sectoral allocation of domestic productive resources would be unaffected as well).
Finally, we turn to the third question: Is there evidence that commitment motives are in fact important for understanding real-world trade agreements? Here the evidence is thin but suggestive. An early paper that offers empirical support for the commitment theory is Staiger and Tabellini (1999). They look for evidence that governments make different tariff choices across distinct GATT environments that are distinguished by the degree of commitment that GATT rules provide for those choices. Focusing on US tariff choices made under the GATT Escape Clause (where GATT should not have helped provide commitment) and made in the Tokyo Round of GATT negotiations (where GATT rules could have helped provide commitment), Staiger and Tabellini find that US tariff decisions were more responsive to the production distorting impacts of the tariffs in the latter decisions as compared to the former decisions, in line with what their commitment model would predict. Tang and Wei (2009) adopt a different approach to this question. They consider the differences in the growth performance of developing countries that joined the GATT/WTO under two different kinds of accession rules, rules that applied in GATT prior to the creation of the WTO and that did not require acceding countries to undertake extensive policy commitments, and rules that applied subsequent to the creation of the WTO in 1995 and that typically required substantial policy commitments on the part of the acceding government. Tang and Wei find that the post-accession growth performance of the developing countries that acceded to the GATT/WTO under the latter set of rules was significantly better than that of non-acceding countries and countries that acceded to GATT under the former rules even after a number of attempts to control for possible differences in reforms across countries, and they attribute this difference in growth performance to the domestic commitment role played by the WTO.112

Summarizing, we conclude from our survey of the commitment theory strand of the trade agreement literature that this theory provides some reason to be less negative about PTAs than our survey of the terms-of-trade theory alone would suggest, but it provides no reason to be less supportive of the WTO. We therefore see the commitment theory as nudging the answer to the motivating question of our survey toward a view that PTAs and the WTO may be complementary, but in light of the relatively thin empirical support to date for the commitment theory’s role in actual trade agreements, probably not as yet nudging the answer very far in this direction.

6.2 The Delocation/Profit-Shifting Theory

Another alternative to the terms-of-trade theory can be found in a growing body of literature emphasizing firm delocation and profit-shifting as sources of international externalities that might

112Further evidence in line with the predictions of the commitment theory of trade agreements is provided by Limão and Tovar (2011) (see note 106), who employing data on Turkish tariffs and NTBs find evidence in line with their theory that Turkey was more likely to bind its tariffs in the WTO and to bind them more tightly in industries where it had low bargaining power relative to the lobbies it faced; and by Liu and Ornelas (2014) (see note 107), who find evidence supporting the two key predictions of their model that PTAs are more likely to be formed by unstable democracies and participation in PTAs helps to stabilize these democracies. Finally, in their cross-country study of emerging economies, Bown and Crowley (2014) provide evidence that these countries changed how they conduct their trade policy (through antidumping and safeguards) by taking on tariff binding commitments when joining the WTO in 1995.
give rise to and shape the design of trade agreements. This strand of the literature shares with the terms-of-trade theory a focus on the internalization of international policy externalities as the primary task of trade agreements; but the delocation/profit-shifting theories reject the implication of the terms-of-trade theory that terms-of-trade externalities are the only – or even the most important – cause of the inefficiency that a trade agreement can correct. Instead, according to these theories, non-cooperative trade policy choices are inefficient because, when left on their own, governments use trade protection to inefficiently “delocate” firms or shift firm-profits from foreign locations to the domestic market.

The delocation/profit-shifting theories build on the unilateral incentives for trade policy intervention that arise when the assumption of perfect competition is relaxed, incentives that were first identified by Brander and Spencer (1985), Spencer and Brander (1983) and Venables (1985, 1987). Brander and Spencer argued that trade policy intervention could be used to shift firm profits toward the intervening country when firms possess market power and use their market power in the presence of entry barriers to earn profits in equilibrium. Venables showed that in a world where profits are dissipated by free entry it is still possible to gain unilaterally with trade policy intervention as long as there are international transport costs, due to the delocation/entry-exit effect of this intervention on foreign and domestic firms. Ossa (2011) was the first to explore the consequences of the delocation effect for the purpose and design of trade agreements, while Mrazova (2011) develops the profit shifting rationale for trade agreements.

Ossa (2011) considers a monopolistically competitive setting in which firms producing differentiated products compete for sales in both the home and foreign markets under conditions of free entry, and where exporting the product abroad involves shipping costs. Venables (1987) established that a firm-delocation motive for trade policy arises in such an environment: if the domestic country offers protection to its importers or a subsidy to its exporters, foreign firms can be “delocated” to the home market and domestic consumers save on trade costs and enjoy a lower overall price index as a result. This domestic gain, however, comes at the expense of foreign consumers, whose price index rises. Hence, the firm-delocation effect represents a negative international policy externality.

What Ossa demonstrates is that in his model the transmission of the firm-delocation effect from the domestic to the foreign country can be interpreted as traveling through local-market prices, not the terms of trade. Intuitively, in Ossa’s model each country is impacted directly by the local price in the other country’s market, because each country could enjoy the savings in transport costs if it could have more of the world’s firms (and the production of their individual varieties) located locally rather than abroad; and the equilibrium pattern of firm location across countries depends on local prices in both countries via the free-entry condition.

A main thrust of Ossa (2011) is therefore that one does not have to believe that terms-of-trade effects of trade policy are important in order to understand the purpose of trade agreements.

---

113 As we discussed in section 4.3, the problem for a trade agreement to solve in models featuring the delocation/entry-exit effect can be given a terms-of-trade interpretation under certain conditions. But the papers we next discuss impose conditions so that this interpretation is not valid, which is why we discuss these papers here rather than in the context of our earlier discussion of the terms-of-trade theory. We return to this point further below.
In fact, Ossa argues that a number of the prominent design features of the GATT/WTO (e.g., reciprocity and MFN) can be equally interpreted as sensible if governments are instead attempting to internalize delocation externalities with their trade agreements. Accordingly, and based on Ossa’s results, if anything the delocation theory of trade agreements could be said to strengthen support for the GATT/WTO as a well-designed institution, as it broadens the interpretation of the problem that the GATT/WTO is well-designed to solve.

Still, while the delocation theory does not appear to undercut support for the GATT/WTO in some fundamental way, might it at least elevate support for PTAs above the fairly dim view of PTAs that emerges from the terms-of-trade theory? For example, does the delocation theory swing support from shallow to deep integration, in which case like the commitment theory the delocation theory might then be interpreted as supporting the need for PTAs to complement the GATT/WTO? This is still an open question in the literature, but the results of DeRemer (2013a) suggest that the answer to this question may be a qualified “No.” In particular, DeRemer demonstrates that shallow integration can work in a delocation model where governments have both tariffs and wage subsidies at their disposal. But DeRemer also shows that the particular form of the “market access preservation rule” that makes shallow integration work in his model bears little relationship to any of the closest analogies in the GATT/WTO. Together with Ossa’s (2011) findings, DeRemer’s results therefore suggest that the GATT/WTO is probably less-well designed to solve problems associated with international delocation externalities than it is to solve the terms-of-trade problem, but with some selective fixes there is no reason in principal that it could not be optimized in this direction, and therefore no particular reason to believe that PTAs are needed to help shoulder the load.

Similarly, Mrazova (2011) develops a model of trade agreements based on the notion that such agreements help to internalize an international profit-shifting externality, and uses this model to explain the WTO ban on export subsidies that as we have observed is puzzling from the perspective of the terms-of-trade theory (similar to Potipiti’s, 2012, use of the commitment theory to explain this puzzle). Mrazova’s argument is based on self-enforcement considerations: she shows that it can be easier to enforce commitments in a repeated game model of trade agreements when import tariffs are the only trade policy instrument allowed under the agreement. DeRemer (2013b) abstracts from self-enforcement issues but employs a profit-shifting model of trade agreements to argue that the evolution of GATT/WTO domestic and export subsidy rules can be better understood from the perspective of profit-shifting models of trade agreements than from the perspective of the terms-of-trade theory (along the lines of Bron and Ruta’s, 2013, use of the commitment model to explain the evolution of domestic subsidy rules in the GATT/WTO). To generate his domestic subsidy result, DeRemer assumes that governments wield a particular form of entry subsidy, and it is an open question whether the result would hold for more general forms of domestic subsidy (or for domestic

---

114 This is not to say that the predictions of the terms-of-trade theory and the delocation theory are the same. For example, as Ossa (2011) demonstrates, the particular definition of reciprocity that is suggested by his model differs from that suggested by the terms-of-trade theory, and the implications of reciprocity and MFN together are somewhat different across the two modeling approaches as well (see especially Ossa, forthcoming, on this point).
policy instruments more generally). But at a minimum DeRemer’s result illustrates that profit-shifting externalities can be usefully employed to help interpret the evolution of GATT/WTO subsidy rules. And finally, Ossa (forthcoming) assesses the importance of profit-shifting and terms-of-trade externalities in a calibrated quantitative model of trade agreements, and finds that together these externalities are sizable enough to account broadly for the observed magnitude of multilateral tariff liberalization under the GATT/WTO.

Again, like the delocation theory, the results from the profit-shifting theory seem to provide, if anything, further support for the GATT/WTO approach to liberalization. And at least to date, the profit-shifting theory has not generated, and does not appear likely to generate, particular results that would indicate heightened support for PTAs. These conclusions are further supported from the perspective of the findings of Bagwell and Staiger (2009, 2012a). According to their findings, as long as governments have both import tariffs and export tax/subsidies at their disposal, the underlying problem for a trade agreement to solve in the delocation and profit-shifting models can still be given a terms-of-trade externality interpretation, though novel local-price externalities do arise when export policies are not available to governments (perhaps ruled out as part of the trade agreement as in Mrazova, 2011). This means that we should expect key results across these theories to be similar, at least when the use of export policies is not ruled out, and Bagwell and Staiger confirm this for a number of the standard predictions of the terms-of-trade theory.

We conclude from our survey of the delocation/profit-shifting theories of trade agreements that this strand of the literature, if anything, bolsters the case for the GATT/WTO approach to liberalization, as it provides a broader base from which to interpret as sensible many of the core design features of the GATT/WTO; and while more work is needed to tease out the implications of these theories for the desirability of PTAs as a form of liberalization, at present these theories provide no specific reasons to think that PTAs offer an attractive alternative or complement to the GATT/WTO approach. Combined with the fact that, with the notable exception of Ossa’s (forthcoming) calibration exercise, there is to date no evidence on the empirical importance of these theories, we do not view the delocation/profit-shifting strand of the trade agreement literature as providing further qualifications to the conclusions we have drawn so far concerning the relative merits of PTAs and the GATT/WTO.

115 It is for this reason that we interpret Deremer’s (2013b) domestic subsidy result somewhat more narrowly as applying to certain subsidies but probably not to domestic policies more generally, while we view the analogous results of Brou and Ruta (2013) concerning domestic subsidies as suggesting (from the perspective of the commitment theory) broader implications for deep integration.

116 This is not to say that the delocation and profit-shifting models don’t offer new insights. For one thing, to the extent that they are successful, the GATT/WTO attempts to limit export subsidies make analyses of these models without export policies an empirically relevant case to consider. For another, even when export policies are available and a standard terms-of-trade interpretation can be given to the problem, the deviation from perfect competition that is featured in the delocation and profit-shifting models can yield novel predictions (as, for example, in Bagwell and Staiger’s, 2012b, use of the linear Cournot delocation model discussed in an earlier section to offer a terms-of-trade theory interpretation of limits on export subsidies).
6.3 The Offshoring Theory

It is by now well-documented that trade in intermediate inputs dominates modern trade flows, that many of these inputs appear to be highly specialized to their intended use, and that this has not always been so (see, for example, Johnson and Noguera, 2014, as well as the discussion in Antràs and Staiger, 2012b). This rise in the importance of “offshoring” raises the question whether the rules and norms of the GATT/WTO, conceived at a time when the nature of trade was quite different, are still relevant today. There are two issues that have been addressed in the literature. First, the rise in offshoring has been accompanied by a significant rise in foreign investment. Adopting the perspective of the terms-of-trade theory of trade agreements, Blanchard (2007, 2010) argues that this investment might naturally reduce the magnitude of the terms-of-trade externality problem that according to the terms-of-trade theory a trade agreement is needed to solve, and in this way offshoring may act to reduce the need for a GATT/WTO-type institution. We have discussed Blanchard’s work above in the context of our survey of the terms-of-trade literature. In this section we focus on a second issue associated with the rise in offshoring: its potential to alter the mechanism of international price determination. This issue has been emphasized by Antràs and Staiger (2012a, 2012b), who argue that the rise of offshoring may have changed the way that international prices are determined and thereby changed the nature of the international externality that a trade agreement must address, with implications for the design of effective trade agreements.

Antràs and Staiger (2012a, 2012b) distinguish between international prices that are determined by standard market clearing conditions and prices that are determined by bilateral bargaining between foreign suppliers and domestic buyers. The former mechanism is featured in the terms-of-trade theory, and it underpins a property that is key for the terms-of-trade theory’s affinity with shallow integration: the tariff is the first-best policy for terms-of-trade manipulation and the international cost-shifting that is implied. This is a key property because, as we discussed earlier in our survey, according to the terms-of-trade theory the only “problem” for a trade agreement to “solve” is terms-of-trade manipulation, and as the tariff is the first-best policy for terms-of-trade manipulation it then follows that the tariff is in fact the only policy that is distorted in the Nash equilibrium: behind-the-border measures are set efficiently under Nash choices.

But when international prices are determined by bilateral bargains between buyers and sellers located in different countries and these bargains are undisciplined by market clearing conditions, as Antràs and Staiger (2012a, 2012b) argue is increasingly the case with the rise in offshoring, the tariff is no longer the first-best policy for cost-shifting, and governments will typically find it unilaterally optimal to distort many of their policy choices – border but also behind-the-border – in an effort to manipulate international prices and shift costs onto their trading partners. This leads to Antràs and Staiger’s first claim: through its implications for international price determination, the rise in offshoring is likely to erode the effectiveness of the GATT/WTO shallow integration approach. And as Antràs and Staiger demonstrate, a second claim follows when governments have political economy motives: in the presence of offshoring, and in stark contrast to the predictions of the terms-of-trade theory, these motives introduce additional policy inefficiencies that a trade
agreement can address. Taken together the implication of the Antràs and Staiger findings is that the rise of offshoring may usher in a new world in which a collection of individualized deep-integration agreements is needed to guide government to internationally efficient policy choices. Strikingly, that sounds a lot like the recent wave of PTAs.

Evidently, the offshoring theory of trade agreements has strong implications for the relative merits of PTAs and the GATT/WTO, both diminishing the attractiveness of the GATT/WTO and elevating the attractiveness of PTAs. And unlike the delocation/profit-shifting theory which largely reinforces the message of the terms-of-trade theory, and the commitment theory which as we have noted can comfortably be viewed as complementary to the terms-of-trade theory, the offshoring theory seems more fundamentally at odds with the terms-of-trade theory and the implied support for the GATT/WTO approach to liberalization. A key question for our survey is therefore whether there is empirical support for the offshoring theory. In fact we are unaware of any direct empirical evidence relating to this theory, and only a few pieces of indirect evidence, which are supportive and suggestive, but hardly definitive.\footnote{We are aware of two pieces of evidence that offer some indirect support for this theory. A first is presented in Antràs and Staiger (2012b), and relates to apparent difficulties in achieving negotiated reductions of tariffs on imported differentiated inputs for a set of countries acceding to the WTO. And a second is contained in Orefice and Rocha (2014), who find evidence that an important predictor of whether two countries sign a deep-integration PTA is the share of their bilateral trade that is comprised of parts and components.} Hence, while the offshoring theory of trade agreements has the potential to substantially alter our conclusions about the relative merits of PTAs and the GATT/WTO, in light of the lack of empirical evidence to date on the relevance of the theory it would be premature to place much weight on its implications at this time. Clearly, however, empirical research in this area is warranted.

6.4 Summing Up Thus Far

Having surveyed the three additional strands of the trade agreement literature, we are therefore left with the following view: while qualified along some important dimensions, the terms-of-trade theory’s implication, that strong support for the GATT/WTO is warranted while a cautious view of PTAs should be taken, survives largely intact. The commitment theory serves to elevate support for PTAs in certain circumstances beyond what the terms-of-trade theory would suggest, while the delocation/profit-shifting theory seems to reinforce the terms-of-trade theory’s support for the GATT/WTO. The offshoring theory could potentially overturn the implications of the terms-of-trade theory in this regard, but there is to date insufficient evidence to justify a change in position based on the predictions of this theory.

We next survey the literature on dispute settlement in trade agreements, and consider the relative merits of PTAs and the GATT/WTO as institutions for solving trade disputes.
7 Dispute Settlement

Such is the apparent success of WTO dispute settlement that it is often referred to as the “crown jewel” of the multilateral trading system. Over a relatively short period, economists, political scientists, and legal scholars have developed a range of positive and normative approaches to explore important research questions in this area.\textsuperscript{118} For example, legal scholars now parse the textual language of each newly-arriving WTO legal decision – so as to draw potential precedent implications for international and domestic law, as well as for public policy – with the same voracity as the American bar devours fresh Supreme Court rulings or Europeans tackle judgments from the European Court of Justice. Furthermore, legal and economic scholars have already teamed for more than a decade to annually publish joint inter-disciplinary assessments of each year’s new WTO jurisprudence.\textsuperscript{119}

This section begins by characterizing how members have used the WTO dispute settlement system to date. Our second section describes research that provides theoretical perspectives on the role of the WTO dispute settlement system. A third section examines the relationship between PTA implementation and WTO dispute settlement activity, while a fourth section discusses recent disputes in PTAs that are associated with the enforcement of non-trade policies. Throughout, we also identify promising directions for future research. A final section offers concluding thoughts on the relative merits of a multilateral dispute settlement system.

7.1 WTO Dispute Settlement: Patterns in Use

WTO disputes involve state-to-state level interaction. Here we characterize some of the data on the nearly 500 formal disputes initiated between 1995 and 2014 in order to first clarify what the disputes tend to be about, who they tend to be between, how the process works, and the typical outcomes that arise. Due to a number of potential issues related to sample selection, a central argument of the literature described below is that dispute settlement cannot be appreciated or evaluated based solely on an appeal to information regarding the disputes that are observed – i.e., the procedures also have important “off equilibrium” implications that do not necessarily arise through formal disputes. Nevertheless, evidence that the WTO members are putting trust in the system by frequently triggering its use over important policies or significant amounts of trade is potentially supportive of the idea that dispute settlement is playing a significant role.\textsuperscript{120}

\textsuperscript{118}Examples of seminal legal scholarship over GATT law begin with Dam (1970), Jackson (1969) and Hudec (1990). The WTO’s legal process is explained in more detail by Palmeter and Mavroidis (2004). Davis (2012) provides a recent overview of political science research on WTO dispute settlement.

\textsuperscript{119}Beginning in 2001, Horn and Mavroidis (2003) initiated a series of annual assessments with the backing of the American Law Institute, subsequently extended by Bown and Mavroidis (2013), which pairs legal scholars with economists to jointly analyze each of the formal WTO Appellate Body (and non-appealed Panel) decisions that arise. Cumulatively these legal-economic assessments have covered nearly 100 different individual dispute decisions to date.

\textsuperscript{120}Put differently, at this moment in time, the WTO members do not seem to be avoiding use of the system or clamoring to develop an alternative system. The GATT experience of the 1980s in particular suggests that that is not necessarily always the case (Bhagwati and Hudec, 1990).
7.1.1 WTO disputes are frequently, but not always, about allegations of excessive import protection

We begin by appealing to an assessment drawn from a database of WTO dispute settlement information compiled and made publicly available by Horn and Mavroidis (2011). Of the formal disputes that have arisen to date, the typical topic concerns a plaintiff ("complainant") member alleging that the defendant ("respondent") country has imposed an excessive level of import protection.

Horn, Johannesson, and Mavroidis (2011) examine disputes taking place between 1995 and 2010 and report that nearly 95 percent of all WTO disputes concern trade in goods, with the remainder split between WTO commitments over services trade or intellectual property rights protection. Within the disputes over goods, there is wide variation as to the WTO Agreements that complainants invoke. The most frequently invoked agreements include the 1947 GATT, followed by the 1995 Agreements on Antidumping, Subsidies and Countervailing Measures, Agriculture, Technical Barriers to Trade (TBT), WTO, Safeguards, and Sanitary and Phytosanitary (SPS) Measures. New import restrictions under antidumping, countervailing duties, and safeguards in particular are some of the most frequently challenged policies facing formal dispute.

China, which acceded to the WTO in 2001, is an interesting case study. China’s accession terms required it to take on significant commitments to domestic (economic) and trade policy reform. And because it is such a large importer, exporter, and has so many trading partners, there were expectations that China would become significantly involved in WTO litigation (Bown, 2010). Within a relatively short period of time, China has become an important WTO litigant; more than 40 percent of disputes initiated between 2007 and 2011, for example, featured China as either a respondent or complainant, with an almost 2 to 1 ratio of respondent to complainant.121

China’s WTO disputes have also turned out to be quite diverse. There are some disputes filed against China that fit the “typical” mold – i.e., they feature an allegation of excessive import protection in a politically sensitive sector such as autos, steel, or agriculture, and address a commonly-challenged policy such as antidumping. However, China’s recent experiences also include disputes over a number of relatively new issue areas. For example, trading partners have used the WTO to legally challenge China’s export restrictions over various raw materials and “rare earth” metals that are especially important in electronic goods. There have been challenges to China’s tax and subsidy policies and its allegedly lax protection of intellectual property rights. Finally, trading partners have initiated disputes over China’s import restrictions on foreign providers of services; examples include financial information services (e.g., Bloomberg, Dow Jones, Thomson-Reuters), electronic payment services (e.g., Visa, MasterCard, American Express), and audio-visual services (e.g., movie studios, media and publishers, software providers).

More generally across the WTO caseload, disputes that involve an allegation of excessive import protection rarely involve one country challenging another country’s increase of an applied MFN

121 Russia also became active in formal dispute settlement quite quickly after its August 2012 accession. At the time of writing, Russia had already filed its first two complaints, it had had four disputes brought against it as a respondent, and was a formal interested third party in 17 other disputes.
tariff above its legal binding. Instead, the complainant country’s typical allegation is that the respondent has implemented excessive import protection through an NTB policy of one variety or another. Then almost by definition, the dispute arises because of a disagreement in interpretation of whether the respondent’s policy is legally viable. The respondent and complainant typically disagree over whether the policy was implemented in order to achieve some other (non-trade) objective that is justifiable under the rules or exceptions of the WTO agreement. Nevertheless, the appearance – if not reality – of two countries having different interpretations of the terms of the WTO’s rules and exceptions does raise the possibility that dispute settlement may be doing more than simply enforcing the agreement. The process of rulings and renegotiations may also address the incompleteness of the GATT/WTO contract. Section 7.2.2 describes advances in this particular area of research.

An early dispute settlement literature was motivated by recognition that most disputes initiated immediately in the aftermath of implementation of the Uruguay Round commitments (e.g., 1995-1998) seemed to involve the trading interests of only high income countries. The policy concern was that the newly arising legal and institutional costs of engaging dispute settlement – given the increased “legalization” of the dispute settlement process under the WTO relative to its immediate GATT predecessor – might be too burdensome for developing country members with limited legal capacity to bear. That is, despite a new system that was based on law as opposed to diplomacy, the WTO system’s additional legal complexity may have inadvertently created a hurdle impeding the ability of poorer countries to enforce their market access interests.122

Beginning with Horn, Mavroidis and Nordstrom (2005), the evidence found strong correlations between dispute settlement use with a country’s overall levels of exports and the diversity of its trading partners. The more a country traded and the more bilateral trading relationships involving the country, the greater the scope for potential frictions to arise that would result in a formal dispute. Subsequent research on disputes from this early period that also explored the potential role of other factors - such as legal and retaliatory capacities - that might affect dispute initiation tended to confirm the central importance of the result that high levels of trade were positively correlated with the triggering of disputes.123

Nevertheless, Bown and Reynolds (forthcoming) provide a recent characterization of the bilateral trade in disputed products for a sample of disputes making up more than 70 percent of the WTO caseload between 1995 and 2011 and find evidence of vast heterogeneity in the levels of trade and scope of products involved. They report that roughly 14 percent of WTO disputes involve bilateral trade in disputed products of tiny amounts – e.g., less than $1 million per year. However,

122 Bown (2009) provides a more comprehensive and in depth treatment of these and related issues confronting developing country access to WTO dispute settlement.
123 Bown (2005a), for example, provides evidence linking higher import penetration ratios to US antidumping duties in the first stage, and higher levels of bilateral trade affected by those US duties subsequently positively associated with the trading partners decision of whether to formally challenge them through GATT/WTO dispute settlement in a second stage. In a separate study examining a cross-country sample of WTO disputes that concern policies imposed on a national treatment basis – in which the policy negatively affected all trading partners – Bown (2005b) finds that higher pre-policy levels of bilateral exports of the disputed products are positively associated with potential litigants' decision of whether to formally engage in the dispute settlement process.
they also find that 15 percent of disputes involve bilateral trade in products over more than $1 billion per year and that the complainant’s disputed product exports are, on average, larger than the disputed product exports of interested third parties and non-participants in the disputes. The heterogeneity in trade stakes across the dispute data raises important questions for research to address regarding potential determinants of dispute settlement use. As we describe below, new theoretical research may help advance our understanding of these forces.

Bown and Reynolds (forthcoming) also provide summary statistics that suggest that WTO member countries have trusted the dispute settlement system to assess policies that cover significant amounts of trade. Over 1995-2011, WTO dispute settlement investigations collectively scrutinized nearly $1 trillion in goods imports, which works out to an average rate of $55 billion per year, or roughly 0.5 percent of world imports in 2011. These are arguably economically significant amounts of trade for dispute settlement to address, without even yet considering the impact of the potential precedent arising through WTO jurisprudence, as well as the “off-equilibrium” impacts of dispute settlement.

7.1.2 The WTO legal process and outcomes

Once a complaining country initiates a dispute and engages in a mandatory period of consultations with the respondent, if that fails to resolve the issue the countries can undertake a legal process whereby their legal representatives make arguments and present evidence to a WTO dispute settlement Panel. Nearly 200 of nearly 500 WTO disputes initiated to date have already resulted in formal, first stage legal rulings through Panel Reports. Of the disputes receiving a Panel Report, more than half have subsequently been appealed and received rulings from the WTO’s standing Appellate Body.

Each WTO dispute can contain numerous “claims” made by the complainant against the respondent; the data also reveal substantial variation across disputes as to the total number of claims filed. While it is generally recognized that there is a “pro-trade bias” in WTO legal rulings – in the sense that the Panel or Appellate Body finds that the respondent has done something wrong in almost every dispute that receives a formal ruling – there are a number of selection issues associated with this interpretation which we discuss more formally below. Furthermore, at the level of legal rulings over particular claims made, Horn, Johannesson, and Mavroidis (2011) report that the complainant “wins” only 57 percent of the claims over which the Panel ultimately rules. Furthermore, there are frequently numerous claims in these disputes over which Panel declines to rule for reasons of judicial economy.

Horn, Johannesson, and Mavroidis (2011) also provide information on the average length of time that the conditional set of disputes take to work their way through the sequential steps of the WTO dispute settlement process. On average, disputing countries remain in consultations

---

124 These are conservative statistics because they do not include roughly 30 percent of the WTO caseload of disputes that are either tied to export policies, services policies, TRIPS, or policies that affect all imports and are not linked to specific product codes. The overall level of trade directly affected by WTO disputes is likely to be much higher.
for nearly 6 months, and then another 15 months elapse before the WTO’s first Panel ruling. It is therefore typically almost two years between dispute initiation and the first legal decision. While the cases that are appealed receive an Appellate Body report relatively quickly (3 months) thereafter, another 9-11 months typically elapse before expiration of the reasonable period of time necessary for the respondent to bring its disputed policy into compliance with rulings and before the potential for discussion of compensation due to noncompliance can occur. Overall, the process takes almost three years on average between the initiation of the dispute and the outcome whereby the respondent country is required to bring itself into compliance with legal rulings.

Finally, the default form of compensation in the WTO dispute settlement system in the event of noncompliance is tariff retaliation by the complainant subject to limits determined by WTO arbitrators; and authorized retaliation can only begin after the legal process described above is exhausted. Overall, and despite such a high volume of disputes that have taken place during the WTO period, retaliation has rarely been an equilibrium outcome. Fewer than 15 disputes have resulted in the WTO judges even having to articulate the permissible level for formal retaliation by the complainant in the event of noncompliance. Many fewer cases have resulted in the complainant country actually implementing the retaliation that the WTO may have authorized.125 Some high profile exceptions notwithstanding, there are very few examples of egregious non-compliance with rulings of the sort that result in the explicit WTO authorization of retaliation.126 Nevertheless, some of these exceptions have identified limitations of the current system and are thus the subject of areas of formal research that we describe below.

7.1.3 Dispute use and retaliatory capacity

The terms-of-trade approach to trade agreements suggests that trade agreements are designed to facilitate an escape from a terms-of-trade driven Prisoners’ Dilemma problem. This perspective suggests that cooperation in trade policies is achieved in the context of a repeated game, where observed deviations may lead to a severe off-equilibrium-path punishment such as Nash reversion that corresponds to a complete unraveling of the agreement. A different kind of retaliation is featured in the WTO dispute settlement system, as authorized retaliation in this context is arguably on the equilibrium path (i.e., something that is part of the agreement and occurs in practice) and

125Bown and Ruta (2010) describe the calculation of economic retaliation levels for the roughly 10 disputes that had made it all the way through the WTO dispute settlement process between 1995 and 2008 whereby arbitrators determined the level of permissible retaliation. The Bown and Pauwelyn (2010) volume provides a broader set of research contributions on the retaliation-setting experiences under the WTO from other perspectives in law, political science, and economics, as well as from policymakers.

126The EC – Beef Hormones dispute involved a sustained period of retaliation by the United States, as did EC-Banana Regime dispute between the EU, US and Latin American banana exporting countries before it was eventually resolved. US – Upland Cotton has not resulted in compliance but in a payoff (financial transfer) from US to Brazil. US – Internet Gambling, discussed by Irwin and Weiler (2008), has not resulted in compliance or implemented retaliation by the tiny island nations of Antigua and Barbuda against the United States. Finally, a separate case study, described in more detail by Bown and Prusa (2011), involves the repeated challenges by a number of different WTO members to the US use of “zeroing” in antidumping investigations. The US reformed its procedures in such a piecemeal manner after each subsequent WTO ruling that more than 15 different disputes over 10 years have been initiated over zeroing in the attempt to resolve the issue.
commensurate in nature.\textsuperscript{127} In addition, as discussed in Section 4.2, commensurate retaliation may occur along the equilibrium path in less formal ways, as for example when a privately informed government imposes an anti-dumping duty while recognizing that doing so increases the likelihood that its exporters will face a similar duty in the future. The terms-of-trade approach to trade agreements thus suggests that governments’ decisions to impose certain trade policies along the equilibrium path may be influenced by their perceptions of the retaliatory capacities of their affected trading partners. We next discuss empirical approaches providing evidence broadly supportive of this general perspective.

Bown (2002, 2004b) relies on dispute settlement data to explore the potential enforcement implications of the terms-of-trade theory. He provides theoretical and empirical approaches, respectively, to examine why a country would knowingly implement import protection through a policy that would result in a formal dispute when there are specific GATT/WTO provisions designed to accommodate such policy updating without a dispute. The theory exploits differences across the GATT/WTO-mandated limits to the compensatory retaliation response under different legal provisions and their implications in an environment characterized by trading partners with different capacity constraints on retaliation. The empirical application examines a cross-country sample of policies imposed between 1973 and 1994 and a framework in which governments are given the choice between GATT-consistent and inconsistent policies of import protection, the latter which would subsequently result in a formal dispute. Conditional on a country selecting to impose some additional import protection in the first stage, the evidence suggests that heightened retaliation capacity by affected trading partners makes the policy-imposing country more likely to implement policy changes in a way that conforms to the rules of trade agreements.\textsuperscript{128}

Blonigen and Bown (2003) present a related empirical analysis that focuses on US antidumping policy and provides further evidence that an enforcement system based on retaliation capacity is likely to endogenously affect the structure of policies that countries impose in the first place. That study examined US policies implemented during the GATT and early WTO period (1980-1998) and found that bilateral retaliation capacity influenced how new US import restrictions were implemented along two different dimensions. In particular, US industries were less likely to request antidumping import restrictions be imposed against trading partners for which they have industry-level export exposure to retaliation; this likely affects the potential policy actions that a government is ultimately asked to consider implementing. Second, conditional on receipt of a request for additional import protection, the US government was less likely to impose duties on trading partners for which overall US exports would subsequently be exposed to bilateral retaliation under a potential WTO dispute.

Beyond concerns over retaliation, there are other potentially important contributing explana-

\textsuperscript{127}For further discussion of the different interpretations of retaliation in GATT/WTO, see Bagwell and Staiger (2002, Chapter 6).

\textsuperscript{128}Bown (2004c) presents related evidence on retaliation capacity impacting the outcomes of trade disputes as well, which is consistent with the notion that such incentives are likely to affect policy choices ex ante. See also Bown (2004a) for evidence that the outcomes of such bilateral disputes are extended to third country exporters in a manner consistent with successful application of the MFN rule.
tions for the variation in how countries implement changes in trade policy, and the extent to which they may be concerned with whether such policy changes get caught up in formal WTO dispute settlement. One particular explanation is the contractual incompleteness of the WTO agreements. Below we discuss a literature that interprets WTO dispute settlement from this perspective.

7.2 WTO Dispute Settlement: Theoretical Perspectives

At the most basic level, the terms-of-trade approach to trade agreements suggests that the WTO represents a codification of a set of cooperative strategies for governments engaged in the repeated play of a terms-of-trade driven Prisoners’ Dilemma game. As noted, cooperation in such a setting is possible only if governments understand that off-equilibrium-path deviations may lead to a breakdown in the agreement and a corresponding reduction in cooperation. In this general context, what, then, is the role of the WTO dispute settlement system? This is an important and understudied question. In this section, we highlight two theoretical perspectives on the role of the WTO dispute settlement system: the system may enhance cooperation by increasing transparency (i.e., by generating and disseminating information) or by helping to “complete” the WTO contract. We also discuss research concerning possible reforms of the dispute settlement system.

7.2.1 Transparency and Information

As discussed in Section 4.2, cooperation is easier to achieve in Prisoners’ Dilemma settings when behavior is transparent or public. The decision of GATT contracting parties to concentrate protection mainly into tariffs (rather than quotas) can be understood in this light. Nevertheless, governments may face monitoring impediments and not be perfectly informed about the full range of trade-policy conduct of other governments. Consequently, WTO rules that generate and disseminate public information about trade-policy conduct may facilitate greater cooperation. The WTO Trade Policy Review Mechanism (TPRM), under which the WTO Secretariat conducts periodic reviews of the trade policies of member governments, may be evaluated in this context. In this section, we describe research in which the WTO dispute settlement system likewise may facilitate cooperation by generating and disseminating information.

As we discuss in Section 4.2, Maggi’s (1999) model of multilateral enforcement indicates that cooperation may be enhanced when third-party transparency is present. His general point is that greater cooperation can be achieved under a multilateral enforcement mechanism in which third-party countries stand ready to retaliate. If a trade agreement is designed to ensure that any deviation is observed by all member governments, then the resulting off-equilibrium-path punishment could take an immediate multilateral form and thus be more severe. In this context, we can imagine that a dispute settlement body that publicly identifies an off-equilibrium-path deviation could play a role in facilitating cooperation, by ensuring that all member governments are aware of the transgression. In particular, when even “bilateral” deviations are made observable at a multilateral level, governments may be able to achieve more cooperative tariffs along the equilibrium path as part of a self-enforcing agreement.
Park (2011) explores a related information role for the WTO dispute settlement system. In his two-country model, each government has available a publicly observed trade policy (e.g., a tariff) and also a “hidden” trade policy (e.g., a non-tariff barrier) that is imperfectly observed by the other government. Importantly, each government privately observes a signal of the other government’s hidden trade policy. In the absence of a dispute settlement system, the governments attempt to cooperate in this “private monitoring” setting. A key feature of optimal cooperation is that, after observing a suspicious signal, a government increases its publicly observed trade policy, which in turn publicly initiates a trade-war phase. The prospect of a trade-war phase discourages opportunistic behavior with respect to the hidden policy; however, suspicious signals are sometimes observed even when the hidden trade policy is set at cooperative levels, and so the trade-war episodes are a feature of optimal cooperation in the absence of a dispute settlement system. Relative to this interesting benchmark, a dispute settlement system can facilitate greater cooperation in his model by providing a public signal and enriching the forms of punishment that are feasible. In this general manner, a dispute settlement system may enhance cooperation by changing the information structure of the game.

Empirically, a potentially useful environment to ultimately examine implications of the WTO’s information-dissemination role may turn out to be the Great Recession of 2008-2009. For despite the highly synchronized and sudden global collapse in economic activity and trade flows, it is now well understood that a global surge in new trade protection of the scale of even earlier recessions, let alone the Great Depression of the 1930s, did not occur (Bown and Crowley, 2013b). Given the relatively moderate trade policy response, it is unsurprising that the WTO has also not been flooded with a subsequent onslaught of newly initiated disputes. However, one contributing explanation may be the sharp increase starting in 2009 of the multilateral monitoring efforts to improve information dissemination on trade policy changes, including by the WTO Secretariat (through its Trade Policy Review Body), the World Bank, and establishment of the independent Global Trade Alert.\footnote{Bown (2011b) provides a more complete discussion of these three enhanced initiatives for additional trade policy monitoring and surveillance that arose during the Great Recession, with emphasis on of the motivations underlying the World Bank’s temporary trade barriers information dissemination activities.}

Our discussion in this section has reviewed research arguing that the WTO dispute settlement system can facilitate greater cooperation by generating and disseminating information about trade-policy conduct. We turn next to a related but distinct set of research that emphasizes a role for WTO dispute settlement in “completing” the WTO contract.

7.2.2 Contract completion

Disputes sometimes arise as a result of disagreement about whether particular conditions are met under the agreement. Legal scholars especially stress the view that countries invoke dispute settlement to address instances in which the original terms of the agreement were vague or incomplete, perhaps because it was too costly to write all possible contingencies clearly into the original agreement. Dispute settlement may hence present a forum to address legitimate differences of opinion
about what behavior the contract was intended to induce. In this section we briefly describe research that adopts this perspective.\footnote{There is also a growing literature that explores the role of dispute settlement procedures in facilitating ex-post renegotiation of trade agreements where commitments may be viewed as either property rules or liability rules. For papers in the literature that adopt this perspective, see Maggi and Staiger (forthcoming) where there are no disputes in equilibrium and Beshkar (2010, 2013) and Maggi and Staiger (2013) where disputes arise in equilibrium. See also Lawrence (2003) and the legal discussion of Schwartz and Sykes (2002).}

Maggi and Staiger (2011) provide a first treatment that examines a variety of roles that dispute settlement might play in a trade agreement modeled as an incomplete contract. Trade takes place between two countries in partial equilibrium in a single industry. The importing country has the policy option of free trade or protection, and the exporting country has no policy instruments. The importing country makes its trade policy decision after a realization of state variables that affect welfare levels under free trade or protection. Realization of the state variables is observed by both countries and by the dispute settlement process – what we refer to here as the “court.” However, it is too costly to describe the states in an ex-ante contract, and the court does not observe the joint payoff to the two countries under the realization of the state variables. As a result, the two countries cannot write a complete state-contingent agreement and the contract must therefore be incomplete. Maggi and Staiger focus on three forms of contract incompleteness that can arise in this environment: the contract may leave gaps, it may be overly rigid, or it may use vague language which leaves the obligations under the contract ambiguous in some states of the world. The research explores a number of ways in which the agreement might articulate different roles for the court which vary in their degree of court “activism,” including limiting the court to enforcing clearly stated obligations in the contract, interpreting ambiguous obligations contained in the contract, filling gaps in the contract, and modifying clearly stated obligations in the contract. The model focuses on a court that is both costly to use and which makes errors in decisions. Importantly, the contract and the court mandate are chosen together as part of the optimal design of the institution. Finally, both a one-period model and a two-period extension are considered to explore implications of allowing the court to establish precedent.

Maggi and Staiger (2011) derive a number of results assessing different degrees of court activism. First, it is never optimal to allow the court to modify clearly stated obligations in the contract. Second, as the court’s decisions become more accurate, its mandate changes from non-involvement (beyond a pure enforcement role) to more activist roles, such as interpreting vague clauses and even filling in the contract where it is simply silent. Third, while precedent can improve efficiency by reducing expected future litigation costs on issues that would have been litigated in any event, it can also introduce negative effects through the costly increase of additional disputes arising over issue areas that would not otherwise have been litigated. Overall, the benefits of precedent outweigh the costs when governments are impatient and when the court is more likely to make mistakes.

Finally, this model can also be used to describe other important features of the disputes that arise, especially in relation to the characteristics of the court. First, there is a tendency of the court to exhibit a pro-trade bias in its legal rulings if the litigation costs to the complainant (exporter)
are high relative to the respondent. However, this pattern arises due to selection effects associated with the importing country being more likely under these conditions to actually have been at fault in the dispute in the first place. Second, and in parallel to the broader insight from the enforcement literature, the off-equilibrium influences of the court are what generate its beneficial impacts. And third, also in parallel with the enforcement literature, the frequency with which countries trigger disputes provides little information on the performance of the court or the value of the dispute settlement process to the overall agreement, a result that occurs in the model because of the interaction between the optimal choice of the contract and the optimal mandate of the court.

7.2.3 Dispute Settlement Reform Proposals

Despite the apparent success of WTO dispute settlement, there have been a number of proposals for reform of the enforcement system. Here we consider the literature that has evaluated some of these proposals. To help evaluate such proposals, it is also important to highlight the possible motives behind why such reforms may be desirable in the first place. One concern is the undesirable efficiency properties of tariff retaliation because, when implemented, the retaliatory tariffs introduce another round of deadweight loss and distortions. A second concern may be the “equity” properties of dispute settlement that arise for countries with bilateral trade imbalances that may not have a sufficiently credible enforcement threat to induce policy compliance.

A first proposal is to replace the current system of trade retaliation with a system of financial compensation; this could also be motivated by the realization that financial transfers have emerged in a couple of instances as part of voluntary settlements in actual WTO disputes.131 Limão and Saggi (2008)132 analyze such a proposal in a self-enforcing, repeated-game framework. They introduce “fines” (financial transfers) by allowing countries to exchange bonds when they sign the trade agreement. Their analysis reveals at least three key points. First, in order for fines to support low cooperative tariffs and provide compensation in the case of a policy deviation, the fines must be backed by an additional policy instrument that is not controlled by the deviating country. The natural instrument to consider is a tariff, and thus tariff retaliation continues to remain a central component to the system. Second, while a system of fines and a system of tariff (only) retaliation are equivalent if there are never policy deviations, Pareto improvements arise in the face of shocks that result in policy deviations and disputes. The intuition is simply that fines are a more efficient transfer instrument because tariffs impose larger losses on the affected country than there are gains to the imposing country. Third, and despite this potential Pareto improvement, a system of bonds and fines does not improve enforcement relative to retaliatory tariffs. The marginal benefit of deviating on bonds is exactly offset by the marginal cost to that country losing its own bond and, in

---

131 As compensation of the US – Section 110(5) of the US Copyright Act dispute, the United States agreed to pay European copyright holders 1.3 million euros annually. See Bronckers and van den Broek (2005), which also provides a more complete legal articulation of the proposal. As a settlement in the US – Upland Cotton dispute, the United States agreed to transfer $147.3 million annually as a form of technical assistance and capacity building to the government of Brazil (USTR, 2010).

132 See also Bagwell and Staiger (2005b). Extensions include Limão and Saggi (2013).
the face of a dispute, the violating country does not have an incentive to return the other country’s bond.

Mexico introduced a separate proposal for dispute enforcement reform based on the idea that countries could be allowed to trade their right to retaliate to other trading partners in lieu of implementing retaliation themselves (WTO, 2002). This option may be of special interest to smaller countries which perceive less benefit from retaliating on their own. Motivated by the Mexican proposal, Bagwell, Mavroidis and Staiger (2007) analyze tradeable retaliation rights in a trade agreement setting by considering different auction formats.\(^{133}\) They begin with a basic auction, in which two competing importers of the product on which retaliation can take place bid for the right to retaliate. As they demonstrate, the basic auction is an auction with (positive) externalities: the losing country prefers that the other bidding country win and impose a retaliatory tariff in comparison to the scenario in which no retaliation occurs. The reason is that the retaliatory tariff imposed by the winning country lowers the world price of the common import good and thereby generates a terms-of-trade gain for the losing country. Due to this positive externality, free-riding is a potential concern, and auction failures (where neither country bids) and misallocations of retaliation rights (due to pooling at the reserve price) can occur. They then consider an extended auction, in which the respondent country is allowed to bid and potentially retire the right of retaliation against it. Both positive and negative externalities can arise in this auction, though an interesting result is that the respondent country always wins the auction and the retaliation right is always retired without the inefficiencies of retaliation. The extended auction thus suggests a potential means through which monetary compensation might be extended from a large (respondent) country to a small (complainant) country.

Drawing normative inference from these different auctions requires further thinking about the underlying motivation of the reform to the enforcement mechanism – e.g., is it to ensure compensation, to encourage respondent compliance with rulings, or to enhance efficiency? The results suggest that the basic auction generates lower expected revenue for the complainant than the extended auction in which the respondent country is also allowed to participate. On the other hand, the compliance and efficiency criteria favor the basic auction under some circumstances.

7.3 WTO Dispute Settlement and PTAs

The stumbling-block/building-block literature reviewed in Section 5 indicates that PTA formation can impact the extent to which multilateral tariff cooperation can be achieved in a self-enforcing agreement. A related but distinct question concerns the relationship between PTA implementation and WTO dispute settlement activity. While there is little formal econometric work assessing implications of repeated-game models for dispute settlement activity,\(^{134}\) there are many examples of PTA implementation resulting in policy changes that have subsequently spilled over into PTA

---

\(^{133}\) See also Chen and Potipiti (2010), who derive the optimal auction design for this setting.

\(^{134}\) The closest work in spirit, but which does not involve trade disputes, is Prusa and Teh (2010), which provides evidence that PTA member countries tend to increase application of antidumping import restrictions against PTA non-members relative to PTA members in the period after implementation of the PTA.
non-members challenging those policy changes formally under the multilateral system’s dispute settlement procedures. These examples highlight this as a likely area of continued conflict at the intersection of preferential liberalization and multilateral commitments and thus an important area for additional research.

The GATT period was replete with disputes arising after countries took on new PTA commitments involving tariffs and trade-related policies that led to adjustments of the PTA members’ external trade policy commitments toward non-members. The European Economic Community in particular faced formal GATT disputes in 1973 after the accession of UK, Ireland and Denmark (initiated by Canada), in 1982 after the accession of Greece (initiated by the US), and in 1987 after the accession of Spain and Portugal (initiated by Argentina).  

Other and more recent examples of disputes involve new PTA members allegedly adjusting their non-tariff policies toward non-members, both immediately as well as long after implementation of the PTA, in the latter case due to PTA rules constraining the conduct of policy toward members in particular. The MERCOSUR reciprocal tariff cuts and customs union between Argentina, Brazil, Paraguay and Uruguay in the early 1990s provide two case studies in particular.

Argentina’s preferential tariff cuts in footwear under MERCOSUR led to an import surge from Brazil and served as the precipitating event behind one important WTO dispute. Argentina’s response to the preferential import surge was to subsequently impose a safeguard on footwear beginning in 1997. However, because of a MERCOSUR legal requirement that safeguards cannot be applied against other PTA members, Argentina exempted imports from Brazil from the policy. Not surprisingly, the policy failed to stem the import surge from Brazil - though it was effective against imports from non-members - and the result was that the EU and Indonesia filed a formal WTO dispute against Argentina. The WTO rulings in the dispute were some of the first WTO jurisprudence establishing additional conditions under which countries could viably apply safeguard measures.

A second example escalated from Brazil’s imposed ban on imports of retreaded tires from all sources in 2000; Brazil later claimed a public policy motive for the ban based on environmental and human health protection. However, Uruguay challenged Brazil’s ban under MERCOSUR’s dispute settlement provisions, and the subsequent legal ruling required that Brazil remove the ban’s application on retreaded tire imports from MERCOSUR partners. Brazil’s imports from some of the newly exempted MERCOSUR partners subsequently increased, and the EU - another exporter of retreaded tires still subject to the Brazilian import ban - filed a WTO dispute. This

135 Hudec (1993) provides an expanded discussion of Canada’s dispute over cereals exports (pp. 460-61), US’s dispute over fruit exports (pp. 496-498) and Argentina’s dispute over corn and sorghum exports (p. 550).

136 Bown, Karacaovali and Tovar (forthcoming) provide a discussion of the Argentina – Footwear dispute. Furthermore, this dispute is also noteworthy as its legal decisions resulted in the important WTO jurisprudence establishing the principle of “parallelism” whereby safeguard-imposing countries became required to only impose safeguard measures against trading partners whose imports they had actually included in the safeguard’s injury investigation. There are, however, also important counter-examples of countries actually facing WTO disputes brought by PTA partners after they have imposed a safeguard in a way that increased the level of import protection against PTA partners relative to non-partners. An example is the Dominican Republic - Safeguard Measures dispute; for an analysis see Bown and Wu (2014).
dispute also resulted in important WTO jurisprudence – this time for the overlap of trade and environmental policy.\footnote{Bown and Trachtman (2009) provide an analysis of the Brazil — Retreaded Tyres dispute.}

These are only two from the many examples of WTO disputes arising due to policy changes that countries make after implementation of a PTA, many of which are likely motivated by the policy changes imposing externalities on PTA non-members. With the proliferation of “WTO-extra” provisions arising under the new PTAs, this may be an increased area of importance for dispute settlement cases.

### 7.4 PTA Dispute Settlement and Deep Integration

While the GATT approach emphasizes shallow integration, many PTAs pursue deep integration linkages between domestic and trade policies. As we discuss in Section 4.4, these linkages may have implications for the most-cooperative tariffs that can be enforced; in particular, Limão’s (2007) work raises the possibility that optimal cooperation in a linked setting with non-pecuniary international spillovers could entail a reduction in trade-policy cooperation, if the linked agreement “borrows” enforcement power from trade policy to achieve non-trade objectives.\footnote{History provides many examples of countries “borrowing” enforcement through trade retaliation to achieve other objectives. The highest profile stem from countries implementing trade sanctions for allegedly political or national security purposes (Hufbauer et al, 2009) – recent examples include sanctions imposed on countries including North Korea, Iran, and Russia.} In this section, we focus on a related but distinct theme and consider PTA disputes that are associated with the enforcement of non-trade policies. We regard this area as an especially promising area for future research.

Our discussion is motivated by recent developments. For the first time, countries are now using trade policy to explicitly enforce the non-trade policy commitments that countries have undertaken bilaterally, outside of the WTO, in other agreements. Some of these examples stem directly from the “trade and...” provisions - such as environmental or labor standards - that were introduced into the public lexicon and came to prominence in the 1990s when NAFTA was first under consideration (Bhagwati and Hudec, 1996). We describe three recent examples here.

In 2010, the United States initiated its first ever formal dispute against a trading partner for the PTA partner’s failure to implement sufficiently high labor standards that it had committed to uphold as part of the PTA. As background, the CAFTA-DR that the US signed in 2004 included a number of “WTO-plus” provisions of the kind described by Horn, Mavroidis, and Sapir (2010) - including labor standards, environmental standards, and foreign direct investment provisions - that would subsequently be enforceable under the PTA’s own dispute settlement procedures. When Guatemala allegedly failed to enforce its own labor laws, the United States initiated a formal CAFTA-DR dispute (USTR, 2011). Such a dispute has the potential to escalate to tariff retaliation.

The second example stems from Mexico filing and winning a NAFTA dispute over the US market for commercial trucking services and implementing $2 billion of authorized retaliation for the US failure to remove restrictions out of public safety concerns. As background, when the US, Canada, and Mexico implemented the NAFTA in 1994, the US agreed to remove restrictions on commercial...
truck and cargo shipping services provided by Mexican firms. However, the US announced in 1995 that it would not remove the restrictions out of alleged safety concerns that the Mexican trucks posed for the US public. Mexico subsequently initiated a formal dispute under NAFTA and in 2001 a NAFTA Panel found against the US restrictions. After eight years of the United States failing to implement the recommendations of the Panel ruling, Mexico announced in 2009 that it would seek compensation by implementing NAFTA-authorized tariff retaliation over more than $2 billion of US imports. Ultimately this retaliation on US exporters galvanized sufficient political pressure within the United States to allow for policy reform in 2011 whereby Mexican trucks were allowed to service the US market and Mexico ended its period of retaliation (Department of Commerce, 2011).

The third example involves the United States implementing trade sanctions in 2012 against Argentina by removing the lower tariff preferences the US had previously offered under the Generalized System of Preference (GSP) program. The US does not have a PTA with Argentina, and the trade retaliation did not follow from a formal dispute, and it was not authorized by any particular dispute settlement process. The source of bilateral friction between the two countries is not even a trade matter; instead, it stems from a dispute over an alleged expropriation of US investments in Argentina, as such it is covered by a bilateral investment treaty between the two countries. This particular investment dispute was litigated under the International Centre for Settlement of Investment Disputes (ICSID), which ruled against Argentina and determined that Argentina should compensate US investors in the form of a financial transfer for damages. The implementation of US trade retaliation beginning in 2012 was an attempt to enforce Argentina’s foreign direct investment commitments and was due to Argentina’s failure to pay roughly $300 million that it owed US investors since 2005-6 (USTR, 2012).

The Guatemala (labor standards), US-Mexico (trucking services and safety standards), and Argentina (foreign direct investment) experiences are three recent examples highlighting the need for renewed research to evaluate the dispute settlement implications of existing, as well as newly proposed “deeper integration” trade agreements – e.g., the mega-regional agreements such as the Trans-Pacific Partnership (TPP) and the Transatlantic Trade and Investment Partnership (TTIP). The negotiations of such agreements are purported to focus less on tariff cuts and more on further cooperation of behind-the-border policies – not only those that might be characterized as non-tariff barriers (or disguised protectionism), but also those that have the efficiency-enhancing purpose of addressing legitimate externalities and domestic policymaking concerns. The need for additional research and understanding is particularly heightened if these three recent experiences also signal a tendency for agreements to move over time beyond soft law and toward hard law in which such deeper integration cooperation on non-trade policies would not only be subject to litigation under formal dispute settlement, but also subject to enforcement through tariff retaliation.
7.5 Summary

Our review highlights central features of the WTO dispute settlement system, the theoretical frameworks that attempt to identify a specific role for dispute settlement procedures, the complicated relationship between WTO dispute settlement and PTA implementation, and some aspects of PTA dispute settlement in the context of deeper-integration initiatives. In this concluding section, we briefly reflect on some of the relative merits of the WTO’s multilateral dispute settlement system.

While WTO dispute settlement is widely praised, it is challenging to determine a solid benchmark against which to measure its success. Certainly the willingness of member governments to utilize WTO dispute settlement procedures to examine disputed policies is suggestive that these procedures play an important role. Furthermore, it is also the case that for many bilateral trading relationships, the WTO’s multilateral system may be the only game in town. Indeed, one half to two thirds of world trade occurs between countries that are not in a common PTA and thus between countries without an obvious alternative forum under which to resolve their potential disputes.\(^{139}\)

One possible relevant benchmark is to compare the current WTO dispute settlement system to its most immediate predecessor, i.e., the multilateral dispute settlement system under the GATT in the late 1980s. The GATT system, which was relatively toothless and based more on diplomacy than law, spurred the rogue path of unilateralism that the United States undertook through its Section 301 actions during the period (Bhagwati and Hudec, 1990). US unilateralism was a signal of its strong displeasure with the old system and at least partially served to illustrate what a world without binding multilateral disciplines could look like. Ultimately these events contributed to ushering in the WTO’s new dispute settlement system in 1995, and the evolving system has arguably performed well in comparison to the GATT system.

A second possible benchmark is to compare the WTO dispute settlement system to the systems that are emerging in PTAs. In principle, PTA dispute settlement systems might be customized to better address the deep-integration initiatives that such agreements sometimes pursue. However, it is far from clear that dispute settlement in PTAs can handle the job on its own. First, there are numerous cases in which purely bilateral issues arising between partners in PTAs with relatively well-functioning dispute settlement systems (e.g., NAFTA) could not be resolved internally and ultimately spilled over into WTO dispute settlement anyway. Second, there are other examples (e.g., MERCOSUR), in which PTA dispute settlement decisions pushed PTA members to make policy choices that imposed externalities on non-members, thus leading those countries to initiate WTO disputes. While the empirical record of PTA dispute settlement use is scant, these sorts of anecdotal examples at least suggest that PTA dispute settlement may cause as many problems as they help resolve.

More research is required to better understand the tradeoffs, incentives, and forces also at work at the particular intersection of preferential and multilateral commitments. Such research may be

\(^{139}\)The WTO (2011, p. 64) reports that 65 percent (49 percent) of world trade in 2008 was extra-PTA trade excluding (including) the European Union. The trade included in the construction of these ratios does include some trade between non-WTO members, so not all of it would therefore be subject to WTO dispute settlement.
of special value now, since it is unlikely that a single overarching rule - such as, ‘wherever there might be a conflict in commitments or obligations, WTO law dominates PTA law’ - can be relied upon. Indeed, some of the major proponents of new obligations arising under PTAs - such as the US and EU - are also some of the most significant litigants and contributors to WTO jurisprudence. As a thought experiment, consider once again the US-Guatemala dispute over labor standards or the US-Argentina retaliation over investment provisions and compensation. When would it make sense for Guatemala or Argentina to use the WTO to challenge potential US trade sanctions as a violation of its WTO commitments, since there are no explicit WTO provisions authorizing the US to raise its tariffs for such reasons in the first place?

While the totality of these considerations tends to provide support for the WTO dispute settlement system, it is also clear that much more research is needed. The relative merits of multilateral and preferential dispute settlement systems is thus an important subject that warrants further theoretical and empirical analysis.

8 Conclusion

The world trading system seems to be at a crossroads. The most recent wave of globalization suggests that something fundamentally different has been taking place. Recent history has featured the rise of global supply chains and offshoring, coupled with the increased economic clout of a number of large emerging economies such as Brazil, Russia, India, and China. The emphasis of international cooperation has apparently changed as well, as there is a seeming momentum shift away from the multilateral and nondiscriminatory framework of the GATT/WTO in favor of discriminatory arenas under new PTAs, and away from negotiations emphasizing shallow integration and toward negotiations stressing increasingly deep integration. Understanding the implications of these shifts is important. What is on the line is a choice over which international institutions will set the future rules of globalization and shape the trade-offs we face in a globalized world economy.

The best path forward may depend on how we have arrived at this particular crossroads. There are two possibilities suggested by the trade agreements literature.

One possibility is that the rise of offshoring and the BRICS has somehow affected the kinds of rules needed to avoid the law of the jungle. In this scenario, the GATT/WTO approach of MFN tariff bargaining coupled with shallow integration as a way to enhance market access is itself in trouble, and its problems are not fixable. According to this possibility, countries must now negotiate deeper constraints in what traditionally was considered the realm of domestic policy making; and as a consequence, greater restrictions on national sovereignty are now an inevitable feature of globalization, with PTAs the most efficacious institution for carrying out the task.

The second possibility is that the rise of offshoring and the BRICS has created major new challenges for the WTO, but the kinds of rules needed to avoid a return to the law of the jungle are still fundamentally the same: the classic GATT/WTO approach to shallow integration may be in trouble, but its problems are fixable. According to this possibility, a trade-off between
sovereignty and globalization is largely avoidable, but only if the WTO is supported and its approach strengthened.

Our review of the literature favors the second possibility. The WTO is not passé. Subject to some caveats, it receives strong support from basic economic principles. The approach can be improved upon, but the effort to do so is worth it, and the stakes are high. This position is only enhanced when also taking into consideration the WTO’s relatively successful track record of resolving bilateral frictions through its system of dispute settlement.

It is likely that the task of shallow integration is not yet completed and there is important work left to be done, especially for some of the major emerging economies and less developed countries. However, creative adaptations to the GATT/WTO’s historically successful principles of reciprocity and nondiscrimination could seemingly be re-deployed - albeit in a more guided and targeted way - to address these challenges.

We close our review with a brief discussion of one potential approach to strengthening the WTO in response to these challenges that has received recent attention in the literature, namely, the possibility of pursuing critical mass agreements (CMAs) or plurilateral agreements (PAs). Such agreements can be thought of as a hybrid between the preferential trade agreements and what has otherwise been the WTO’s “Single Undertaking” approach. A CMA is defined as an agreement in a WTO-covered area between a subset of WTO member countries whereby, because WTO disciplines apply, the benefits the members offered to one another under the CMA must be extended to all other WTO members on an MFN basis. In order to prevent free riding, CMAs therefore may be only likely to occur between major subsets of large countries. The main example of a successfully concluded CMA occurring under the WTO is the 1997 Information Technology Agreement which cut tariffs to zero in products covered under the agreement, and for which the original negotiations were concluded between only 29 WTO members, though this has subsequently grown to more than 70. A PA, on the other hand, is an agreement between a subset of WTO member countries in an area where either WTO disciplines are not applied or where they are extended (e.g., ‘WTO-extra’), and to which therefore the benefits offered to one another would not need to be extended to other WTO members. Examples under the WTO include the Agreement on Government Procurement and the Agreement on Civil Aviation.

As Hoekman and Mavroidis (2013, 2014) point out, PAs and PTAs have important similarities and differences. Similarities include that both can be applied on a non-MFN basis without vio-

---

140 Plurilateral agreements were a common outcome of GATT rounds prior to the Uruguay Round. For example, the conclusion of the Tokyo Round in 1979 led to a number of plurilateral agreements adopted mainly by industrialized countries, in issue areas such as subsidies and countervailing measures, technical barriers to trade (standards), import licensing procedures, government procurement, customs valuation, antidumping, bovine meat, dairy, and trade in civil aircraft. The Kennedy Round of negotiations in the 1960s also brought forward a plurilateral code on antidumping.

141 Negotiations using CMA/PA approaches in new issue areas include attempts to liberalize trade in environmental goods and additional liberalization in services under a proposed TiSA (Trade in Services Agreement). Negotiations outside of the WTO among a small subset of mostly high-income countries have taken place for additional intellectual property rights protection and resulted in the anti-counterfeiting agreement (ACTA).

142 For additional discussions of plurilateral agreements, or what is sometimes referred to as potential “variable geometry” under the WTO, see also Lawrence (2006) and Levy (2006).
lating WTO rules and both seem to be going beyond the shallow integration approach of traditional GATT/WTO disciplines and into deep integration. On the other hand, important differences may make PAs more appealing than PTAs from the perspective of the multilateral system. First, PAs are “open” in that other WTO members should (in principle) be allowed an explicit path to accede to the PA in the future, whereas PTAs do not typically have an open accession process for potentially interested trading partners. Second, attempts to bring PAs into the WTO system would potentially make the content of these agreements much more transparent, which may be especially important for non-signatory countries. Third, problems arising between PA signatories would be addressed through litigation taking place under the WTO, thus more likely completing the contract in a coherent way, as opposed to the potential fragmentation of international jurisprudence that might take place otherwise arising under PTA dispute settlement provisions. Combined, Hoekman and Mavroidis suggest these features of PAs could make them less likely to impose externalities on third countries than PTAs. Nevertheless, the full theoretical implications of such alternative approaches have yet to be fully explored by the literature, and thus further analysis along these lines is a ripe area for additional research.
References


119


Kim, In Song. 2013. “Political cleavages within industry: firm level lobbying for trade liberalization.” Mimeogr., MIT.


Stockholm School of International Economics.


Table 1: GATT/WTO - 60 Years of Tariff Reductions

<table>
<thead>
<tr>
<th>Implementation period</th>
<th>Round covered</th>
<th>Weighted tariff reduction</th>
<th>Weights based on MFN imports (year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1948</td>
<td>Geneva (1947)</td>
<td>-26</td>
<td>1939</td>
</tr>
<tr>
<td>1950</td>
<td>Annecy (1949)</td>
<td>-3</td>
<td>1947</td>
</tr>
<tr>
<td>1952</td>
<td>Torquay (1950-51)</td>
<td>-4</td>
<td>1949</td>
</tr>
<tr>
<td>1956-58</td>
<td>Geneva (1955-56)</td>
<td>-3</td>
<td>1954</td>
</tr>
<tr>
<td>1962-64</td>
<td>Dillon Round (1961-62)</td>
<td>-4</td>
<td>1960</td>
</tr>
<tr>
<td>1968-72</td>
<td>Kennedy Round (1964-67)</td>
<td>-38</td>
<td>1964</td>
</tr>
<tr>
<td>1980-87</td>
<td>Tokyo Round (1973-79)</td>
<td>-33</td>
<td>1977 (or 1976)</td>
</tr>
</tbody>
</table>

Source: WTO (2007), Table 5. Notes: MFN tariff reduction of industrial countries for industrial products, excluding petroleum. Tariff reductions for the first five rounds refer to the United States only. The calculation of average rates of reductions are weighted by MFN import values.

Table 2: Applied tariff rates of selected GATT/WTO Members, 1952 and 2005

<table>
<thead>
<tr>
<th>Economy</th>
<th>1952</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>17</td>
<td>4.2</td>
</tr>
<tr>
<td>Benelux</td>
<td>9</td>
<td>4.2</td>
</tr>
<tr>
<td>Denmark</td>
<td>5</td>
<td>4.2</td>
</tr>
<tr>
<td>France</td>
<td>19</td>
<td>4.2</td>
</tr>
<tr>
<td>Germany</td>
<td>16</td>
<td>4.2</td>
</tr>
<tr>
<td>Italy</td>
<td>24</td>
<td>4.2</td>
</tr>
<tr>
<td>Sweden</td>
<td>6</td>
<td>4.2</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>17</td>
<td>4.2</td>
</tr>
<tr>
<td>European Union (EU-25)</td>
<td>-</td>
<td>4.2</td>
</tr>
<tr>
<td>Canada</td>
<td>11</td>
<td>3.8</td>
</tr>
<tr>
<td>United States</td>
<td>16</td>
<td>3.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WTO member economy</th>
<th>MFN applied rate, simple average</th>
<th>Binding rate, simple average</th>
<th>Binding coverage</th>
<th>Coverage of applied duties &gt; 15 percent</th>
<th>Maximum applied rate</th>
<th>MFN applied rate, agriculture only</th>
<th>TTB coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>G20 High-income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>2.7</td>
<td>10.0</td>
<td>97.1</td>
<td>0.1</td>
<td>28.0</td>
<td>1.2</td>
<td>1.0</td>
</tr>
<tr>
<td>Canada</td>
<td>4.3</td>
<td>6.9</td>
<td>99.7</td>
<td>6.9</td>
<td>551.0</td>
<td>16.2</td>
<td>1.2</td>
</tr>
<tr>
<td>European Union</td>
<td>5.5</td>
<td>5.2</td>
<td>100.0</td>
<td>5.1</td>
<td>605.0</td>
<td>13.2</td>
<td>3.0</td>
</tr>
<tr>
<td>Japan</td>
<td>4.6</td>
<td>5.2</td>
<td>99.7</td>
<td>3.8</td>
<td>692.0</td>
<td>16.6</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>5.1</td>
<td>11.3</td>
<td>100.0</td>
<td>0.4</td>
<td>427.0</td>
<td>6.2</td>
<td>NA</td>
</tr>
<tr>
<td>South Korea</td>
<td>13.3</td>
<td>16.6</td>
<td>94.6</td>
<td>10.4</td>
<td>887.0</td>
<td>52.7</td>
<td>0.5</td>
</tr>
<tr>
<td>United States</td>
<td>3.4</td>
<td>3.5</td>
<td>100.0</td>
<td>2.7</td>
<td>350.0</td>
<td>4.7</td>
<td>6.8</td>
</tr>
<tr>
<td><strong>G20 Emerging</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>12.5</td>
<td>31.9</td>
<td>100.0</td>
<td>36.0</td>
<td>35.0</td>
<td>10.5</td>
<td>3.2</td>
</tr>
<tr>
<td>Brazil</td>
<td>13.5</td>
<td>31.4</td>
<td>100.0</td>
<td>36.2</td>
<td>55.0</td>
<td>10.1</td>
<td>1.9</td>
</tr>
<tr>
<td>China (2011)</td>
<td>9.6</td>
<td>10.0</td>
<td>100.0</td>
<td>14.6</td>
<td>65.0</td>
<td>15.6</td>
<td>1.3</td>
</tr>
<tr>
<td>India</td>
<td>13.7</td>
<td>48.6</td>
<td>73.8</td>
<td>19.6</td>
<td>150.0</td>
<td>33.5</td>
<td>6.5</td>
</tr>
<tr>
<td>Indonesia</td>
<td>7.0</td>
<td>37.1</td>
<td>96.6</td>
<td>1.6</td>
<td>150.0</td>
<td>7.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Mexico</td>
<td>7.8</td>
<td>36.1</td>
<td>100.0</td>
<td>13.8</td>
<td>254.0</td>
<td>21.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Russia</td>
<td>10.0</td>
<td>7.8</td>
<td>100.0</td>
<td>11.7</td>
<td>292.0</td>
<td>13.3</td>
<td>NA</td>
</tr>
<tr>
<td>South Africa</td>
<td>7.6</td>
<td>19.0</td>
<td>96.4</td>
<td>20.6</td>
<td>&gt;1000</td>
<td>8.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Turkey</td>
<td>9.6</td>
<td>28.6</td>
<td>50.3</td>
<td>10.8</td>
<td>225.0</td>
<td>41.2</td>
<td>4.9</td>
</tr>
<tr>
<td>*<em>Developing, other</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangladesh (2011)</td>
<td>14.4</td>
<td>169.2</td>
<td>15.5</td>
<td>40.1</td>
<td>25.0</td>
<td>17.2</td>
<td>**</td>
</tr>
<tr>
<td>Burma</td>
<td>5.6</td>
<td>83.4</td>
<td>17.6</td>
<td>5.0</td>
<td>40.0</td>
<td>8.6</td>
<td>**</td>
</tr>
<tr>
<td>DR of the Congo</td>
<td>NA</td>
<td>96.2</td>
<td>100.0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>**</td>
</tr>
<tr>
<td>Egypt</td>
<td>16.8</td>
<td>36.7</td>
<td>99.4</td>
<td>19.2</td>
<td>&gt;1000</td>
<td>66.7</td>
<td>NA</td>
</tr>
<tr>
<td>Ethiopia† (observer only)</td>
<td>17.3</td>
<td>** **</td>
<td>50.8</td>
<td>35.0</td>
<td>22.4</td>
<td>** **</td>
<td></td>
</tr>
<tr>
<td>Iran† (observer only, 2011)</td>
<td>26.6</td>
<td>** **</td>
<td>45.7</td>
<td>400.0</td>
<td>30.4</td>
<td>** **</td>
<td></td>
</tr>
<tr>
<td>Nigeria (2011)</td>
<td>11.7</td>
<td>119.1</td>
<td>19.1</td>
<td>39.0</td>
<td>35.0</td>
<td>15.5</td>
<td>**</td>
</tr>
<tr>
<td>Pakistan</td>
<td>13.5</td>
<td>59.9</td>
<td>98.7</td>
<td>36.1</td>
<td>100.0</td>
<td>15.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Philippines</td>
<td>6.2</td>
<td>25.7</td>
<td>67.0</td>
<td>3.1</td>
<td>65.0</td>
<td>9.8</td>
<td>0.1</td>
</tr>
<tr>
<td>Thailand</td>
<td>9.8</td>
<td>27.8</td>
<td>75.0</td>
<td>22.6</td>
<td>142.0</td>
<td>21.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Vietnam</td>
<td>9.5</td>
<td>11.4</td>
<td>100.0</td>
<td>24.7</td>
<td>135.0</td>
<td>16.1</td>
<td>NA</td>
</tr>
</tbody>
</table>

Source: tariff data taken from WTO, ITC, and UNCTAD (2013) and temporary trade barrier (TTB) data taken from Bown (2014). Notes: parentheses indicate data availability for year other than 2012. *selected other developing countries chosen as those with 2012 populations greater than 50 million. ** indicates non-user (or un-reported user) of the policy instrument. NA = not available. G20=Group of 20. † indicates WTO non-member.