2 The GATT/WTO

In this chapter I provide a broad overview of the General Agreement on Tariffs and Trade (GATT) as well as its successor the World Trade Organization (WTO), and of the basic modeling framework that will provide my foundation for an economic interpretation of GATT's design features, its successes, and ultimately its shortcomings. I begin by describing GATT's design and a brief history of how it came to be, and I then present the modeling framework.

2.1 The Design of the GATT/WTO

Origins

The direct historical antecedents of what would eventually serve as the de facto constitution of the world trading system of the twentieth century arose at a time of crisis.[1] Trade barriers had become increasingly restrictive in the decade following World War I and reached a climax when the United States enacted the Smoot-Hawley Tariff Act of 1930, increasing average US tariffs from 38 to 52 percent. US trading partners responded, and soon tariff rates among all the major powers were generally on the order of 50 percent. As Hudec (1990, 5) explains, "The postwar design for international trade policy was animated by a single-minded concern to avoid repeating the disastrous errors of the 1920's and 1930's."

In 1934, the US Congress passed the Reciprocal Trade Agreements Act (RTAA). Under the RTAA, the United States for the first time engaged in bilateral reciprocal tariff bargaining with a sequence of trading partners, and it combined this bilateral bargaining approach

1. The material in this section builds from chapter 3 of Bagwell and Staiger (2002) and section 4 of Bagwell and Staiger (2010a).

with unconditional most-favored-nation (MFN) treatment, according to which exports from each country with whom the United States had an agreement under the RTAA would automatically receive the lowest ("most favored nation") tariff rate that the United States offered to any exporting country. It is widely acknowledged that much of the GATT architecture was inspired by prior US experience with the RTAA.

What is less well appreciated is the way in which the RTAA was itself influenced by the successes and failures of the many international attempts that came before it to address the problem of high and rising trade barriers. During the decade following World War I, the United States took part in a number of multilateral bargaining efforts to address this issue, each largely unsuccessful. In describing the evolution away from multilateral bargaining and toward the bilateral bargaining approach that would eventually be embodied in the RTAA, Tasca (1938, 7) attributes the lack of success of these earlier multilateral attempts to the complexity of multicountry bargaining:

> The adoption of a policy of bilateral actions does not preclude the use of multilateral conventions to liquidate trade barriers. During the post-war period various attempts to proceed upon this basis have met with little success. It is the method itself which possesses weaknesses in certain respects. . . . The complexities involved in such a program of concerted action arise in part out of the fundamental variations in national tariff systems. This means that practically only horizontal reductions in tariffs can be considered feasible. But the differences in the economies concerned and their varying positions in the world economy demand reductions in trade barriers according to the circumstances in each case. Moreover, the diffusion of responsibility grows with the number of prospective contractants. Nations became less concerned with the failure of a projected plurilateral pact and more with the possibility of yielding more in the way of concessions than other nations.

As Tasca observes, these repeated failures of multilateral bargaining led to a conscious decision on the part of the United States to experiment with bilateral bargaining under the RTAA.

The RTAA was remarkable not only because it adopted a bilateral bargaining approach to the problem, but because it marked the first time that the United States combined bilateral tariff bargaining with unconditional MFN.[2] Yet while the approach embodied in the RTAA was novel from the US perspective, from the perspective of Europeans it was not. As Tasca (1938, 135) observes, for decades before, the

2. The United States had, since 1922, adopted an unconditional MFN approach, but it maintained an "autonomous" (i.e., unilateral) tariff up until the RTAA (see Tasca 1938, 116–121).

approach of combining bilateral tariff bargaining with MFN treatment "formed the essential basis of the commercial policies of numerous European countries." In fact, it appears that the design and implementation of the RTAA built on lessons learned from the European experience in at least two important ways.

First, the European experience with bilateral tariff bargaining established the practical necessity of granting unconditional MFN. As Wallace (1933, 629) writes

> After the World War, France experimented with the idea of abandoning the most-favored-nation clause ... By 1927 France was again driven back to the granting of most-favored-nation treatment, either de jure or de facto. The reason is not far to seek. When a country, by exclusive tariff bargains, institutes discriminations against third countries, then the greater these discriminations the greater will be the pressure against that country for their removal. In each successive negotiation it finds that the firmest demand of the other country is for equality of treatment, present and future, guarded by a most-favored-nation clause or its equivalent.

In effect, the European experience with bilateral tariff bargaining taught the important lesson that a country's current bargaining partners would require the assurance that any future bilateral deals that it struck with other countries would not substantially erode the value of the concessions being granted, and that the most practical way to provide assurance against such "concession erosion" was with a promise of unconditional MFN. The promise of unconditional MFN was included in the RTAA in part to address the concession erosion issue.[3]

Second, the European experience provided an object lesson in the power of the perverse incentive to raise tariffs and adopt so-called bargaining tariffs to better position oneself for future negotiations. According to Wallace (1938, 630):

> This padding of tariff rates in anticipation of negotiations is a chief reason why half a century of bargaining has meant on the whole higher and higher tariff rates in Europe instead of lower and lower rates.

This also informed the design of the RTAA. As Tasca (1938, 179) observes:

> The United States Tariff Commission in submitting recommendations on tariff bargaining declared, "The Congress should formulate restrictions designed to

3. Other arguments articulated at the time for adopting a policy of unconditional MFN included the perceived "multilateralization" benefits that this inclusion was expected to engender and a reduction in the risk of war (see, for example, Culbert 1987 and Rhodes 1993).

prevent the inclusion in reciprocity agreements of illusory concessions; that is, the removal of trade barriers or the reduction of tariff rates when such barriers and rates had been raised in anticipation of tariff bargaining, the amount of the concessions being smaller than or not greater than the previous increases in barriers and rates. Specifically, it is suggested that the Congress prescribe that all concessions included in the reciprocity agreements, on both sides, be made from the rates and relating to the barriers in effect at a date which shall be fixed by the Congress."

The lessons learned from the European experience with bilateral tariff bargaining may therefore have contributed to the success of bilateral tariff bargaining under the RTAA by helping the United States avoid the twin problems associated with concession erosion and bargaining tariffs that plagued the European efforts before it. But, as it happened, the adoption of unconditional MFN would itself introduce a different potential issue for the RTAA, one that was related to the earlier problem of bargaining tariffs that the Europeans had experienced but took a slightly different form: While in the European experience this issue had taken the form of the unilateral positioning of pre-negotiation tariffs, under the RTAA the analogous issue became how to design bilateral agreements with early negotiating partners to best preserve bargaining power for later agreements with other negotiating partners. This task was made difficult by the unconditional MFN requirement, which automatically granted "for free" to other potential bargaining partners any tariff concessions granted to early negotiating partners.

The preservation of bargaining power for later negotiations became a major preoccupation of the United States under the RTAA. Describing the tactics used by the United States in this regard, Tasca (1938, 146–147) notes:

There are, then, five methods being utilized by the United States to assure the compatibility of the unconditional most-favored-nation clause with a conventional tariff bargaining program. By far the most basic is the chief supplier formula. This is reenforced by the reclassification of commodities in the tariff schedules of the Act of 1930. The use of partial reductions in successive agreements, the simultaneous negotiations with groups of countries and the withdrawal clause are subsidiary to the first two. They play the part of supporting beams in those instances in which the chief supplier is not entirely applicable to existing conditions.

In effect, by granting tariff concessions to a negotiating partner only on those products for which the partner was the principal ("chief") supplier, possibly combined with product reclassification for tariff purposes to heighten the dominance of the partner in these products, it

was thought that much of the free-rider potential created by unconditional MFN could be eliminated. And where free-riding remained a substantial possibility, three additional tactics were available: splitting the concession into a sequence of partial tariff reductions negotiated with different countries in successive agreements; attempting to engage groups of countries in simultaneous negotiations; and threatening to withdraw or modify the earlier agreement if free-riding continued.

Beckett (1941) reviews the US experience under the RTAA and emphasizes the difficulties involved in preserving bargaining power in the presence of unconditional MFN, even when the chief supplier rule is applied. As she describes, split concessions often became the preferred method to prevent undue loss of bargaining power in an early negotiation:

> It is important to notice that the use of the chief supplier rule involves certain special difficulties. A problem arises, for instance, when, during the process of negotiation with small countries, it is impossible to isolate any commodities in which the other country is our chief supplier...A further difficulty appears when two or more countries supply almost exactly the same quantity of a given commodity or when two countries are the chief suppliers of the commodity in alternative years. If a substantial reduction in duty is granted in the trade agreement with one country, bargaining power with the other country is lost. To avoid such embarrassment, simultaneous negotiations of two agreements can be attempted. More often a split concession is granted: that is, a small reduction in duty is made in the agreement with the first country and an additional reduction in the agreement with the second country. By this procedure bargaining power with the second country is preserved. (Beckett 1941, 23)

Tasca (1938, 146) also emphasizes the importance in this regard of the various withdrawal clauses that were included in the RTAA:

> If the major benefits of a duty concession fall to a third country and "in consequence thereof an unduly large increase in importation" takes place, the contractants may withdraw the concession or impose a quantitative restriction upon imports of that item. Concessions are granted by the United States only after careful study in order to gauge the effects upon the whole economy; if these calculations should fail, then there exists a remedy in resort to this clause. But what is more significant, this withdrawal clause forestalls any third country from reaping any considerable benefit from a concession which might in any manner lessen its incentive to promulgate a pact with the United States.

The practice of granting split concessions became the most frequently observed manifestation of bargaining tariffs under the RTAA, while the threat to withdraw or modify a concession was typically kept in the

background but seen as providing an important means of maintaining bargaining leverage for later negotiations.

In short, tariff bargaining under the RTAA exhibited a number of central features. The approach was decidedly bilateral, chosen only after the United States had considered, attempted, and ultimately rejected multilateral tariff bargaining. Prior European experience with concession erosion and bargaining tariffs influenced the design and implementation of the RTAA along important dimensions. And unconditional MFN, the chief supplier rule, split concessions, and withdrawal/modification clauses were understood to be central to the operation of reciprocal tariff bargaining under the RTAA.

Between 1934 and 1947, the United States successfully concluded separate bilateral agreements with 29 countries. Encouraged by its success in the bilateral arena with the RTAA, the United States sought to build on the key components and establish a multilateral institution. In 1946, negotiations began for the creation of an International Trade Organization (ITO). As with the RTAA, under the ITO it was expected that negotiations between governments would result in reciprocal and mutually advantageous reductions in tariffs, and the principle of nondiscrimination would then ensure that the reduced tariffs would be extended to all member countries. In 1947, GATT was negotiated and was intended to serve as an interim agreement, but the ITO was never ratified by the US Congress.

Stated Purpose

What is the stated purpose of GATT? According to its preamble, the objectives of the contracting parties include "raising standards of living, ensuring full employment and a large and steadily growing volume of real income and effective demand, developing the full use of the resources of the world and expanding the production and exchange of goods."

As for the means to achieve this purpose, the preamble of GATT states that "reciprocal and mutually advantageous arrangements directed to the substantial reduction in tariffs and other barriers to trade and to the elimination of discriminatory treatment in international commerce" would contribute toward these goals. The objectives stated in the preamble to the WTO are broadened to include the exchange of not only goods but also of services, and to acknowledge the additional objectives of sustainable development, the protection and preservation of the environment, and the greater inclusion of developing countries to share in the gains from the growth of trade. But the means to achieve

this purpose as stated in the WTO preamble are identical to those in the GATT preamble (with the phrase "international trade relations" in the WTO preamble replacing "international commerce" from the GATT preamble).

Perhaps surprisingly to economists, free trade is not a stated objective of GATT or the WTO. This reflects the fact that, as a "member-driven organization" that serves as a trade policy negotiating forum for member governments with diverse interests, priorities, and needs, the GATT/WTO is designed with the aim of securing mutually beneficial agreements among these governments, and free trade is not necessarily something to which all member governments will aspire.

In total, there were eight rounds of GATT negotiations that together spanned almost 50 years. The primary focus of the earlier rounds was the reduction of import tariffs on goods. In the final GATT round, known as the Uruguay Round, governments took on several new issue areas (e.g., investment, services, and intellectual property) and formed the WTO. The WTO has sponsored a ninth round, the Doha Round, launched in 2001 and as yet uncompleted. The WTO embraces the rules and agreements made in GATT negotiations, but it is also a full-fledged international organization with an explicit organizational charter and a unified dispute-settlement system. In effect, with the creation of the WTO, participating governments fulfilled their original quest with the ITO for an official international organization that would set and administer the rules of the world trading system.

Architecture

GATT/WTO member governments are obliged to abide by a set of rules. In GATT, these rules were laid out in a series of 39 articles. The WTO has incorporated these GATT articles and extended the principles embodied in them to a number of new issue areas. I now provide an overview of the GATT/WTO legal structure by focusing on the principles embodied in these articles.

It is helpful to distinguish between three broad elements: substantive obligations, exceptions to those obligations, and dispute settlement procedures. The substantive obligations of a GATT/WTO member relate to tariff commitments, MFN treatment, and a general "code of conduct" in the international-trade arena. Broadly speaking, these provisions oblige the member governments to concentrate national protective measures into the form of tariffs, to apply them on a nondiscriminatory basis to other members, and to honor any tariff "bindings" made in a GATT/WTO negotiation, where the tariff binding refers to

a legal maximum level above which a country agrees not to raise its tariff.

As mentioned, the GATT/WTO also provides for certain exceptions to these obligations. One class of exceptions is for "original" actions, such as when a member seeks to suspend an obligation temporarily, or to permanently withdraw a previous concession through renegotiation. The rationale for including exceptions of this nature is that a government is more likely to make a substantial tariff commitment if it knows that the legal system has "safeguards" allowing its concessions to be modified or withdrawn under appropriate conditions. Of course, a tariff commitment would lose its meaning if exceptions for original actions were not subject to some disciplining structure. In part for this reason, and in part to maintain a balance between the rights and obligations of the members, GATT/WTO rules permit as well a second class of exceptions for "retaliatory" actions. Specifically, if a government modifies or withdraws a previous concession, then GATT/WTO rules recognize that a cost may be borne by a trading partner. This partner may then seek "compensation" from the government for the harm done (e.g., a tariff reduction from the government on some other good), and if this fails the partner is allowed to achieve "self-help compensation" through retaliation. The meaning of retaliation is that the trading partner can reciprocate by withdrawing a concession of a "substantially equivalent" nature.

The third element mentioned above is GATT/WTO dispute settlement procedures. Here, a central issue is the determination whether the actions by one country serve to "nullify or impair" the benefits expected under the agreement by another country. Nullification or impairment includes actions taken by one country "which harmed the trade of another, and which 'could not reasonably have been anticipated' by the other at the time it negotiated for a concession" (Jackson 1997, 115). In the typical "violation complaint," a country is alleged to have failed to comply with one or more of its GATT/WTO obligations, leading to a prima facie case of nullification or impairment.

An important distinction arises between the procedures associated with safeguard exceptions and those that are associated with nullification or impairment. The safeguard procedures provide explicitly for the *lawful* suspension of obligations or withdrawal of negotiated concessions, and these procedures specify as well the permissible retaliatory responses of trading partners. By contrast, the dispute settlement procedures govern retaliation against a country that takes a

harmful action that its trading partners could not have anticipated under GATT/WTO rules. In the typical complaint, at issue is whether the offending country has violated GATT/WTO rules, and retaliation here may then be more directly concerned with the enforcement of rules.

The procedure for settling disputes consists of three stages: First, there is a consultation phase among the involved parties; second, a GATT/WTO panel (and, after appeal, the appellate body) conducts an investigation and issues a ruling and recommendation; and as a last resort, authorization of retaliation occurs. Resolution is often achieved in the first stage, or it may follow the panel ruling. If the panel finds that nullification or impairment has occurred, then it recommends that the offending country correct any illegal measures. The offending country may be unwilling to do so, however. In this case, it may seek a negotiated resolution by offering the harmed country compensation through MFN tariff reductions on some other goods. If compensation is not offered, or if it is offered but rejected, then the harmed country may follow through with the last-resort response: an authorized and discriminatory suspension of tariff concessions. In practice, the number of authorized retaliations has been small, though this number has grown in the WTO era. As Rhodes (1993, 109) observes, however, the threat of authorized retaliation is often the catalyst for a resolution of the dispute in the earlier stages.

It is notable that, while authorized retaliation in the context of dispute resolution is allowed to be discriminatory, it is nevertheless generally limited to the suspension of concessions of a substantially equivalent nature. One might have thought that the GATT/WTO would authorize and coordinate maximal retaliation against a member government found to be in violation of the rules by the GATT/WTO's own dispute settlement body. But in fact, as the early report of the US International Chamber of Commerce (quoted in chapter 1) observed, the GATT dispute settlement procedures keep a *lid* on permissible retaliation levels, and this is how the GATT/WTO dispute settlement system works to avoid a trade war. This point was reflected in a statement made by one of the drafters of the original GATT articles governing retaliation in the context of dispute settlement, as found in Petersmann (1997, 82–83):

> The drafting history of Article XXIII:2 confirms that it was designed to limit the customary law right of unilateral reprisals, whose exercise had contributed so much to the "law of the jungle" in international economic affairs during

the 1930's, and to introduce, as stated by one of the drafters, "a new principle in international economic relations. We have asked the nations of the world to confer upon an international organization the right to limit their power to retaliate. We have sought to tame retaliation, to discipline it, to keep it within bounds. By subjecting it to the restraints of international control, we have endeavored to check its spread and growth, to convert it from a weapon of economic warfare to an instrument of international order."

Indeed, Schwartz and Sykes (2002) argue that the major innovation in the dispute settlement procedures of the WTO relative to GATT was the addition of a mechanism for arbitrating the magnitude of authorized retaliation so that an effective lid on retaliation could be maintained.

Finally, it is often observed that, along with MFN, reciprocity is a pillar of the GATT/WTO architecture. In the GATT/WTO, the principle of reciprocity refers to the ideal of 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. The preceding discussion contains two instances in which the notion of reciprocity arises. First, as I have observed, when governments negotiate in GATT/WTO rounds, they do so with the stated goal of obtaining mutually advantageous arrangements through reciprocal reductions in tariff bindings: In this context, it is often observed that governments approach negotiations seeking a "balance of concessions," whereby the market access value of the tariff cut offered by one government is balanced against an "equivalent" concession from its trading partner. This first instance of reciprocity therefore refers to changes in tariffs in a liberalizing direction. Second, when a government seeks to renegotiate its tariff commitments and modifies or withdraws a previous concession as an original action, and more generally whenever a government takes an action that nullifies or impairs the benefits expected under the agreement by another government, GATT/WTO rules permit substantially affected trading partners to retaliate in a reciprocal manner, by withdrawing "substantially equivalent concessions." This second instance of reciprocity refers to changes in tariffs in an upward direction.

The balance achieved through reciprocity in tariff negotiations and the role of retaliation in preserving this balance is reflected in the remark by a drafter of the GATT articles governing retaliation as quoted in Jackson (1969, 170–171):

> What we have really provided, in the last analysis, is not that retaliation shall be invited or sanctions invoked, but that a balance of interests once established, shall be maintained.

And the unique role of retaliation in the GATT legal system as a means of preserving reciprocity is pointed out by Dam (1970, 80–81):

> The best guarantee that a commitment of any kind will be kept (particularly in an international setting where courts are of limited importance and, even more important, marshals and jails are nonexistent) is that the parties continue to view adherence to their agreement as in their mutual interest. . . . Thus, the GATT system, unlike most legal systems . . . is not designed to exclude self-help in the form of retaliation. Rather, retaliation, subjected to established procedures and kept within prescribed bounds, is made the heart of the GATT system.

2.2 The Purpose of Trade Agreements

I now present the outlines of a basic modeling framework that will provide my foundation for an economic interpretation of GATT's design features, its successes, and ultimately its shortcomings.[4] In this section, I develop the model to answer one simple but fundamental question: What problems would governments want a trade agreement to help them solve? The answer to this question clarifies the purpose of a trade agreement and can help guide its design to serve that purpose.

To provide an answer, I abstract from possible domestic commitment problems that a government might face that could lead to *domestic* inefficiencies in its unilaterally chosen policies and that it might seek to solve with help from a trade agreement as an external commitment device.[5] I focus instead on characterizing the possible *international* inefficiencies that might arise under unilaterally chosen policies and that a trade agreement could address. A useful starting point for this purpose is the standard two-country, two-good general equilibrium model of trade familiar from any undergraduate international trade course.

The General Equilibrium Trade Model

The standard general equilibrium model of trade has two countries, home (no *) and foreign (*), who trade two goods that are normal goods in consumption and produced in perfectly competitive markets under conditions of increasing opportunity costs. I denote by x the natural import good of the home country and by y the natural import good of the foreign country, and I define $p \equiv p_x/p_y$ and $p^* \equiv p_x^*/p_y^*$ to

4. The material in this section builds from Bagwell and Staiger (2002, chap. 2).
5. On the possibility that trade agreements might help solve domestic commitment problems, see the literature reviewed in Bagwell and Staiger (2002, 32–34).

be, respectively, the local relative price in the home and foreign market. With τ the home-country import tariff and τ^* the foreign-country import tariff, each expressed in ad valorem terms and assumed to be set at nonprohibitive levels, it then follows that $p = (1+\tau)p^w \equiv p(\tau, p^w)$ and $p^* = p^w/(1+\tau^*) \equiv p^*(\tau^*, p^w)$, where $p^w \equiv p_x^*/p_y$ is the "world" (i.e., untaxed) relative price. The foreign terms of trade is then given by p^w while the home terms of trade is given by $(1/p^w)$. I am assuming for now that governments possess tariffs as their only tax/subsidy instrument. This ensures that both producers and consumers face the same local relative price in the market within which they reside. In later chapters, I will introduce into the model a richer array of government policies that include the possibility of regulatory standards as well as production and/or consumption taxes/subsidies; in the presence of the additional tax/subsidy policies, consumers and producers residing in the same market may face different local prices.

Production possibilities in each country are defined by a production possibilities frontier, which with Q denoting production, I represent by the decreasing and concave function $Q_y(Q_x)$ in the home country and $Q_y^*(Q_x^*)$ in the foreign country, defined over the feasible values of production of x in each country. Production in a country occurs at the point on the production possibilities frontier where the marginal rate of transformation between x and y is equal to the local relative price, allowing home and foreign production functions to be represented as $Q_i = Q_i(p)$ and $Q_i^* = Q_i^*(p^*)$ for $i = \{x, y\}$. Consumption depends on both the local relative price—which defines the trade-off faced by consumers and, in determining the point on the production possibilities frontier at which the economy operates, also implies the level and distribution of factor income in the economy measured at local prices—and on tariff revenue, which is distributed lump-sum back to consumers in the country where it is collected. I denote by R the tariff revenue collected in the home country and by R^* the tariff revenue collected in the foreign country, each measured in units of the country's export good at local prices. National consumption in the home and foreign country can then be written as $D_i = D_i(p, R)$ and $D_i^* = D_i^*(p^*, R^*)$ for $i = \{x, y\}$, where tariff revenue is defined implicitly by $R = [D_x(p, R) - Q_x(p)][p - p^w]$ or $R = R(p, p^w)$ for the home country and by $R^* = [D_y^*(p^*, R^*) - Q_y^*(p^*)][1/p^* - 1/p^w]$ or $R^* = R^*(p^*, p^w)$ for the foreign country, and where each country's tariff revenue is an increasing function of its terms of trade under the normal-goods assumption. This allows national consumption to be written

as $C_i(p, p^w) \equiv D_i(p, R(p, p^w))$ and $C_i^*(p^*, p^w) \equiv D_i^*(p^*, R^*(p^*, p^w))$ for $i = \{x, y\}$, with C_i decreasing in p^w and C_i^* increasing in p^w.

To express the trade balance and equilibrium conditions of the model, I define home-country imports of x and exports of y by $M_x(p, p^w) \equiv C_x(p, p^w) - Q_x(p)$ and $E_y(p, p^w) \equiv Q_y(p) - C_y(p, p^w)$, respectively. Similarly, foreign-country imports of y and exports of x are defined by $M_y^*(p^*, p^w) \equiv C_y(p^*, p^w) - Q_y^*(p^*)$ and $E_x^*(p^*, p^w) \equiv Q_x^*(p^*) - C_x^*(p^*, p^w)$, respectively. For any world price, we also have

$$p^w M_x(p(\tau, p^w), p^w) = E_y(p(\tau, p^w), p^w) \text{ and} \tag{2.1}$$

$$M_y^*(p^*(\tau^*, p^w), p^w) = p^w E_x^*(p^*(\tau^*, p^w), p^w), \tag{2.2}$$

which are the balanced trade conditions, where I now make explicit the dependence of the local price on the tariff and the world price. The equilibrium world price, $\widetilde{p}^w(\tau, \tau^*)$, is then determined by the requirement of market clearing for good y:

$$E_y(p(\tau, \widetilde{p}^w), \widetilde{p}^w) = M_y^*(p^*(\tau^*, \widetilde{p}^w), \widetilde{p}^w), \tag{2.3}$$

with market clearing for good x implied by (2.1), (2.2), and (2.3).

Thus, given any pair of tariffs, the equilibrium world price is determined by (2.3), and the equilibrium world price and the given tariffs then determine in turn the local prices and thereby the production, consumption, import, export, and tariff revenue levels. I focus on the standard case and therefore assume that the Lerner and Metzler paradoxes[6] are ruled out so that

$$\frac{\partial \widetilde{p}^w(\tau, \tau^*)}{\partial \tau} < 0 < \frac{\partial \widetilde{p}^w(\tau, \tau^*)}{\partial \tau^*} \text{ and}$$
$$\frac{dp(\tau, \widetilde{p}^w(\tau, \tau^*))}{d\tau} > 0 > \frac{dp^*(\tau^*, \widetilde{p}^w(\tau, \tau^*))}{d\tau^*}. \tag{2.4}$$

For future reference, I note that the first set of inequalities in (2.4) implies that, if the home tariff τ were reduced by a small amount, there exists a small reduction in the foreign tariff τ^* that would hold the equilibrium world price $\widetilde{p}^w$ constant.

Government Objectives

I now turn to the specification of government objectives. The trade policy objectives of real-world governments are diverse, and it is

6. Bagwell and Staiger (2016, 499–501) consider the implications for the purpose of trade agreements when the Metzler and/or Lerner paradoxical cases arise.

important to allow for this diversity when considering the purpose of a trade agreement, lest the purpose ascribed to the agreement is unduly limited by the trade policy objectives ascribed to governments. Even in the simple model of a world economy presented here, there are many possible motives for government trade policy intervention that could be entertained.

For example, a government might care only about the level of national consumption and hence the level of real national income when choosing its tariffs, either because it is unconcerned about the distribution of income and consumption among its citizens or because it has lump-sum redistributive instruments to handle these concerns. The preferences of such a government in the home country could be represented in the model with the objective function $G(C_x, C_y)$, with G increasing in both arguments. Notice that, as $C_i(p, p^w)$ is decreasing in p^w for $i \in \{x, y\}$ as indicated above, I can also write this objective function as

$$G(C_x(p, p^w), C_y(p, p^w)) \equiv W(p, p^w), \tag{2.5}$$

where W is decreasing in p^w; similarly, for the foreign government, I can write

$$G^*(C_x^*(p^*, p^w), C_y^*(p^*, p^w)) \equiv W^*(p^*, p^w), \tag{2.6}$$

where W^* is increasing in p^w given that $C_i^*(p^*, p^w)$ is increasing in p^w.

But real-world governments often view tariffs as a tool to address distributional concerns.[7] Why would these governments use tariffs for this purpose when it is well known that there are other policy interventions that are, in principle, better suited for this task? One reason could be that in practice, these governments lack not only the policy ideal of lump-sum taxes but also any of the other policy instruments that, if available, would typically dominate tariffs as tools for influencing the distribution of income and preserve the economist's case for free trade. In the context of this limited set of policy options, tariffs might then be the best available policy response to address these concerns.[8]

7. These concerns likely reflect a combination of a desire of governments to serve some notion of social welfare, such as that embodied in the "conservative social welfare function" introduced by Corden (1974), and political economy motives that serve politically favored groups (as in Grossman and Helpman 1994).

8. There are a variety of reasons why, as a practical matter, such nontariff instruments may not be available to governments. They include administrative costs and funding

Or it could be that, even though some of these policy instruments are technically available to governments, the welfare of their citizens is determined by more than simply the material standard of living that can be attained with a given level of consumption; it might depend as well on the *manner* in which the income to support this level of consumption is attained, with the receipt of lump-sum transfers or direct subsidy payments diminishing personal dignity in a way that earning income at market prices—even if not the prices that would prevail under free trade—would not.

In any event, the fact is that many governments use tariffs to address distributional concerns and, more broadly, as tools of industrial policy, and therefore they choose tariffs to affect the sectoral pattern of production in their economies for reasons that go beyond how that production translates into real national income and thereby national consumption levels. In terms of the model, these governments would appear to have preferences over where on the production possibilities frontier their economy operates, independent of the national consumption levels that are attained. Such government preferences for the home country could be represented in the model by the objective function $G(C_x, C_y, Q_x, Q_y(Q_x))$. The distribution and level of factor income measured in local prices would be pinned down for a given choice of Q_x and therefore $Q_y(Q_x)$ on the production possibilities frontier. Conditional on the aggregate level of national consumption C_x and C_y, the home government would then have its own preference ranking over the choice of Q_x and $Q_y(Q_x)$ as reflected in the function G. For given Q_x and $Q_y(Q_x)$ and the factor incomes that are implied, it is again natural that G is increasing in C_x and C_y, because when factor incomes are fixed, increasing C_x and C_y amounts to increases in tariff revenue according to the national budget constraint. Notice again that I can write this objective function as

$$G(C_x(p, p^w), C_y(p, p^w), Q_x(p), Q_y(Q_x(p))) \equiv W(p, p^w), \tag{2.7}$$

where W is decreasing in p^w. And similarly for the foreign government, I can write

requirements that, when taken into account, could make these instruments impractical or at least less attractive than tariffs. See also Rodrik (1987), Drazen and Limão (2008), and Limão and Tovar (2011) on additional reasons why governments may choose to use tariffs for purposes of redistribution. I discuss the possible role of tariffs as a tool of industrial policy more generally in chapter 7.

$$G^*(C_x^*(p^*, p^w), C_y^*(p^*, p^w), Q_x^*(p^*), Q_y^*(Q_x^*(p^*))) \equiv W^*(p^*, p^w), \quad (2.8)$$

where W^* is increasing in p^w.

More generally, a government's preferences over the sectoral pattern of production in its economy could arise for reasons of national security, or from the societal benefits of maintaining a robust middle class with access to stable and good-paying jobs that are more prevalent in one sector than they are in another, or from the desire to preserve employment in a region that is dependent on a particular sector, or from the avoidance of sector-specific negative externalities of an "eyesore" variety. Any of these nonpecuniary features could be embedded in the model without changing the formal structure that I have outlined above, as long as they do not invalidate the competitive equilibrium conditions that the model assumes or lead to transborder nonpecuniary externalities. And, for each of these cases, I can once again write the associated home-government objective function as in (2.7), with W decreasing in p^w, and similarly I can again write the associated foreign-government objective function as in (2.8) with W^* increasing in p^w.

Evidently, in all of the cases I have described, government preferences can be represented in the model with the home-country and foreign-country objective functions expressed in the form $W(p, p^w)$ and $W^*(p^*, p^w)$, respectively, where W is decreasing in p^w and W^* is increasing in p^w and where the difference across these various government objectives translates into differences in how W varies with p and how W^* varies with p^*.[9] To capture all these possibilities in a unified framework, I will therefore follow Bagwell and Staiger (1999, 2002) and represent the trade policy objectives of the home and foreign government with the general functions $W(p, p^w)$ and $W^*(p^*, p^w)$, with the only structure placed on W and W^* that, holding its local price fixed, each government is assumed to achieve higher welfare when its terms of trade improve:[10]

$$\frac{\partial W(p, \tilde{p}^w)}{\partial \tilde{p}^w} < 0 \text{ and } \frac{\partial W^*(p^*, \tilde{p}^w)}{\partial \tilde{p}^w} > 0. \quad (2.9)$$

9. See also Bagwell and Staiger (1999; 2002, 18–21) for an inventory of the formal models of trade policy determination in the economics literature that are captured by this structure.

10. See Bagwell and Staiger (2002, 19–20) for a description of the change in the home and foreign tariff that would increase $\tilde{p}^w$ while holding fixed an economy's local price. Throughout, I also impose standard regularity conditions so that all second-order conditions are globally satisfied and all partial derivatives of W and W^* are finite.

The Purpose of a Trade Agreement

I now turn to the central question of this chapter: What problems would governments want a trade agreement to help them solve? In the absence of a trade agreement, I assume that each government would set its trade policy to maximize its objective function, taking as given the tariff choice of its trading partner. This yields the following home and foreign reaction functions:

$$\textit{Home Reaction Function}: W_p \frac{dp}{d\tau} + W_{p^w} \frac{\partial \widetilde{p}^w}{\partial \tau} = 0 \tag{2.10}$$

$$\textit{Foreign Reaction Function}: W^*_{p^*} \frac{dp^*}{d\tau^*} + W^*_{p^w} \frac{\partial \widetilde{p}^w}{\partial \tau^*} = 0, \tag{2.11}$$

where subscripts denote partial derivatives. The joint solution to (2.10) and (2.11) defines the noncooperative (Nash) tariff pair (τ^N, τ^{*N}). Notice that under (2.4) and (2.9), the home-country reaction function (2.10) implies $W_p < 0$ while the foreign-country reaction function (2.11) implies $W^*_{p^*} > 0$. I will return to this feature of noncooperative tariffs below.

Under a trade agreement, by contrast, I assume that the two governments negotiate to a position on the efficiency frontier, where this frontier is defined by

$$\max_{\tau,\tau^*} W(p(\tau, \widetilde{p}^w), \widetilde{p}^w) \tag{2.12}$$

$$s.t.\ \ W^*(p^*(\tau^*, \widetilde{p}^w), \widetilde{p}^w) \geq \overline{W}^*,$$

with $\overline{W}^*$ denoting any feasible level of foreign welfare. The efficiency frontier is characterized by solving (2.12) for each value of $\overline{W}^*$, and it traces out the locus of Pareto efficient tariff pairs (τ^E, τ^{*E}). The associated first-order conditions are

$$W_p \frac{dp}{d\tau} + W_{p^w} \frac{\partial \widetilde{p}^w}{\partial \tau} + \lambda \left[\left(W^*_{p^*} \frac{\partial p^*}{\partial p^w} + W^*_{p^w} \right) \frac{\partial \widetilde{p}^w}{\partial \tau} \right] = 0 \tag{2.13}$$

$$\left[W_p \frac{\partial p}{\partial p^w} + W_{p^w} \right] \frac{\partial \widetilde{p}^w}{\partial \tau^*} + \lambda \left[W^*_{p^*} \frac{dp^*}{d\tau^*} + W^*_{p^w} \frac{\partial \widetilde{p}^w}{\partial \tau^*} \right] = 0, \tag{2.14}$$

where λ is the Lagrange multiplier on the constraint in (2.12). Solving (2.13) for λ and substituting the result into (2.14), together with the price definitions, yields the condition that defines the locus of efficient tariffs:

$$\left[\tau W_p + W_{p^w}\right]\frac{\partial \widetilde{p}^w}{\partial \tau^*} - \left[\frac{\left[W_p \frac{dp}{d\tau} + W_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau}\right] \times \left[W^*_{p^*}\frac{dp^*}{d\tau^*} + W^*_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau^*}\right]}{\left[\frac{1}{\tau^*}W^*_{p^*} + W^*_{p^w}\right]\frac{\partial \widetilde{p}^w}{\partial \tau}}\right] = 0. \tag{2.15}$$

A familiar special case of the efficiency locus defined by (2.15) arises when governments care only about the level of national consumption and hence the level of real national income when choosing their tariffs. In this case, as I have noted above, we then have that the home and foreign welfare functions $W(p, p^w)$ and $W^*(p^*, p^w)$ can be written in the particular form given in (2.5) and (2.6), respectively, and it is straightforward to show that (2.15) then simplifies to the Mayer (1981) locus of efficient tariffs defined by $(1+\tau) = 1/(1+\tau^*)$. The Mayer locus includes the point of reciprocal free trade $\tau = 0 = \tau^*$, but it also includes a locus of other efficient pairs of tariffs in which an import tariff in one country is exactly offset by an import subsidy of the same magnitude in the other country. To understand the conditions for efficiency along the Mayer locus, notice that at any point on the locus we have

$$p = (1+\tau)\widetilde{p}^w(\tau, \tau^*) = \frac{1}{(1+\tau^*)}\widetilde{p}^w(\tau, \tau^*) = p^*.$$

Hence, along the Mayer locus, tariffs are adjusted to maintain equality in relative local prices between the home and foreign countries, with different tariff pairs resulting in different world prices and therefore different distributions of income across trading partners through shifts in the (positive or negative) tariff revenue collected by each country. When $W(p, p^w)$ and $W^*(p^*, p^w)$ are not assumed to conform to the particular structure in (2.5) and (2.6), equation (2.15) still determines the efficient relationship between home and foreign tariffs, but it need not be the case that this relationship equates relative local prices across trading partners, and it need not be the case that this relationship is satisfied by reciprocal free trade.

Continuing now with the general government preferences $W(p, p^w)$ and $W^*(p^*, p^w)$ as described above, a first question is whether the noncooperative tariff choices are efficient. If they are, then assuming that the two governments have entered into negotiations voluntarily, there is nothing for a trade agreement to do since it cannot offer a Pareto improvement over the noncooperative outcome. Using (2.10) and (2.11) together with (2.4) and (2.9), and also using the fact that the noncooperative tariffs imply $W_p < 0$ and $W^*_{p^*} > 0$, it is straightforward to confirm that the first-order condition for efficiency given in (2.15) is

violated when evaluated at the noncooperative tariff pair (τ^N, τ^{*N}) defined by (2.10) and (2.11); more specifically, the left-hand side of (2.15) is strictly negative. This implies that, regardless of which of the underlying motives for tariff intervention included in the general government objective functions $W(p, p^w)$ and $W^*(p^*, p^w)$ is operative, noncooperative tariffs are *too high* relative to the efficiency locus.[11] And as Bagwell and Staiger (1999; 2002, chap. 2) demonstrate, starting at the Nash equilibrium, mutual gains for governments are therefore possible only if they both cut their tariffs. Clearly, this case for tariff liberalization in a trade agreement has nothing to do with the economist's case for free trade, since it arises regardless of the underlying motives for trade protection captured in the general government objective functions $W(p, p^w)$ and $W^*(p^*, p^w)$, and as discussed above, many of those motives would violate the assumptions that underlie the case for free trade as an efficient outcome.

We may now ask, Why are noncooperative tariffs inefficiently high? If we can identify the reason, then we can say that addressing this reason is the problem that governments want a trade agreement to help them solve. We can say this because by solving this problem, a trade agreement would bring countries to the efficiency frontier, and at that point there is no possibility of further Pareto gains for the governments.

To proceed formally, we need to characterize the difference between the Nash first-order conditions in (2.10) and (2.11) and the first-order conditions for efficiency given in (2.15). To aid in this characterization, it is useful to pick a specific point on the efficiency locus and compare the conditions that define that pair of efficient tariffs to the conditions that define the pair of Nash tariffs.

A point on the efficiency locus that is particularly illuminating for this purpose is the point that Bagwell and Staiger (1999) call the "political optimum," defined as the tariff pair (τ^{PO}, τ^{*PO}) that satisfies

$$\textit{Home Political Optimum}: W_p \frac{dp}{d\tau} = 0 \Leftrightarrow W_p = 0 \tag{2.16}$$

$$\textit{Foreign Political Optimum}: W^*_{p^*} \frac{dp^*}{d\tau^*} = 0 \Leftrightarrow W^*_{p^*} = 0, \tag{2.17}$$

11. In particular, the fact that the left-hand side of (2.15) is strictly negative when evaluated at the noncooperative tariff pair (τ^N, τ^{*N}) means that τ^{*N} is too high relative to the level of τ^* that would be efficient in combination with τ^N. Analogously, τ^N is too high relative to the level of τ that would be efficient in combination with τ^{*N}. It is in this sense that noncooperative tariffs (τ^N, τ^{*N}) are too high relative to the efficiency locus.

where the second equality in (2.16) and in (2.17) follows from the second set of inequalities in (2.4). In the special case where governments care only about the level of national consumption and hence the level of real national income when choosing their tariffs, and where the government objectives therefore take the particular form in (2.5) and (2.6), the politically optimal tariffs correspond to reciprocal free trade, a point on the Mayer locus. That politically optimal tariffs are efficient as well under the general government objective functions $W(p, p^w)$ and $W^*(p^*, p^w)$ described above can be immediately confirmed using (2.16) and (2.17) by noting that, when evaluated at the tariff pair (τ^{PO}, τ^{*PO}), the condition for efficiency (2.15) is satisfied:

$$\left[\tau W_p + W_{p^w}\right]\frac{\partial \widetilde{p}^w}{\partial \tau^*} - \left[\frac{\left[W_p \frac{dp}{d\tau} + W_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau}\right] \times \left[W^*_{p^*}\frac{dp^*}{d\tau^*} + W^*_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau^*}\right]}{\left[\frac{1}{\tau^*}W^*_{p^*} + W^*_{p^w}\right]\frac{\partial \widetilde{p}^w}{\partial \tau}}\right]$$
$$= W_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau^*} - W_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau^*} = 0.$$

But comparing (2.16) and (2.17) to (2.10) and (2.11), it is now also apparent that the noncooperative tariffs fail to reach the political optimum because of the presence of a single term, $W_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau}$, in the home-country reaction curve and a single term, $W^*_{p^w}\frac{\partial \widetilde{p}^w}{\partial \tau^*}$, in the foreign-country reaction curve. These terms represent the incentive each country has when choosing its tariff noncooperatively to manipulate the terms of trade in its favor and thereby to shift a portion of the costs of its tariff intervention onto its trading partner.

For the home government, this term is the product of two negative terms: the term $\frac{\partial \widetilde{p}^w}{\partial \tau}$, which is strictly negative as long as the home country is large and therefore has market power on world markets; and the term W_{p^w}, which is also negative and reflects the negative income effect of a terms-of-trade deterioration holding local prices in the home economy fixed. And as this product is itself positive, its presence in (2.10) drives the home noncooperative tariff choice higher than the tariff that would imply $W_p = 0$, ensuring that at the noncooperative tariff, we in fact have $W_p < 0$ (as I have observed).

For the foreign government, this term is the product of two positive terms: the term $\frac{\partial \widetilde{p}^w}{\partial \tau^*}$, which is strictly positive as long as the foreign country is large and therefore has market power on world markets; and the term $W^*_{p^w}$, which is also positive and reflects the positive income effect of a terms-of-trade improvement holding local prices in

the foreign economy fixed. And as this product is itself also positive, its presence in (2.11) drives the foreign noncooperative tariff choice higher than the tariff that would imply $W_{p^*}^* = 0$, ensuring that at the noncooperative tariff, we in fact have $W_{p^*}^* > 0$ (as I have observed).

The fact that these terms lead the home and foreign government to choose tariffs in the noncooperative equilibrium that imply $W_p < 0$ and $W_{p^*}^* > 0$ is also revealing. As Bagwell and Staiger (1999; 2002, chap. 4) show, if each government were offered the opportunity to alter its tariff from the noncooperative level without impacting its terms of trade, it would choose to *cut* its tariff: The home tariff cut would decrease the local relative price p in the home economy according to the second inequality in (2.4), leading to a rise in home welfare in the amount $\Delta W = W_p[-\frac{\partial p}{\partial \tau}] > 0$; and the foreign tariff cut would increase the local relative price p^* in the foreign economy according to the second inequality in (2.4), leading to a rise in foreign welfare in the amount $\Delta W^* = W_{p^*}^*[-\frac{\partial p^*}{\partial \tau^*}] > 0$. Viewed in this light, it is then clear that it is the ability of each government to shift some of the costs of its tariff onto its trading partner through terms-of-trade movements that drives each government to choose the overly high tariffs that obtain in the noncooperative equilibrium.

Hence, regardless of which of the underlying motives for tariff intervention included in the general government objective functions $W(p, p^w)$ and $W^*(p^*, p^w)$ is operative, the purpose of a trade agreement is the same: to eliminate the unilateral incentive that governments have to manipulate their terms of trade and thereby help governments escape from a terms-of-trade-driven prisoner's dilemma.

Bagwell and Staiger (1999, 2002) make this same point, but from the other direction. They observe that the Nash first-order conditions (2.10) and (2.11) would be converted to the conditions (2.16) and (2.17) if the terms-of-trade manipulation terms $W_{p^w}\frac{\partial \tilde{p}^w}{\partial \tau}$ and $W_{p^w}^*\frac{\partial \tilde{p}^w}{\partial \tau^*}$ were dropped from (2.10) and (2.11), respectively. Further, they demonstrate that the conditions (2.16) and (2.17) define a point on the efficiency frontier, which they refer to as the political optimum. They then observe that the politically optimal tariffs can be interpreted as the tariffs that would arise under unilateral choices in a hypothetical world in which governments are not motivated by the terms-of-trade implications of their trade policy choices, in the sense that the home government acted as if $W_{p^w} \equiv 0$ and the foreign government acted as if $W_{p^w}^* \equiv 0$. And by showing that the tariffs selected unilaterally by governments with these hypothetical preferences would satisfy (2.16) and (2.17) and thus be

efficient, where the evaluation of efficiency is undertaken with respect to the actual government preferences, they conclude that when governments have objectives that can be represented by the general form $W(p, p^w)$ and $W^*(p^*, p^w)$ subject to (2.9), the only rationale for a trade agreement is to eliminate the unilateral incentive that governments have to manipulate their terms of trade.

Whether politically optimal tariffs are seen as a particular point on the efficiency frontier that can be usefully compared to the first-order conditions defining the noncooperative tariffs, as I have emphasized here, or as a useful hypothetical thought experiment for noncooperative tariff choices, as in the original Bagwell and Staiger (1999, 2002) presentation, is immaterial. As long as politically optimal tariffs as defined by (2.16) and (2.17) are efficient in a given environment, we can conclude from the Nash first-order conditions (2.10) and (2.11) that the purpose of a trade agreement in that environment is to eliminate the unilateral incentive that governments have to manipulate their terms of trade.[12]

Positive but Also Normative?

Now is a good time to pause and consider a question that has been lurking behind the approach that I have adopted for identifying the purpose of a trade agreement. I have accepted the sovereign right of each national government to define its own policy preferences. I have then characterized the task that a trade agreement must accomplish if it is to eliminate the international inefficiencies associated with unilateral policy choices as judged by the preferences of the member governments. I have called this task the purpose of a trade agreement. Because the GATT/WTO is a member-driven organization and the members are national governments, this seems a reasonable approach from which to draw positive conclusions about the purpose of a trade agreement. But does this approach also have normative implications? Is it enough for the world trading system to serve the interests of its member *governments*? Can a case for the legitimacy of the GATT/WTO be built around a demonstration that it is well designed to serve these interests, where by "legitimacy" I have in mind a "right to rule" concept along the lines articulated by Buchanan and Keohane (2006)?[13]

12. Notice that I have said nothing here about whether a trade agreement would actually implement the political optimum, only that the politically optimal tariffs are useful as a comparator to noncooperative tariffs when evaluating the purpose of a trade agreement.
13. Buchanan and Keohane (2006, 411) define legitimacy in the case of global governance institutions as "the right to rule, understood to mean both that institutional agents are

If national governments were always and everywhere the faithful servants of their citizens, where the desires of their citizens were aggregated into policy directives for the governments through political processes that their citizens saw as legitimate, then the answers to these questions would clearly be "yes." But most real-world governments operate far from this ideal. And so, in the real world, the answers are not so clear.

Looking to the international political economy literature for guidance on these questions provides a mixed view. On the specific question of what determines the legitimacy of an international institution, Peter (2017) notes that there are two approaches in the literature: a "state-centered" approach and a "people-centered" approach. Beitz (1979, 408) describes the state-centered approach as one in which "international society is understood as domestic society writ large, with states playing the roles occupied by persons in domestic society." In the people-centered approach, it is instead the welfare of individuals that is taken as the basis for the determination of an international institution's legitimacy (Buchanan 2003). If the purpose of a trade agreement that I have identified above can be interpreted as having normative relevance, then establishing a claim of legitimacy for the GATT/WTO based on a demonstration that it is well designed to serve this purpose falls squarely on the state-centered approach. Under this interpretation, like the preferences of consumers in a domestic context, the preferences of national governments are taken as sovereign in the international context, and the legitimacy of a trade agreement is judged on its ability to deliver efficient outcomes where efficiency is assessed using the preferences of the member governments.[14] This interpretation seems tenuous, but what are the viable alternatives?

One possibility would be to dispense completely with the nation-state as the unit of observation for normative purposes and to evaluate the legitimacy of the GATT/WTO based on how close the agreement comes to maximizing a global social welfare function defined over the welfare of individuals. This would amount to a people-centered approach. For example, the GATT/WTO's design might be judged with

morally justified in making rules and attempting to secure compliance with them and that people subject to those rules have moral, content-independent reasons to follow them and/or to not interfere with others' compliance with them." See also Franck (1990).

14. To be clear, while this approach can be described as state-centered, it is otherwise distinct from the approaches to evaluating legitimacy featured in the international political economy literature and reviewed in Peter (2017), as it uses a different set of (state-centered) criteria.

a criterion based on a utilitarian ideal, where global welfare is measured by the sum of the utilities across all individuals in the world and where each individual's utility enters that sum with an equal weight.[15] Or a Rawlsian criterion, under which global social welfare is only as high as the utility of the least-well-off individual on the planet, might be used to judge the design of the agreement. As a general matter, it is of course important to know how an agreement performs according to these normative benchmarks. But as a means to evaluate the legitimacy of the GATT/WTO, these benchmarks seem unworkable, because to proceed with such an evaluation would require that a consensus emerge regarding the correct normative benchmark, and it seems unlikely that such a consensus could ever exist.[16]

Another possibility for assessing legitimacy would be a hybrid approach somewhere in between the state-centered and people-centered approaches, maintaining the nation-state as the unit of observation but including more interests from each nation in the global social welfare function than simply the interests of each member government. Such an approach might, for example, mirror the "tripartite" structure of national representation in the International Labor Organization (ILO), where each member country is represented by three national interests: its government, its workers, and its employers. The analogue for assessing the legitimacy of the GATT/WTO might be to include in the global social welfare function used in that assessment representatives of government, exporter, and importer interests in each member country (or possibly government, producer, and consumer interests). But again, a consensus on the appropriate representation would be needed to make this approach workable.[17]

15. See Maggi and Ossa (2020) for an approach to evaluating the normative properties of a trade agreement along these lines.

16. Partly the difficulty in reaching a consensus on this matter rests with the fact that it involves value judgments over which there will always be disagreements. And partly the difficulty can be traced to disagreements over factual matters, such as the importance of market failures and the array of policy instruments that real-world governments have to pursue their objectives.

17. There is also another issue raised by moving away from a state-centered approach to evaluating the legitimacy of a trade agreement: If interests beyond those of the member governments are to be represented in a trade agreement, how are commitments that serve those interests but not also the interests of the member governments to be enforced? This issue seems germane for the GATT/WTO, where enforcement ultimately comes down to tariff retaliation and governments hold the levers of this enforcement mechanism, and it may explain why under the ILO's tripartite representation (unique among United Nations agencies) no member state is under any obligation to ratify any ILO convention or recommendation (see Johnston 1970, 90).

In light of these consideration, it is useful to think of the question of the legitimacy of the world trading system as applying at two levels. First, at the international level there is the question of whether the GATT/WTO can be seen as legitimate from the perspective of the member governments. And second, at the national level there is the question of whether the member governments can be seen as legitimate from the perspective of their own citizens. If both questions can be answered in the affirmative, then the GATT/WTO can be said to be legitimate from both the state-centered and the people-centered perspective. But as trade agreements are fundamentally government-to-government contracts, the key question of legitimacy for the GATT/WTO as an international institution—and the only question whose answer it has any meaningful control over—relates to the first question, not the second.

My approach in this book is to therefore focus on the answer to the first question—Does the GATT/WTO have the moral authority to make rules and attempt to secure compliance with those rules from its member governments—and to acknowledge that an answer to this question can provide only part of the answer to the larger question of the legitimacy of the world trading system. But it is an important part of the answer. If this first question *cannot* be answered in the affirmative, then it is hard to see how the GATT/WTO could remain viable, since it would presumably lack support from the governments that are its members. And if this question can be answered in the affirmative, then the central *international* task in designing a constitution for the world trading system has been accomplished with the design of the GATT/WTO. And with this state-centered task accomplished, attention could then be focused on the task of establishing that each national government satisfies agreed criteria for legitimacy, thereby ensuring that the world trading system, so designed, could be said to be legitimate from a people-centered perspective as well.

Generality

Thus far I have emphasized the wide array of government objectives that are consistent with the conclusion that the purpose of a trade agreement is to eliminate the unilateral incentive that governments have to manipulate their terms of trade. But I have maintained a very particular and simple economic environment within which to derive these results. How dependent is this conclusion on the economic environment within which governments operate? An immediate implication of the discussion above is that this conclusion does depend on governments having

a complete set of trade taxes at their disposal. This can be seen from the definition of politically optimal tariffs, which in general requires the use of both τ and τ^* to satisfy the two conditions in (2.16) and (2.17).[18] As has been emphasized by Ossa (2011) and Bagwell and Staiger (2012, 2015, 2016), when limitations are placed on the trade taxes that governments possess, different roles for a trade agreement can arise. That said, some of the most salient restrictions on trade tax/subsidy instruments are associated with commitments made as a *result* of trade agreements (e.g., to restrict the use of export subsidies), and it is not clear that such restrictions should be taken as given when attempting to identify the underlying purpose of trade agreements, as is my intent here.

Beyond the assumption that governments have a complete set of trade taxes, however, the conclusion that the purpose of a trade agreement is to eliminate the unilateral incentive that governments have to manipulate their terms of trade is surprisingly robust to alternative economic environments. It holds in a many-country version of the model that I have outlined, provided that tariffs are imposed on a nondiscriminatory (MFN) basis (Bagwell and Staiger 1999, 2002), and it holds in partial equilibrium versions of these models (Bagwell and Staiger 2001a). It holds in competitive environments for trade in goods or trade in services when governments have access to regulatory standards and/or additional domestic tax/subsidy policies (see Bagwell and Staiger 2001b; Staiger and Sykes 2011, for trade in goods; Staiger and Sykes 2021, for trade in services). And it holds in models of Cournot or monopolistic competition with homogeneous firms (Bagwell and Staiger 2002, chap. 9; 2012a; 2012b; 2015) and in models of monopolistic competition with heterogeneous firms (Bagwell and Lee 2020; Campolmi, Fadinger, and Forlati 2020; Costinot, Rodriguez-Clare, and Werning 2016, 2020). For this reason, it is useful to adopt a common shorthand for referring to models that share this prediction

18. An exception is when the government objective functions take the particular form in (2.5) and (2.6) and the politically optimal tariffs correspond to reciprocal free trade, a point on the Mayer locus. In this case, if only one of the two governments had access to a tariff, it could still be concluded that the purpose of a trade agreement is to eliminate the unilateral incentive that this government has to manipulate its terms of trade, because at the political optimum neither government imposes a tariff, so it is immaterial that one of them does not have access to a tariff. And Staiger and Sykes (2021) show that the lack of available trade taxes that can arise with certain types of services trade does not change the purpose of a trade agreement; I review their findings in chapter 9. Also, to be clear, notice that for the arguments in the text to remain valid, there is no requirement that governments have a complete set of *tax* instruments, only that they have a complete set of *trade* taxes.

about the purpose of a trade agreement, and I will follow Bagwell and Staiger's (2002) terminology and sometimes make use of the phrase "terms-of-trade theory of trade agreements" as a catchall for models of this kind.[19]

This is not to say that preventing terms-of-trade manipulation is the only possible purpose for a trade agreement. Indeed, as I noted at the outset of this chapter, I am intentionally abstracting from the possibility that a trade agreement could serve as a policy commitment device for its member governments when those governments struggle to make policy commitments to their private sectors on their own. And as I will review in later chapters, the arguments I have made here do not extend to all economic environments; as I alluded to in chapter 1, some of the environments where these arguments do not extend may be more important in the twenty-first century than they were in the twentieth century, raising the possibility of an evolution of the purpose of trade agreements over time. But as I have illustrated here, these arguments do apply in a remarkably broad set of circumstances, suggesting that a trade agreement that is designed well to solve the terms-of-trade manipulation problem will be a very useful trade agreement to its member governments. From this perspective, it is therefore meaningful when assessing the reasons for GATT's success and determining the basis for its legitimacy to evaluate the degree to which its design features are well equipped to serve this function. It is to this evaluation that I now turn.

19. But see Grossman (2016) for a different perspective on this terminology.

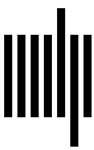
A World Trading System for the Twenty-First Century

By: Robert W. Staiger

Citation:

A World Trading System for the Twenty-First Century
By: Robert W. Staiger
DOI: 10.7551/mitpress/13574.001.0001
ISBN (electronic): 9780262371292
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Palatino LT by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data
Names: Staiger, Robert W., author.
Title: A world trading system for the twenty-first century / Robert W. Staiger.
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Series: Ohlin lectures | Includes bibliographical references and index.
Identifiers: LCCN 2022000723 (print) | LCCN 2022000724 (ebook) | ISBN 9780262047302 (hardcover) | ISBN 9780262371292 (pdf) | ISBN 9780262371308 (epub)
Subjects: LCSH: International economic relations. | Commercial treaties. | Foreign trade regulation. | Foreign trade regulation—Developing countries. | Globalization.
Classification: LCC HF1365 .S73 2022 (print) | LCC HF1365 (ebook) | DDC 337.1—dc23/eng/20220310
LC record available at https://lccn.loc.gov/2022000723
LC ebook record available at https://lccn.loc.gov/2022000724