CHAPTER 7

The Purpose of Trade Agreements

G.M. Grossman[1]
[1]Princeton University, Princeton, NJ, United States

Contents

Abstract

In this chapter, I review the literature on governments' motivations for negotiating and joining international trade agreements. I discuss both normative explanations for trade agreements and explanations based on political-economy concerns. Most of the chapter focuses on the purpose of multilateral agreements, but I do discuss briefly the reasons we might see governments forming preferential or regional trade agreements that exclude some countries.

Keywords

Trade agreements, Trade pacts, International cooperation, Multilateralism, Regionalism

JEL Classification:

F13, F53, K33

Handbook of Commercial Policy, Volume 1A
ISSN 2214-3122, http://dx.doi.org/10.1016/bs.hescop.2016.04.016

1. INTRODUCTION

Britain and France inked the first modern trade agreement on January 23, 1860. The so-called Cobden-Chevalier Treaty promised that France would eliminate all import prohibitions on British manufactured goods while capping most duties at 30% (25% after 1865). Britain in turn agreed to remove import barriers entirely from all but 48 French commodities while reducing dramatically its tariffs on French wine and brandy (Ashley, 1904). Notably, each country promised to grant the other most-favored-nation (MFN) consideration with regard to any tariff concessions it might subsequently grant to other trading partners. There followed a veritable explosion of *bilateral* trade pacts, with an additional 56 treaties having been signed within 15 years. By 1875, virtually all of Europe was party to a low-tariff zone by dint of a web of agreements that included the linchpin MFN clause.

When the General Agreement on Tariffs and Trade (GATT) went into effect on January 1, 1948, it marked the first of a sequence of *multilateral* trade agreements. The GATT incorporated more than 45,000 tariff concessions by its original 23 signatories, while also providing a broader framework for regulating international trade. Seven subsequent "rounds" of negotiations by these and additional participants led to innumerable further tariff cuts and to the introduction of rules governing various nontariff barriers to trade. The Uruguay Round, which was signed by 123 "contracting parties" and took effect on January 1, 1995, created the World Trade Organization (WTO), while also extending trade rules to many services, harmonizing treatment of intellectual property, and establishing procedures for dispute settlement. By November 30, 2015, the WTO had grown to include 162 members that together conduct more than 96% of world trade.[a] Meanwhile, the multilateral agreement lives side-by-side with 267 different bilateral and regional trade agreements that the WTO reported to be in effect as of February 1, 2016.[b]

This chapter reviews the economics literature that poses the question, why do countries negotiate and accede to international trade agreements? The chapter focuses mostly on the motivation for multilateral agreements, in order to minimize overlap with Chapter 14 by Limão (2016) on preferential agreements. However, Section 5 does cover some literature that addresses the incentives countries have to negotiate bilateral or regional agreements alongside (or instead of) multilateral agreements in a many-country world. Also, this chapter focuses on the broad *purpose* of trade agreements, leaving discussion of their design for Chapter 8 by Bagwell and Staiger (2016).[c] I do consider both

[a] Current members are listed in http://www.wto.org/english/thewto_e/whatis_e/tif_e/org6_e.htm. Trade coverage of WTO members is reported in http://www.wto.org/english/thewto_e/acc_e/cbt_course_e/c1s1p1_e.htm.

[b] See http://www.wto.org/english/tratop_e/region_e/region_e.htm.

[c] This chapter covers some of the same ground as the excellent survey by Maggi (2014) in the *Handbook of International Economics*, vol. 4.

the incentives that large countries have to create trade pacts *de novo* and the incentives that small countries have to sign existing agreements.

The chapter is organized as follows. Section 2 discusses research that sees trade agreements as addressing international externalities that arise in competitive economic environments. It includes the case of both welfare-maximizing governments and politically motivated governments, both in situations with well-functioning markets and with distorted markets. One controversy in the literature concerns whether the sole motivation for trade agreements in competitive markets is to eliminate the temptation governments have to manipulate their terms of trade. I discuss this debate, concluding that it is more a matter of semantics than substance. In Section 3, I review the literature that identifies the different types of international externalities that can arise in imperfectly competitive market environments. These externalities reflect governments' incentives to influence firm location, to shift or extract profits away from foreign monopolists or oligopolists, or to affect imperfect contracting in international outsourcing relationships. I briefly discuss agreements to protect intellectual property that are motivated by externalities in the international innovation process.

In Section 4, I turn to an alternative purpose that has been suggested for trade agreements, namely to aid governments in committing not to intervene in favor of domestic special interests. The commitment motive arises when optimal policies are not time consistent; that is, when governments know they would be tempted to adjust policies *ex post* away from the levels that they prefer *ex ante*. I argue that commitment is unlikely to be the reason that two governments will sit down to negotiate a trade agreement, but that it might very well explain why some countries accede to existing agreements.

Finally, in Section 5, I turn from the purpose of multilateral agreements to that of regional and bilateral agreements. In this section I discuss only research that bears on the motivation that governments have to negotiate preferential agreements in addition to—or instead of—multilateral agreements, in a many-country world. Preferential agreements may serve to promote allocative efficiency among signatories, to improve members' terms of trade vis-à-vis nonmembers, to transfer rents to special interests via trade diversion, or to facilitate a dynamic process of multilateral trade liberalization.

2. INTERNATIONAL EXTERNALITIES FROM UNILATERAL TRADE POLICIES

Trade treaties are a formal expression of intergovernmental cooperation. Governments relinquish their sovereign rights to choose their own trade (and other) policies in exchange for similar concessions by others. Why might a government be willing to compromise its sovereignty? In a word, the answer is *interdependence*. The policies imposed by any government affect the well-being not only of its own citizens, but also of those in other countries. No matter what the objectives of the policy makers—be they

benevolent, autocratic, or politically motivated—each has an interest in the choices made by its trading partners. With unilateral policy choices, governments may fail to take into account the impact of their actions on interests abroad. A trade agreement provides a means to internalize these externalities. Of course, to identify the incentives for concluding a treaty, one must begin by identifying the nature of the potential externalities, that is, by predicting the trade policies that would prevail in the absence of cooperation.

2.1 Welfare-Maximizing Policy Makers

Johnson (1953) was the first to analyze the strategic interdependence between countries' tariff-setting decisions. Johnson conceived of tariffs as being the outcome of a static game played by a pair of welfare-maximizing governments and he proceeded to provide an early application of the then-novel concept of a Nash equilibrium.

Suppose initial that there are two countries and two goods. The goods are competitively produced in each country by firms that have access to strictly convex technologies. Suppose further that aggregate welfare in each country can be represented by a strictly quasiconcave function of the country's aggregate consumption bundle. Let t be the *ad valorem* tariff rate imposed by the home country on its import good. Let t^* be the *ad valorem* rate imposed by the foreign country on its respective import good, which of course is the home country's export good. We can solve for the competitive equilibrium as a function of t and t^* and then write $W(t, t^*)$ and $W^*(t, t^*)$ as the resulting aggregate welfare levels in the home and foreign countries, respectively.

Fig. 1 depicts the best-response functions of the two welfare-maximizing policy makers, with t on the horizontal axis and t^* on the vertical axis. The inverted u-shaped curves such as the one labeled WW represent iso-welfare loci for the home country. These curves peak at the tariff rates that maximize $W(t, t^*)$, given the corresponding values of t^*. The peaks generally fall in the positive quadrant, because the

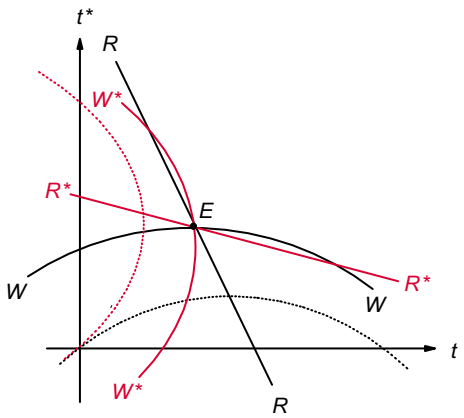

Fig. 1 Nash equilibrium of a tariff-setting game.

Mill-Bickerdike argument for an optimum tariff implies that, for any given foreign policy and economic conditions, home households can gain from a tariff that optimally exploits the country's monopoly power in trade.[d] The curve RR that connects the set of peaks is the home country's best-response function; ie, a function that gives the home country's optimal tariff in response to an arbitrary level of the foreign tariff, t^*. Among any pair of iso-welfare loci for the home country, the curve that lies above the other represents a lower level of home welfare in view of the fact that a departure from the *best* response (a horizontal move to the right or to the left of RR) must result in a welfare loss for this country.

Similarly, the right-parentheses-shaped curves such as W^*W^* represent iso-welfare loci for the foreign country. These curves "peak" in the horizontal direction at the optimal tariffs for the foreign country given the corresponding rates of the home tariff, t. The curve that connects these peaks, R^*R^*, represents the foreign country's best-response function. Among any pair of foreign iso-welfare curves, that to the right represents the lower level of foreign welfare.

As Johnson first noted, a Nash equilibrium occurs at point E, where each government's tariff choice is a best response to that of the other. At this point, neither government can raise aggregate welfare by unilaterally altering its trade policy. The pair of Nash equilibrium tariffs in the Johnson equilibrium are both nonnegative, except possibly in a perverse case such as that described in Footnote d.[e]

Fig. 2 outlines a "lens" to the southwest of point E. At any point in this lens, the aggregate welfare levels in both countries are higher than at point E. In other words, points in this set represent pairs of tariff rates that are Pareto-preferred by the two welfare-maximizing governments to the noncooperative outcome at E. A trade agreement—if one could be negotiated and enforced—that achieves any pair of tariff rates in this lens would be one that both governments prefer to the outcome that occurs without cooperation.

The basis for a trade agreement in this rather simple setting is the negative externality that each government imposes on households in the other country when it imposes its optimal tariff. At point E, a small reduction in the home tariff has virtually no effect on aggregate welfare in the home country, because the optimal tariff just balances on the

[d] Actually, Kemp (1967) shows that, for some values of t^*, the best response by the home country might be an import subsidy; ie, $t < 0$. This can arise only if the foreign offer curve is multi-valued, which in turn requires that the home country's import good is sufficiently inferior in the foreign country's preferences. For ease of exposition, I will neglect this rather obscure possibility.

[e] For the tariffs in the Johnson equilibrium to be nonnegative, it is sufficient that demands in each country can be derived from those of a representative agent and both goods are normal (see Bond, 1990). Dixit (1987) points out that there generally exists another Nash equilibrium, one with prohibitive tariffs in both countries. When one country chooses a prohibitive tariff, it is always a best response for the other to do likewise.

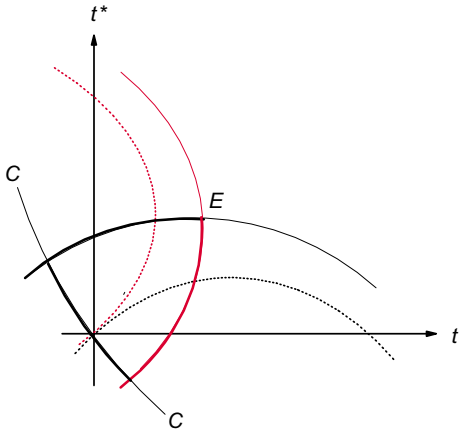

Fig. 2 Efficient agreements.

margin the positive terms-of-trade gain with the negative volume-of-trade loss (see, eg, Dixit, 1985). Any reduction in home welfare generated by a small departure from the best response is second-order small. Meanwhile, a reduction in the home tariff generates a first-order welfare gain for the foreign country inasmuch as it improves that country's terms of trade without generating any distortion of its resource allocation. In other words, each country's optimal tariff is a beggar-thy-neighbor policy that achieves gains for its own citizens (holding constant the other's policy) at the expense of citizens elsewhere and global efficiency. When governments behave unilaterally and noncooperatively, they ignore the harm that their policies impose on citizens outside their borders. This creates the opportunity for mutually beneficial exchange of tariff cuts; points to the southwest of E have approximately the same terms of trade as at E, but higher volumes of trade and a more efficient allocation of the world's resources.

The curve labeled CC in Fig. 2 connects points of tangency between iso-welfare curves of the home and foreign countries. At such points, it is not possible to raise aggregate welfare for one country without reducing it for the other; ie, CC is a locus of Pareto-efficient tariff rates. As Mayer (1981) pointed out, the tariff rates on this curve satisfy

$$(1+t)(1+t^*)=1, \tag{1}$$

because the relative price of the home import good in the home country is $(1+t)p_w$, where p_w is the "world" relative price of this good, while the relative price of the home import good in the foreign country is $p_w/(1+t^*)$.[f] A necessary and sufficient condition

[f] The world price, p_w, is the relative price at a fictitious offshore port where goods are free of all trade taxes. A tariff imposed by the home country raises the internal relative price of this good by a factor $1+t$, whereas a tariff imposed by the foreign country (on the "other" good) reduces the internal relative price there by a factor $1+t^*$.

for global efficiency is that the internal relative prices in the two countries are the same, since this ensures equality worldwide of marginal rates of transformation and marginal rates of substitution.

The figure also shows a portion of this curve in bold. The northwest endpoint of this bold segment corresponds to a pair of tariff rates that yield the home country the same level of aggregate welfare as in the Nash equilibrium at *E*. The southeast endpoint gives the foreign country the same welfare as at *E*. Therefore, all points along the bold portion of *CC* yield a Pareto improvement for the two benevolent governments relative to the noncooperative outcome at *E*. Mayer (1981) concludes that, with efficient and costless bargaining, the governments should agree to some pair of tariffs along *CC* that leave each at least as well off as at *E*. The particulars of the agreement will depend on the bargaining protocol and the countries' relative negotiating proficiency. But, in all cases, both tariff rates are lower under an efficient agreement than at *E*; in other words, the cooperative agreement entails trade liberalization (or promotion) by both countries.[g] The noncooperative tariffs are both positive in the Nash equilibrium as each country attempts to exploit its monopoly power in beggar-thy-neighbor fashion. When the countries cooperate fully, they will choose either free trade, or else a positive tariff in one and a negative tariff in the other in order to achieve the same allocation of resources as with free trade together with a transfer of government revenue.

Some features of the Johnson (1953) and Mayer (1981) analyses are special to the two-good, two-country setting that they consider. Graaf (1949) observed that, with many goods, the vector of a country's optimal tariffs and export taxes could include negative elements, ie, some goods may be subject to import or export subsidies. This conclusion carries over to the Nash equilibrium of a tariff-setting game, and so an efficient trade agreement need not entail a reduction in all trade tax rates. But Bond (1990) shows that if the foreign offer curve is monotone—such that an increase in price reduces foreign imports on average—and if foreign excess demands can be derived from the preferences of a representative agent for whom all goods are normal, then the optimal trade policy for any country generates nonnegative tariff revenue. Under these conditions, a move to universal free trade (which is always efficient with perfect competition and no market distortions) entails an overall reduction of trade taxes and thus a liberalization of trade. Fukushima and Kim (1989) provide conditions, later relaxed by Turunen-Red and Woodland (1991), under which an equiproportionate reduction in all (specific) trade taxes and subsidies (ie, a radial movement toward zero) must raise global welfare in a world with an arbitrary number of goods and subsidies. Under these conditions,

[g] In the figure, the free trade point at the origin generates a Pareto improvement, and all other points of possible agreement involve a positive tariff by one country and an import subsidy by the other. It is possible, however, that one country will fare worse at the free trade point than at *E*. In such circumstances, a Pareto-improving agreement requires an import subsidy in the other country.

multilateral trade liberalization of this sort shifts the world's utility possibility frontier outward, but international transfers of purchasing power might still be needed to ensure Pareto gains for all countries.

2.2 Why a Formal Agreement?

Johnson assumed that, in the absence of any international treaty, policy makers would set their country's tariffs once and for all at the noncooperative levels identified by the static Nash equilibrium. In reality, of course, trade policy decisions are made repeatedly over time. This observation raises the question of whether an actual agreement is needed to achieve the gains from cooperation, or whether a cooperative outcome might be achieved as an equilibrium (without explicit cooperation) in a repeated tariff-setting game. Dixit (1987) discusses the infinitely repeated tariff game involving a pair of welfare-maximizing governments.

The "folk theorem" of repeated games (see Aumann, 1985) ensures that, if the policy makers' discount rates are sufficiently low, any point that is Pareto preferred to point E can be sustained as a subgame-perfect equilibrium in the infinitely repeated game. Each government understands that it is expected to set a given tariff rate repeatedly, such that the pair (t, t^*) is preferred by each to E. Each government further understands that any deviation by the other should invoke it to retaliate by setting its Nash-equilibrium tariff forever afterward. The anticipated punishment is credible, because if one government is expected to play its Nash equilibrium tariff in every period, the best the other can do is to respond similarly, and an indefinite repetition of the static equilibrium is itself an equilibrium of the infinitely repeated game. With this expected punishment, a government that contemplates a deviation will compare the one-period gain from behaving opportunistically against the sustained loss that results from foregoing the Pareto gains from cooperation forever afterward. With a small enough discount rate, the loss must dominate, and so any opportunistic behavior is deterred.[h]

It follows that, if discount rates are low enough, the points along the bold portion of CC in Fig. 2 can be achieved by tacit cooperation in a repeated, noncooperative game, without the need for a formal treaty. So what role does a treaty play? One might think that a formal agreement is needed to sustain high levels of cooperation (ie, points on or close to CC) in situations where the discount rates are not so low, so that governments would otherwise be tempted to behave opportunistically in order to capture short-run gains. However, this answer is not compelling, because if a given cooperative outcome

[h] Bond and Park (2002) show that, if cooperation via repeated play is intended to yield an asymmetric division of surplus relative to the Johnson equilibrium and if the discount rate is not too high, then achieving the Pareto frontier of national welfare levels requires a gradually declining rate of protection in the country that enjoys the lion's share of the gains from cooperation. They offer this observation as an explanation for trade agreements that involve gradual liberalization despite a stationary economic environment.

cannot be sustained by self interest in a repeated game with punishments, it is unclear how a trade agreement would solve the problem. Governments have sovereign rights under international law and there is no higher authority to which they can appeal in case their trade partners behave opportunistically. Rather, a formal international agreement or an international agency can at most threaten poor behavior with (credible) punishments. In this way of thinking, a formal agreement can only achieve those outcomes that are sustainable in an infinitely repeated game.

One role that a formal agreement might play is that of coordination. The infinitely repeated game has many equilibria involving different degrees of cooperation. Any one of them can be sustained as long as both policy makers know what is expected of them and stand ready to punish if the other deviates from its tacitly agreed behavior. How will the governments know and agree upon what behavior is expected of them and what actions constitute deviations that warrant retaliation? After all, there are many tariff rates that are consistent with *some* cooperative behavior along some equilibrium path of the repeated game. And the governments' interests diverge with respect to which of the many equilibria they would like to see played. A formal agreement can be used to achieve a mutual understanding of what is expected, so that misunderstandings do not invoke a return to noncooperative play.

Maggi (1999) explores another potential rationale for a formal institution such as the WTO. His argument explicitly takes into account that multilateral agreements have more than two participants. Maggi argues that an institution such as the WTO can play an informational role by verifying violations of the agreement and by informing third countries when they have occurred. By sharing information about deviations that take place in particular bilateral relationships, the organization can facilitate multilateral punishments for opportunistic behavior that collectively offer greater deterrence than bilateral punishments. Maggi has in mind policy deviations that can be observed directly only by the injured parties (eg, an exporter knows when its market access has been restricted) but not by uninvolved, third parties. As he notes, import tariffs are sufficiently transparent that presumably rate hikes can be observed outside the particular relationship. But other forms of nontariff barriers to trade are more subtle and more difficult for third countries to perceive.

Let us compare the cooperation that can be achieved when a country can observe only the barriers imposed on its own exports with the cooperation that can be achieved when an institution such as the WTO publicizes to all members any violation of the agreement. Consider a stylized trading environment with three countries in which each pair of countries exchanges two goods that are neither produced nor consumed in the third country. In this setting, no country is directly harmed by opportunistic behavior on the part of its trade partners in its bilateral relations with the third country. Nevertheless, in some circumstances, every country can benefit from an agreement that calls for multilateral punishment for any deviation. Such punishments would prescribe reversion

to Nash equilibrium tariffs in bilateral relations with the violating country not only by the country that is directly harmed by its actions, but also by the third country.

The key consideration for whether multilateral punishments facilitate greater cooperation than do bilateral punishments alone is the existence of imbalances of power in the bilateral trading relationships. Maggi shows this point by first examining a case in which all countries are symmetric and each has balanced trade with both of its trading partners. In this case, multilateral punishments are more severe than bilateral punishments, but the temptation to behave opportunistically also is greater. The set of outcomes that can be sustained by multilateral punishments is the same as that with bilateral punishments. On the other hand, if each of the symmetric countries is a net exporter in one of its bilateral relationships and a net importer in the other and if each can apply import tariffs but not export taxes, then each country will be more powerful than its partner in the relationship in which it imports. The greater power reflects the fact that a net importer in a given bilateral relationship stands to lose less from a reversion to Nash equilibrium tariffs than does the net exporter. Then, a regime that publicizes violations and calls for punishments by third parties is more effective in deterring opportunistic behavior than one that entails punishments by only directly injured parties. As Maggi notes, the WTO rules as currently construed do not allow for explicit third-party punishments, but it is conceivable that such punishments occur implicitly by a loss of goodwill. In any case, the paper shows that a multilateral institution that monitors trade behaviors and publicizes violations *could* play a role in sustaining greater cooperation and freer trade.[i]

2.3 Trade Agreements among Politically Minded Governments

When welfare-maximizing governments can act opportunistically in an effort to improve their country's terms of trade, the resulting international externalities open the potential for mutually advantageous trade agreements. But governments' trade policies seem to be motivated as much or more by distributional concerns as by concerns about aggregate welfare. The assumption that governments maximize aggregate welfare is convenient but rather unrealistic in a world with asymmetrically informed voters and with campaign spending funded by contributions from special interest groups. Policy makers who are interested in their own political success might be tempted to choose protectionist policies not (or not only) to manipulate the terms of trade, but rather (or also) to redistribute income to swing voters in the electorate or to groups that offer campaign support. With perhaps limited interest in aggregate welfare, would these politicians be willing to

[i] Park (2011) illustrates another potential role of a formal organization such as the WTO in a world with imperfectly observed trade barriers. In his model, a third party can publicize private signals and thereby initiate a punishment phase after a violation of the agreement. This enhances the cooperation that can be achieved in repeated play.

negotiate with foreign counterparts? By doing so, wouldn't they curtail their own use of a valuable instrument for currying political favor?

These questions cry out for a political-economic theory of trade agreements. Hillman and Moser (1995) provided a first pass at such a theory (see also Hillman et al., 1995). They considered a model with two countries, two goods, and two sector-specific factors of production, one that derives income only from a country's import-competing industry and the other that derives income only from the country's export industry. The governments have objective functions that are expressed in reduced form; each government derives "political support" from the specific factors in its own polity, with overall support an increasing but concave function of support from each group. The support, in turn, reflects the real income of a group of specific factor owners. In short, they posit government objective functions $G\left(I_x, I_y\right)$ and $G^*\left(I_x^*, I_y^*\right)$ for the home and foreign governments, where I_j and I_j^* are the real incomes of the specific factors used to produce good j in the home and foreign countries, for $j = x,y$. The first partials of G and G^* are assumed positive while the second partials are assumed negative. Each government can set the level of a single policy variable, t or t^* respectively, which is the tariff levied on its own imports. Hillman and Moser rule out exports subsidies with casual reference to countervailing duty laws that would neutralize their effects; however, they offer no explanation for why export subsidies are subject to countervail but policy makers are free to choose import tariffs without institutional constraints.

From here, the story proceeds much as in Johnson (1953) and Mayer (1981). In the absence of an agreement, each government sets its tariff to maximize its political support taking the other government's policy choice as given. The noncooperative outcome is the Nash equilibrium pair of tariffs, much like in Fig. 1. The home tariff raises the real income of the specific factor in the home import-competing industry at the expense of the real income of the specific factor in the home export industry. Each government's choice balances the marginal effects on its own political support, but neglects entirely the adverse effect that the tariff has on the other country's export interests (and the positive effect that it has on the other country's import-competing interests). The Nash equilibrium occurs at a peak of an inverted-u shaped indifference curve for the home government drawn in the space of the two tariffs. It also occurs at the peak of a right-parenthesis shaped indifference curve for the foreign government. And, as in Fig. 2, there is a lens to the southwest of the Nash equilibrium such that, for all pairs of tariff rates in the lens, each government could achieve greater political support from its own specific factors than at the noncooperative equilibrium.

Several observations are in order. First, the motivation for a trade agreement again reflects international externalities and interdependence. Without the agreement, each government neglects the impact that its policy choices have on the *political welfare* of the opposite government. But the motivation for the agreement is not necessarily based on an improvement in social welfare, nor are the agreed-upon tariffs likely to be

economically efficient, as they are in Mayer (1981). Rather, an agreement allows the governments to promote "political efficiency" by an exchange of "market access." By reducing its own tariff slightly, the home government has a negligible effect on its own political support, since the Nash tariff was chosen to balance the offsetting effects on domestic interests (including the beneficiaries of any tariff revenue). But the foreign export interests benefit from the improved market access and, although the foreign import-competing interests are harmed, their income drop is buffeted by an improvement in the foreign terms of trade. On the margin, a reduction in the home tariff improves political support for the foreign government at only negligible cost to the home government. The foreign government can reciprocate by cutting its own tariff, thereby improving market access for home exporters and raising the home government's support from its export industry.

Grossman and Helpman (1995b) attempt to provide better microfoundations for the governments' choices in a model with campaign giving by special interest groups. There are $n + 1$ industries, one of which produces the numeraire good from labor alone while the remaining n industries each produce with labor and a factor of production that is specific to the industry. Household welfare is a quasilinear function of consumption of the numeraire good and utility from consuming the various nonnumeraire goods. The home government maximizes a political objective function of the form $G(W, C) = aW + C$, where W is aggregate welfare in the home country and C is aggregate contributions from all special interest groups.[j] Meanwhile, the foreign government maximizes $G^*(W^*, C^*)$, with variables defined analogously. The politics are treated as a two-stage game in which the interest groups offer contribution schedules to their own government in the first stage and the governments choose their trade-policy vectors (import and export taxes and subsidies) in the second stage. Interest groups are assumed to represent the sector-specific factors of production in a subset of industries, namely those industries that manage to overcome free-rider problems and become politically organized. Each contribution schedule links a group's campaign contribution to the government's multidimensional policy choice. In the competition for influence that arises when the governments behave noncooperatively, the interest-group contributions are assumed to be a function of the own government's policy vector. If a cooperative trade agreement is anticipated, then an interest group can also tie its contribution to the other government's choices in the hope of influencing its own government's posture in the trade negotiations. Finally, the "trade war" outcome is taken to be the Nash equilibrium of a second-stage noncooperative game between the governments, whereas the "trade talks" outcome is the Nash bargaining solution to the second stage.

[j] This objective function can be derived from a two-party model of elections with impressionable and nonimpressionable voters, where the former can be influenced by campaign spending and the latter vote purely based on self-interest (see Grossman and Helpman, 1996).

The equilibrium policies in the trade war satisfy

$$t_i - 1 = -\left(\frac{I_i - \alpha_L}{a + \alpha_L}\right)\frac{x_i}{m_i}\frac{1}{e_i} + \frac{1}{e_i^*}, \tag{2}$$

$$t_i^* - 1 = -\left(\frac{I_i^* - \alpha_L^*}{a^* + \alpha_L^*}\right)\frac{x_i^*}{m_i^*}\frac{1}{e_i^*} + \frac{1}{e_i}, \tag{3}$$

where t_i is the import tax at home if good i is imported or the export subsidy if it is exported, I_j is an indicator variable that takes on the value one if industry j is politically organized or zero if it is not, α_L is the fraction of the total population that belongs to some organized interest group or another, x_i is output in home industry i, m_i is home imports (possibly negative), e_i is the elasticity of import demand (negative) or export supply (positive), as the case may be, and symbols with asterisks are analogous variables that apply to the foreign country. Grossman and Helpman interpret the first-expression on the right-hand side of each equation as the political motive for noncooperative trade policy; the competition for influence via contributions induces each government to support those local industries that are politically organized, that have a high ratio of output to trade and that are relatively immune to deadweight loss, as reflected in a small local trade elasticity. The second term in each equation is the standard terms-of-trade motive for positive tariffs and export taxes. The inverse trade elasticity in the partner country reflects a country's monopoly power in trade. Notice that, if no industry is politically organized ($I_i = 0$ and $I_i^* = 0$ for all i) or if the governments give negligible weight to campaign contributions relative to aggregate welfare ($a = a^* = \infty$) then the model predicts the vector of Nash-equilibrium optimal tariffs, as in Johnson (1953).

Starting from the trade policies indicated in (2) and (3), the two governments have an incentive to negotiate a trade agreement for much the same reason as in Hillman and Moser (1995) or, for that matter, Mayer (1981). In the Nash equilibrium, the unilateral policies chosen by each government confer externalities on the other. The pair of noncooperative tariff vectors is economically inefficient, but that is hardly surprising, because the governments are only partly concerned with aggregate welfare. What motivates their cooperation is rather the lack of *political* efficiency. By trading policy concessions, the governments can find an agreement such that each achieves a higher weighted sum of welfare and campaign contributions as compared to the outcome at the Nash equilibrium.

When the organized interest groups design their contribution schedules in anticipation of an efficient trade negotiation (one that will yield a Pareto-efficient outcome for the two governments with respect to their objectives, G and G^*), the negotiated trade policies satisfy

$$t_i - t_{i^*} = \frac{I_i - \alpha_L}{a + \alpha_L}\frac{x_i}{m_i}\frac{1}{e_i} - \frac{I_i^* - \alpha_L^*}{a^* + \alpha_L^*}\frac{x_i^*}{m_i^*}\frac{1}{e_i^*}. \tag{4}$$

The trade talks do not determine the levels of the import and export taxes or subsidies in each country in a given industry, but only the difference between the importing country's tariff and the exporting countries export subsidy, as reflected in (4). This is so, because an equal change in the import tariff rate and the export subsidy rate would not affect internal prices in either country, and so it would not affect consumer surplus, producer surplus, or deadweight loss. An equal rate change only affects the distribution of tax revenues (positive or negative) captured by the two governments, and such a change in one industry can always be compensated by an offsetting rate change in another industry. The relative bargaining power of the two governments determines the aggregate trade tax revenue that each collects, but not the composition of that revenue by source industry.[k]

The net effect of the negotiated trade policies applied on any good reflects the difference in the political strength of the industry's special interests in the two countries. This difference reflects whether or not the specific factors in the industry are politically organized, whether price distortions in the industry cause a great or small amount of deadweight loss, and how willing the industry's own government is to sacrifice aggregate welfare in exchange for campaign contributions. The negotiated policies preserve the political motive that each government has to cater to its local interests, while accounting for the externality that its policies impose on the other. The components of (2) and (3) that reflect the terms-of-trade motive for unilateral policy intervention do not appear in (4), because the exercise of monopoly power via an optimal tariff or export tax enhances aggregate welfare in one country at the expense of that in the other, and such beggar-thy neighbor policies have no place in an efficient bargain between governments, be they economically motivated or otherwise.

2.4 Is it All About the Terms of Trade?

To recap, Mayer (and subsequent authors) identified a terms-of-trade externality that arises when governments unilaterally pursue the maximization of constituents' aggregate welfare. This externality creates an opportunity for mutual gain from cooperation. Hillman and Moser (and others) argued that governments are motivated not only by economic concerns but also—or especially—by political concerns that reflect the distribution of income. In their view, a terms-of-trade externality is not the primary motivation for trade agreements. Rather, governments seek agreements in order to internalize a market-access externality. They argued that when institutional constraints preclude direct government support for export industries (for some unspecified reason), the governments can reap mutual political gains by jointly opening their markets, thereby capturing the political rewards that their respective export interests have to offer.

[k] The indeterminacy arises for much the same reason that all home and foreign tariff rates that satisfy (1) are efficient when tariffs are set by two welfare-maximizing governments.

So, is it a terms-of-trade externality that motivates trade agreements, or something else? Bagwell and Staiger (1999, 2002) make a strong claim that, even in a world with politically motivated governments, trade agreements are all about governments overcoming their temptation to manipulate the terms of trade. But they also claim that, when addressing the terms-of-trade externality, there is a sense in which governments are inexorably led to grant better market access. In short, they claim that there is no meaningful difference between a terms-of-trade externality and a market-access externality.

Bagwell and Staiger begin by specifying two governments' objective functions in reduced form. They write the governments' political objectives as $G(p,p_w)$ and $G^*(p^*,p_w)$, where p and p^* again are the internal (or "domestic") relative prices that producers and consumers face in the home and foreign countries, respectively, and p_w is the world relative price of the good imported by the home country, ie, the international terms of trade. The wedge between p and p_w reflects the home country's trade policies, while that between p^* and p_w reflects the foreign country's trade policies. A special case of these preferences arises when the governments maximize aggregate welfare, since all factor prices (and thus market incomes) and all consumption decisions depend on domestic prices, while a government's revenue depends on the difference between domestic and world prices. The Grossman and Helpman (1995b) specification yields a similar reduced form, once the contributions C and C^* that enter directly into the governments' objectives are replaced by the equilibrium first-stage contribution offers of the special interest groups. Indeed, the Bagwell-Staiger reduced form represents government preferences in a wide range of political-economy models with perfect competition, because in all such models the domestic prices determine the distribution of factor income and consumer surplus, while the wedges between the world price and the internal relative prices determine the intergovernmental distribution of tax revenues.

Consider the effects of a rise in the world relative price of the home country's import good, p_w, holding local prices in both countries constant. This will not alter factor incomes in either country, nor will it change the prices faced by any consumers. The only effect is a reduction in the home government's revenues from its trade policies and an equal increase in government revenues in the foreign country. Plausibly, this event should be viewed favorably by the foreign government and unfavorably by the home government. Accordingly, Bagwell and Staiger posit that $G^*_{p_w} > 0$ and $G_{p_w} < 0$. Concerning the economic environment, they assume only that the Lerner paradox and the Metzler paradox do not arise.[1]

[1] The Lerner paradox applies when an increase in a countries tariff leads to an increase in the world price of its imports, hence a deterioration in the terms of trade. This can happen only if the government spends a disproportionate share of its tariff revenues on the import good. The Metlzer paradox applies when a tariff generates such a large improvement in an importing country's terms of trade that the domestic relative price of the import good actually falls there. For this to happen in the home country, the foreign excess demand for its import good must be sufficiently inelastic.

The market-clearing world price depends on the trade policies chosen by the two governments. Given the trade policies and the market-clearing world price, the pair of domestic prices are determined mechanically. In such circumstances, one might as well imagine that the governments choose their domestic prices directly and the respective trade policies are determined residually, once p_w is determined. Writing p_w as a function of p and p^* to reflect the dependence of equilibrium prices on trade policies, the first-order conditions for the governments' optimal choices of domestic prices at the Nash equilibrium imply

$$G_p(p, p_w) + G_{p_w}(p, p_w)\frac{\partial p_w(p, p^*)}{\partial p} = 0, \tag{5}$$

$$G^*_{p^*}(p^*, p_w) + G^*_{p_w}(p^*, p_w)\frac{\partial p_w(p, p^*)}{\partial p^*} = 0. \tag{6}$$

As before, this outcome is politically inefficient, inasmuch as both G and G^* could be increased by appropriate adjustments in the domestic prices away from those that prevail in the noncooperative equilibrium. This can be seen from the fact that the iso-utility curve for the home policy maker is perpendicular to that for the foreign policy maker at the Nash equilibrium, whereas political efficiency calls for tangency between these two curves.

Next, Bagwell and Staiger designate a benchmark that they term the "politically optimal tariffs." They define these tariffs implicitly as the wedges that arise between domestic and world prices when the internal prices satisfy

$$G_p(p, p_w) = 0, \tag{7}$$

$$G^*_{p^*}(p^*, p_w) = 0. \tag{8}$$

How should we understand these politically optimal tariffs? They are the tariffs the governments would choose if, for some unspecified reason, each were to ignore the effect on its political objective that results from changes in the terms of trade, and if each expected the other to do similarly; that is, if both acted as if $G_{p_w} \equiv G^*_{p_w} \equiv 0$. Were the governments hypothetically to behave in such a manner, there would be no scope for them to negotiate a trade agreement in pursuit of mutual political gains. To see why, consider the effect of a change in the foreign tariff, t^*, on the political welfare of the home government, starting from a pair of local prices that satisfy (7) and (8). A change in t^* would induce an equilibrium adjustment of the home price, p, but to first order this alone would neither benefit nor harm the home government, by (7). The change in t^* would also alter the terms of trade. But the induced change in p_w would generate a pure transfer of welfare between the two governments and so it could not be a basis for mutual gain. In short, if each government were to set its tariff equal to the political optimal, there would be nothing further on which they could agree.

The politically optimal tariffs refer to a thought experiment; Bagwell and Staiger view them as purely hypothetical and do not suggest that governments would have reason to behave this way in any set of circumstances. Since there is no behavioral justification for these tariffs, it is difficult to evaluate whether they provide a sensible benchmark to use in ascribing motives to the governments. Are terms-of-trade externalities an apt description of the motivation for trade negotiations, because no negotiations would take place if governments simply ignored their welfare effects? Perhaps, but without an explanation about why and in what circumstances the governments might do so, this is largely a matter of semantics. The semantic nature of the debate becomes clear when Bagwell and Staiger are asked, Is there no market-access externality such as the one identified by Hillman and Moser? Bagwell and Staiger (2002) in fact regard the terms-of-trade externality and the market-access externality as being two sides of the same coin. They associate the market access that a country affords to its trade partner at world prices p_w with the import demand function, for example, $M(p,p_w)$ for the home country. Now, starting from a noncooperative equilibrium, suppose the home government were to contract its market access marginally by increasing the domestic relative price of the import good. In the absence of any Lerner paradox, this would have to decrease the world price, p_w. At the Nash equilibrium value of p^*, the effect on the political objective of the foreign government is unambiguously negative. So, a contraction of market access imposes a negative externality abroad, and a mutually beneficial trade agreement requires each government to shift out its import demand function, ie, to grant expanded market access to its trade partner.[m]

The important conclusion is that, with perfect competition in world markets, for a wide range of political objective functions governments have an incentive to negotiate a trade agreement that expands world trade. This is the essence of the Bagwell-Staiger insight, more so than whether we describe the agreement as being one in which each country offers a better world price for the initial volume of trade (improved terms of trade) or as one in which each offers greater import demand at the initial world prices (improved market access).

2.5 Market Distortions and Corrective Policies

We have seen that governments have an incentive to negotiate a trade agreement even when their objectives include the support of local special interests. The politically efficient agreement induces governments to take into account the international externalities from their policy decisions and generates improved market access for exporters in both

[m] Bagwell and Staiger (2002) make a global argument. They say that an agreement provides additional market access for a country if its trade partner's import demand function shifts out for *some* world prices. They prove that every mutually beneficial trade agreement must secure additional market access in this sense for both countries.

countries. What if local markets do not function perfectly and governments need to use the policy instruments at their disposal to overcome market inefficiencies? Or what if governments have idiosyncratic preferences for certain types of market outcomes, be they economic or noneconomic objectives? Trade policies are part of the arsenal that policy makers have to correct market failures and to generate preferred outcomes. And other policies besides trade policies—even if "legitimately" motivated—can generate externalities for other governments. How should we think about the purpose of trade agreements in a setting with market distortions, noneconomic objectives, and the possibility that governments might invoke all manner of domestic economic policies?

Bagwell and Staiger (2001, 2002, Chapter 8) offer an elegant answer to this question. Let us now write the objectives of the home and foreign governments as $G(p, p_w, s)$ and $G^*(p^*, p_w, s^*)$, respectively, where p and p^* are the relative prices faced by home and foreign consumers, and s and s^* represent the levels of two arbitrary domestic policies set by the home and foreign governments.[n] The domestic policies might be subsidies to local production of one good or the other, in which case the prices received by producers would differ from those paid by consumers. Or the policies might represent interventions in local factor markets, either in the form of taxes or subsidies for factor employment or factor supply, or else measures of direct regulation of factor usage. The policies could also represent limits on pollution, standards for product quality or safety, restrictions on the use of underage labor, minimum wage rates, etc. Meanwhile $G(\cdot)$ and $G^*(\cdot)$ could incorporate political and distributional objectives of the government, as before, as well as their idiosyncratic preferences over market outcomes, and the government objectives could reflect the extent to which market failings that can be traced to factor-market rigidities or product-market externalities. We rule out only the exercise of monopoly power by firms or households (which we will discuss later) and international nonpecuniary externalities that arise, for example, from cross-border pollution, induced climate change, or local concerns about market outcomes in the other country.[o] As before, we assume away conditions that give rise to the Lerner and Metzler paradoxes, and we assume that the Marshall-Lerner conditions for market stability (ie, that a rise in the relative price of a good causes world excess for the good to fall) are satisfied.

Now, governments set their trade policies and their "other" policies. Equivalently, they choose their domestic prices p and p^* along with s and s^*; in this conceptualization, the trade policies are determined residually. The choices of (p, s) and (p^*, s^*) determine

[n] This representation of government objectives is slightly different from the one used by Bagwell and Staiger (2001, 2002), but it captures the essence of their arguments about both standards and production subsidies provided that the local prices p and p^* are interpreted to be consumer prices and the governments do not have use of consumption taxes and subsidies. Bagwell and Staiger (2006) analyze WTO treatment of production subsidies in a setting that allows for domestic consumption policies.

[o] An example of a local concern for market outcomes in other countries arises when citizens disdain the use of child labor, even if it occurs outside the borders of their country.

home and foreign supplies and demands, given p_w, and so the world relative price must settle at the level that clears the world market. In other words, the equilibrium world price can be written as $p_w(p, s, p^*, s^*)$. Each government has two first-order conditions that guide its unilateral choice of trade and domestic policies; for example, for the home government, these are

$$G_p(p, p_w, s) + G_{p_w}(p, p_w, s)\frac{\partial p_w(p, s, p^*, s^*)}{\partial p} = 0, \tag{9}$$

$$G_s(p, p_w, s) + G_{p_w}(p, p_w, s)\frac{\partial p_w(p, s, p^*, s^*)}{\partial s} = 0. \tag{10}$$

On the other hand, global political efficiency requires that the home government set p and s to maximize $G(p, p_w, s) + \lambda G^*(p^*, p_w, s^*)$, where λ is the relative weight attached to the foreign government's objective in the efficient agreement, an indicator of its relative negotiating ability. Therefore, the globally efficient home policies satisfy

$$G_p(p, p_w, s) + \left[G_{p_w}(p, p_w, s) + \lambda G^*_{p_w}(p^*, p_w, s^*)\right]\frac{\partial p_w(p, s, p^*, s^*)}{\partial p} = 0, \tag{11}$$

$$G_s(p, p_w, s) + \left[G_{p_w}(p, p_w, s) + \lambda G^*_{p_w}(p^*, p_w, s^*)\right]\frac{\partial p_w(p, s, p^*, s^*)}{\partial s} = 0. \tag{12}$$

Clearly, the unilateral policies that the home government would choose in a noncooperative setting, as given by (9) and (10), do not satisfy the requirements for political efficiency in (11) and (12). The inefficiency of the Nash equilibrium arises again from an international externality, as reflected in the neglected terms $\lambda G^*_{p_w}(p^*, p_w, s^*)[\partial p_w(p, s, p^*, s^*)/\partial p]$ and $\lambda G^*_{p_w}(p^*, p_w, s^*)[\partial p_w(p, s, p^*, s^*)/\partial s]$. And clearly the externalities travel to the foreign country via world prices; when the home government acts unilaterally, it ignores the effect that its choices have on world market prices and thereby on the objectives of the foreign government. The foreign government can react by setting its trade and domestic policies to further its own political objectives, to pursue its idiosyncratic preferences over market outcomes, and to offset its local market distortions, but ultimately the joint choices in a Nash equilibrium leave scope for mutual gain.

Bagwell and Staiger make a further observation. Suppose the governments negotiate over market access, as reflected in the location of each country's import demand schedule or, equivalently, as ultimately reflected in the world market price. Once they agree on a value of p_w, the governments can (and should) be left to determine their own mix of trade and domestic policies. That is, suppose the governments conclude that $p_w(p, s, p^*, s^*)$ should be equal to $\bar{p}_w$ under an international agreement. Then unilateral policy choices by the home government *subject to* $p_w(p, s, p^*, s^*) = \bar{p}_w$ will satisfy

$$G_p(p, \bar{p}_w, s) + \gamma\frac{\partial p_w(p, s, p^*, s^*)}{\partial p} = 0, \tag{13}$$

$$G_s(p,\bar{p}_w,s)+\gamma\frac{\partial p_w(p,s,p^*,s^*)}{\partial s}=0, \tag{14}$$

where γ is the Lagrange multiplier on the terms-of trade (or market-access) constraint. By judicious choice of $\bar{p}_w$, γ can be made to equal $G_{p_w}(p,\bar{p}_w,s)+\lambda G^*_{p_w}(p^*,\bar{p}_w,s^*)$, and then the conditions for global political efficiency in (11) and (12) will be satisfied. In short, the international agreement can leave the details about policy mix to each country, as long as each provides the appropriate degree of market access.

Lest this tolerance for national sovereignty be misinterpreted, Bagwell and Staiger emphasize its limitation. Given an agreement about market access, their analysis shows that each government can be allowed to choose its own mix of trade and domestic policies. But this is not the same thing as saying that the governments can come to an agreement about trade policies while leaving each one free to choose whatever domestic policies it prefers. If the governments agree only about trade policies while making no commitments about their domestic policies, the home government, for example, will have an incentive to distort its choice of s to satisfy a condition like (10). That is, it will choose its policy not only to correct the domestic market failure or to achieve its noneconomic objective, but also with an eyes towards the effect on the terms of trade. In so doing, it will partly undo any concession about providing market access that is implied by its restricted choice of trade policy. An agreement that restricts choices of trade policies may be better than no agreement at all, but without a commitment to prescribed levels of market access, the resulting outcomes will not be politically efficient.

2.6 Critiques of the Theory

I have presented a theory of trade agreements founded on the potential gains from cooperation among policy makers that have different constituencies and serve different interests. On the economic side, the theory assumes convex technologies and perfect competition, but it allows for various types of domestic market failures. On the political side, it assumes self-interested policy makers who may have purely benevolent motivations or political motivations and may pursue noneconomic objectives. The theory rests on the sole premise that a noncooperative Nash equilibrium is inefficient for the two sets of policy makers, because unilateral policy choices generate international externalities. The theory presumes governments negotiate trade agreements in order to internalize these externalities.

The theory has been criticized on various grounds, most notably by Ethier (2007, 2013) and Regan (2006, 2015). The criticisms take three, interrelated forms. First, whereas the theory highlights the role of terms-of-trade manipulation in motivating trade negotiations, the rhetoric surrounding actual negotiations makes virtually no reference to the terms of trade and focuses instead on governments' aims to reduce "protectionism." Second, whereas the theory assumes that governments value tariff

revenue, the discussions surrounding trade agreements do not highlight such a concern. Finally, the theory takes the set of political actors that lobby for trade policy as given, whereas it seems that the prospect of a trade negotiation mobilizes export interests that might otherwise remain on the political sideline. In this section, I will describe these criticisms in more detail and outline potential responses.

Recall that Bagwell and Staiger (1999, 2002) identify terms-of-trade manipulation as *the problem* that trade agreements are meant to solve. They argue that, in a hypothetical situation in which governments act unilaterally but behave as if they see no benefit (political or otherwise) from any terms-of-trade improvement, the outcome would be politically efficient and admit no potential gains from cooperation. For this reason, they point to governmental concerns about terms-of-trade manipulation as the motivating force behind trade negotiations. Grossman and Helpman (1995b) make no such claim, but they do observe that the politically efficient trade policies strip away the optimum-tariff components of the noncooperative policies while preserving the components that reflect the difference in the political strengths of the special interests in the import-competing and exporting industries. One might say that the trade talks in Grossman and Helpman also are necessary only to rid the world economy of terms-of-trade manipulation.

Regan (2015) describes a "practitioner's understanding" of trade agreements, based loosely on his reading of historical accounts by participants in trade negotiations. Trade practitioners, he asserts, do not mention any concern about the terms of trade. The words never appear in their accounts of their bargaining experiences. Rather, the practitioners tell a story of reducing "protectionism," which he defines as "any unilateral trade policy that restricts imports to get political support for the government from import-competing producers." In his view, policy makers are willing to forego protectionism in exchange for similar concessions by their partners, because the political support each can attract from export interests exceeds what each stands to lose from its import-competing interests.[p] Regan further notes that a reduction of protectionism in this sense can play no role in motivating trade negotiations in Bagwell and Staiger (1999, 2002) or Grossman and Helpman (1995b), because protectionism remains unabated under the politically efficient agreement.

Next, Regan and Ethier (2013) point to tariff revenues. These are the other side of the coin from terms-of-trade improvements. When a trade tax causes the world price of a country's imports to fall or the world price of its exports to rise, and domestic prices move in the opposite direction in the absence of any Lerner paradox, the terms-of-trade gain is realized in the revenues collected by the government. Were the government not to care about these revenues, it also could not value the change in world prices. Put differently, a

[p] Regan (2015) cites Hillman et al. (1995) as an example of what he has in mind, although, as noted above, the welfare effects of term-of-trade changes are essential to the Hillman et al. demonstration that the political support gains from export interests exceeds the loss from import-competing interests.

policy that raises domestic prices while collecting no revenue (such as a "voluntary export restraint") can only worsen the terms of trade, not improve them. But Regan and Ethier see no evidence in their reading of practitioners' accounts of any negotiator's interest in the revenues generated by import and export taxes.

Finally, Regan and Ethier argue that the key to understanding trade agreements is the boon they provide to export interests. As they recognize themselves, their preferred account relies on an assumption that export interests have no way to capture policy rents in a world of unilateral policy setting in a manner akin to what import-competing interests manage to do. Import-competing interests exchange political support in the form of campaign contributions or otherwise for private benefits from protection. Export interests, they claim, cannot play this game. According to these authors, a trade negotiation is fundamentally about empowering export interests to counter protectionist forces.[q]

Let us evaluate these criticisms in terms of what they imply about the modeling of trade agreements. The fact that practitioners do not mention a concern about the terms of trade *per se* is not determinative about their role in the workings of formal models of international externalities. The terms of trade are instrumental in these models, not a direct policy goal. The Grossman-Helpman model assumes that policy makers are concerned with the welfare of their constituents and with campaign contributions, and that the latter are offered by interest groups to further their economic interests. The politicians need only recognize that a lower price of imports contributes to the welfare of consumers and that a higher price of exports adds to the income of (some) producers in order to behave "as if" they are concerned about the terms of trade. What policy maker would not prefer that imports be cheaper and that exports command higher prices, *all else the same?* This logic underlies the reduced-form specification of preferences in Bagwell and Staiger. The legitimacy of their rather general formulation of government preferences does not rest on whether practitioners use the words "terms of trade" to describe their bargaining goals.

The modeling of trade agreements by Grossman and Helpman and by Bagwell and Staiger does rely on the assumption that governments care about the fiscal consequences of their policy choices. To see why, consider a Grossman-Helpman world in which a negligible fraction of the population is represented by an interest group ($\alpha_L = 0$) and in which the government for some reason places no value of tax revenue, positive or negative. In such a setting, the equilibrium contribution schedules would induce the government to act as if it were maximizing $G = \sum(1 + aI_i)\Pi_i(p_i) + S_i(p_i)$, where a is the weight that the policy maker attaches to campaign contributions relative to producer-plus-consumer surplus (but with no revenue term), $\Pi_i(p_i)$ and $S_i(p_i)$ are producer and

[q] Neither Ethier nor Regan provides a logically consistent formal model of this process. Similar informal statements about this role of trade negotiations can be found in the writings of some trade-focused political scientists and legal scholars, such as, Hudec (1993), Destler (2005), and Pauwelyn (2008).

consumer surplus from the consumption of good i when the domestic price is p_i, and I_i again is a dichotomous variable that indicates whether industry i is politically organized or not. Note that $\partial G/\partial p_i = (1 + aI_i)x_i(p_i) - c_i(p_i)$, where x_i and c_i are output and consumption, respectively, and that $\partial^2 G/\partial p_i^2 = (1 + aI_i)x_i'(p) - c_i'(p) > 0$. Evidently, there can be no interior solution in this setting; each government tries to set as high a domestic price as possible, which it can do by raising toward infinity its export subsidy. Needless to say, this is not a sensible prediction of the model and does not provide a benchmark against which to consider governments' incentives for negotiating a trade agreement. In a political-economy setting in which governments literally do not care about the fiscal implications of their policies, their search for contributions or other forms of political support leads them to subsidize lavishly whichever group has more at stake in the policy decision and a better means of playing the political game.

Regan and Ethier recognize that an assumption of no governmental concern for fiscal deficits generates nonsensical predictions, so they argue instead for a formulation in which the government places no value on any positive revenues generated by trade taxes, but bears a prohibitive cost of making public outlays for import or export subsidies. Regan in particular argues that such an assumption—admittedly inconsistent with the belief that policy makers are "rational actors"—is descriptively realistic. But, even if we accept such irrationality, it is not clear how it would rescue the situation. Suppose a government acts unilaterally to maximize $G = \sum(1 + aI_i)\Pi_i(p_i) + S_i(p_i)$, but does so under the constraint that $R_i \equiv (p_i - p_{wi})[c_i(p_i) - x_i(p_i)] \geq 0$ for all i, where p_{wi} is the world price of good i and therefore R_i is the revenues collected from trade taxes on good i. The constraint says that the revenues from any trade tax cannot be negative, ie, that subsidies to trade are impossible. There are two cases to consider. Suppose first that $x_i(p_{wi}) > c_i(p_{wi})$, ie, that the country would export good i under free trade. In this case, $\partial G/\partial p_i = (1 + aI_i)x_i(p_i) - c_i(p_i) > 0$, so the government would like to raise the domestic price to benefit the exporters, but it cannot do so due to the fact that subsidies are forbidden. Alternatively, suppose that $c_i(p_{wi}) > x_i(p_{wi})$, ie, that the country would import good i under free trade. If $c_i(p_{wi}) > (1 + aI_i)x_i(p_{wi})$, it would wish to reduce the domestic price, but cannot do so without subsidizing imports. If $c_i(p_{wi}) < (1 + aI_i)x_i(p_{wi})$, it will instead wish to raise the domestic price, and will do so until imports are eliminated. So a government that acts unilaterally and that does not care about positive revenues from trade taxes but stays clear of trade subsidies has either zero or prohibitive trade taxes in every sector. This hardly seems a reasonable depiction of the counterfactual outcome in the absence of a trade agreement.

More promising, perhaps, is the Ethier-Regan suggestion that trade negotiations empower export interests that otherwise would be shut out of the political process. They offer no convincing reason why import-competing interests are able to organize themselves and lobby for protection in a noncooperative setting whereas export interests cannot do so. Nor do they offer any explanation as to why the situation suddenly changes

when trade negotiations are contemplated.[r] But Grossman and Helpman (1995b) are silent about which industries are politically organized and which are not, and Bagwell and Staiger (2002) are silent about the stability of their reduced-form government preferences. It is certainly possible that new interest groups become organized when trade negotiations are active, so that the induced government preferences change.[s] Future research might uncover reasons why the onset of trade negotiations encourages the participation of export interests in lobbying that are otherwise unable to influence trade policy outcomes. If so, then policy makers might well wish to enter into trade agreements as a way of to mobilize these interests and thereby counter the forces of protectionism.[t]

Regan (2015) and others have offered another critique of the competitive theory of trade agreements, one that focuses on its predictions rather than its assumptions. The critique concerns the model's implications for how export subsidies ought to be treated in trade agreements and the reality of how modern agreements actually treat such subsidies. Indeed, (Bagwell and Staiger, 2002, Chapter 10) have been puzzled by this same issue.

Consider again the trade policies that result when the home and foreign governments choose t and t^* to maximize $G(W,C) = aW + C$ and $G^*(W^*, C^*) = a^*W^* + C^*$, respectively, in a noncooperative equilibrium of the Grossman-Helpman "trade wars" model.[u] Suppose that good i is an export good for the home country and that, contrary to the discussion in the previous paragraph, its export industry is politically active even before any trade talks take place, ie, $I_i = 1$. The home-country's trade policy in the Nash equilibrium might involve an export subsidy $(t_i > 1)$ or an export tax $(t_i < 0)$. The first term on the right-hand side of (2) captures the force for a subsidy, reflecting the influence bought by the domestic industry. The second term captures the force for a tax, inasmuch as an export tax improves the country's terms of trade and thus contributes to the

[r] Regan (2015) does offer the argument that government support for exporters in the noncooperative setting would require export subsidies and that governments find such subsidies politically intolerable, for some unexplained reason. But as we have just seen, the combination of indifference to positive tariff revenues and intolerance for any trade subsidies leads inexorably to the conclusion that either trade is free or none takes place.

[s] Mitra (1999) develops a model in the spirit of Grossman-Helpman (1994) in which interests group endogenously form prior to the lobbying game by bearing a cost of organizing. However, the equilibrium in his model does have endogenous organization by some export industries, just as it has endogenous organization by some import-competing industries.

[t] This potential argument for a trade agreement has the same flavor as the commitment theories that are discussed in Section 4. That is, a government that has benevolent intentions *ex ante* may anticipate that it will be tempted to cater to organized interests *ex post*. By entering into a trade negotiation, it encourages the participation of the exporters in the political process and thereby induces an outcome that is closer to its *ex ante* preferences.

[u] A similar argument can be made using instead the Bagwell-Staiger reduced-form specification of the governments' objectives, $G(p, p_w)$ and $G(p^*, p_w)$, for which the Nash equilibrium is described by (7) and (8).

country's aggregate welfare (note that $e_i^* < 0$ when the foreign country imports good i). The net effect reflects, *inter alia*, the size of a, the government's weight on contributions relative to welfare. When a is large, $t_i < 1$, as concerns about welfare carry the day. When a is small, $t_i > 1$, as the lobbyists prefer.

Suppose the Nash equilibrium policy involves an export subsidy and that the countries come together to discuss a cooperative trade agreement. A small change (up or down) in the home country's subsidy has no effect on the government's objective, G, because t_i was chosen by the home government to balance the marginal effects of a policy change on contributions and welfare. But a small *increase* in the home export subsidy (starting from the Nash equilibrium) raises the objective G^* of the foreign government. An increase in the subsidy rate conveys a positive externality to the foreign government inasmuch as it reduces the world price of one of that country's import goods, ie, it improves the foreign terms of trade. Since the foreign government has chosen t_i^* to balance the marginal influence of its import-competing industry against the marginal affect on aggregate welfare, dG^*/dt_i must be positive at the Nash equilibrium. In short, the model predicts that trade talks ought to encourage larger export subsidies than in the noncooperative equilibrium. But the actual rules of the GATT and WTO do much the opposite, in fact they forbid export subsidies entirely.[v]

Unfortunately, the literature offers no compelling reason why trade agreements should outlaw export subsidies in a trading environment characterized by perfectly competitive markets. Perhaps the world economy is better described by pervasive imperfect competition, in which case the analysis of Section 3 comes into play. There, as we shall see, limitations on export subsidies can more easily be explained. Or perhaps further research will uncover other ways to resolve what Maggi (2014) has termed the "export subsidy puzzle."

3. INTERNATIONAL EXTERNALITIES WITH IMPERFECT COMPETITION

Until now, I have focused on international externalities that arise in a perfectly competitive trading environment. These externalities give rise to gains from cooperation and so provide incentives for governments to negotiate a trade agreement. I now turn to environments with imperfect competition, where a wider set of externalities prevail.

[v] As noted previously, the national policies for industry i in the cooperative equilibrium are not well determined; only the difference in policies is pinned down by the requirements for political efficiency, as expressed in (4). Accordingly, t_i might fall as the result of trade talks, so long as t_i^* does so as well. Still, it is troubling that the model suggests that marginal increases in export subsidies starting from the Nash equilibrium would enhance the combined political objectives of the two governments and yet the trade agreement calls for zero subsidies.

3.1 Firm-Delocation Externalities

Delocation refers to the exit by producers from some locations coupled with entry by new firms in other locations. The mix of producers' locations in an industry can matter for national welfare whenever firms set prices above marginal cost and transport costs create price differences across markets. In such a setting, a government may have an incentive to pursue trade policies that encourage entry at home and exit abroad, thereby changing the mix of prices faced by local consumers. If the governments act unilaterally, they will neglect the harm that delocation causes to consumers elsewhere.

Venables (1985, 1987) studied the welfare effects of tariffs in models with imperfect competition and free entry. In Venables (1987), for example, two countries produce varieties of a differentiated product under increasing returns to scale. They also produce a homogeneous good under constant returns to scale. The varieties are CES substitutes, as in Krugman (1980), and trade in the differentiated varieties entails shipping costs. Labor is the only factor of production. A zero-profit condition determines the number of varieties produced in each country. Venables shows that when the home country levies a small import tariff, domestic welfare rises even if the tariff revenues confer no social value. The gains come at the expense of foreign consumers. The mechanism for these welfare transfers is that of delocation: The tariff raises the profitability of producing in the home market, which generates additional entry there. Foreign firms lose directly from the incidence of the tariff and indirectly from competition with a greater number of home firms. So, some firms exit abroad. Since home products do not bear tariffs or shipping costs in the home market, they are cheaper than the imports. Therefore, an increase in the relative number of domestic producers reduces the home price index, which raises real incomes at home. Just the opposite is true in the foreign country.

In the model developed by Venables (1987), an export subsidy also can be used to raise home welfare. A subsidy, like a tariff, enhances profitability for home firms while reducing that of foreign firms, so it induces entry at home and exit abroad. Again, the change in the composition of firms benefits consumers at home while harming those abroad. Although Venables does not discuss the net effect on foreign welfare, it is straightforward to show that welfare there might fall; the direct benefit that foreign consumers receive from the subsidized prices can be more than offset by the harm from delocation.

What are the incentives for forming a trade agreement in the presence of delocation externalities? Ossa (2011) addresses this question in a model based on Krugman (1980). He considers trade between two countries that have symmetric preferences and the same production technologies. The countries may differ in their labor endowments and in their trade policies. Consumers have Cobb-Douglas preferences over a homogeneous good and a CES composite of differentiated varieties. The homogeneous good is produced with constant returns to scale and is freely traded. Varieties of the differentiated product require a fixed input of labor as well as a constant per-unit variable input.

The differentiated varieties are traded subject to an iceberg shipping cost and an *ad valorem* import tariff. Firms earn zero profits in an equilibrium with free entry.

Ossa assumes that the iceberg trade costs are sufficiently high that both countries produce differentiated products. Also, the labor endowments are sufficiently large that both countries produce positive amounts of the homogeneous good. As in Venables (1987), he finds that the real price index in each country is decreasing in the country's own tariff and increasing in the tariff of its trading partner, for small enough tariffs. The explanation again is delocation: A tariff at home generates entry by home firms and exit by foreign firms and the direct effect of the higher import prices due to the tariff is more than offset by the indirect effect of the change in the mix of producers.

If tariff revenues do not enter the government's welfare calculus, then the pair of welfare-maximizing governments set prohibitive tariffs in a Nash equilibrium of the policy game. If the revenues do figure in welfare, then the Nash tariffs are positive but finite. The interior solution results from the fact that by further raising an already high tariff, the government sacrifices tax revenue. This cost must eventually outweigh the benefit from further delocation. As usual, the noncooperative tariffs are detrimental to the joint welfare of the two countries. The benefit that each country achieves by delocation comes at the expense of its trading partner and generates further deadweight loss. Efficiency requires tariff reductions; eg, in the case in which tariff revenue does not enter welfare, at least one country's tariff rate must be set to zero under an efficient trade agreement.

Ossa emphasizes that the externality that arises in the noncooperative Nash equilibrium of his model should not be called a terms-of-trade externality. Since the CES demand structure implies a constant markup over marginal cost, the ex-factory prices of export goods are independent of the tariff rate. If we define the terms of trade to be the ex-factory price of an imported variety divided by that of a domestic variety, then the terms of trade are, in fact, independent of trade policy. If, instead, we define the terms of trade as the price index for exported varieties divided by the price index for imported varieties, then—as Ossa shows—a tariff that improves a country's welfare via delocation actually generates a deterioration in its terms of trade. The purpose of a trade agreement, he argues, is to internalize the externality that results when countries use their import policies to alter the composition of domestic and foreign firms in the market.

Bagwell and Staiger (2012b) consider a different but related model with delocation possibilities, namely one based on Venables (1985). Their model has linear demands for a homogeneous product and no income effects on demand. Home and foreign firms engage in Cournot competition in the two markets, which are assumed to be segmented. Trade costs are positive. As in Venables (1985), the equilibrium features two-way trade in identical products. A small tariff in the home country increases the number of firms located there, decreases the number of firms in the foreign country, and by altering the intensity of Cournot competition in the two markets, it raises home welfare at the expense of foreign welfare. In this setting, if governments are limited to using import

tariffs as instruments for delocation, then—just as in Ossa (2011)—the inefficient Nash equilibrium involves positive tariffs in both countries. Moreover, the countries have reason to negotiate a reduction in these tariffs, just as Ossa describes. In a symmetric setting, the efficient tariffs are zero.

In Bagwell and Staiger's model, just as in Ossa's, the countries also have incentive to introduce export policies. A small export subsidy starting from free trade raises a country's welfare just as does a tariff, and for much the same reason (see Venables, 1985). The entry of firms in the country with the subsidy and the exit of firms in the other country results in a change in the intensity of competition in the two markets that favors consumers in the country with the active policy.

Bagwell and Staiger proceed to consider the Nash equilibrium of a policy game in which the two governments can implement both import and export policies. They find, perhaps surprisingly, that in the Nash equilibrium, each country combines a tariff on imports with a *tax* on exports. The key to this finding is their observation that, when deviating from free trade, each government can always find a small tariff and a small export tax that together generate the same internal price and the same consumer surplus as are achieved under free trade but that yield positive tax revenue. In effect, import tariffs and export taxes are complements in their model, because a country that imposes a high tariff and thereby induces entry by a large number of local firms will want to avoid an export subsidy that would transfer a great deal of revenue to foreign consumers. Put differently, with an import tariff in place, an export tax by the home country that causes entry abroad will increase the volume of foreign exports on which the tariff is levied, thereby increasing the home government's revenues. Bagwell and Staiger conclude that, starting from a noncooperative equilibrium that has positive export taxes and high import tariffs, the countries might appear to have no reason to conclude an agreement that limits the use of export subsidies inasmuch as subsidies are absent from this equilibrium. Moreover, an efficient agreement is one in which any positive import tariff maintained by one country is exactly offset by an export subsidy in the other, so that the net effect of the offsetting trade policies on world prices is nil. In this sense, export subsidies must be tolerated (indeed encouraged) to achieve efficiency, unless import tariffs are fully eliminated. However, as the authors emphasize, an efficient agreement cannot leave governments free to choose whatever export subsidies they prefer; once tariffs have been reduced to low levels, the countries will have incentives to overuse these instruments to encourage delocation. While the model does not provide an explanation for bans on export subsidies, it does provide a reason why the use of subsidies should be regulated.

Finally, in Bagwell and Staiger (2015), the authors consider the nature of the externality that motivates a trade agreement in models of firm delocation. They show that the globally efficient policies coincide with the "political optimum," where the latter is defined as the vector of policies that would maximize $G(p, p^*, p_w)$ and $G^*(p, p^*, p_w)$ were the two countries for some reason to ignore the effect of its trade policy choices on the

world price. In this sense, they argue, and contrary to the claim in Ossa (2011), the fundamental purpose of a trade agreement in the delocation model is really to eliminate manipulation of the terms of trade. Maggi (2014) points out, however, that their conclusion very much relies on their assumption of no income effects in the demands for the imperfectly competitive good. Moreover, the meaning of "terms-of-trade manipulation" in a model of Cournot competition with free entry is not entirely clear, at least not to me.

3.2 Profit-Extracting and Profit-Shifting Externalities

International externalities also arise in imperfectly competitive environments when governments can use trade policies to extract monopoly rents from foreign producers or to shift profits from such producers to their domestic rivals. Katrik (1977) and Svedberg (1979) were the first to demonstrate that a government might be able to raise national welfare by imposing a tariff on a good imported from a foreign monopolist. They assumed that the home country confronts the foreign monopoly with a linear demand. In such a setting, a small specific tariff reduces the ex-factory price charged by the monopoly. Although consumers pay more for the good with the tariff in place, part of their payment goes to the home government in the form of tariff revenue. The monopolist's reduction in the ex factory price corresponds to a terms-of-trade improvement, for the home country, which boosts its welfare. Brander and Spencer (1984a) extended the Katrik and Svedberg analyses to include more general demands. When the foreign monopolist operates subject to a constant marginal cost, a specific tariff induces a reduction in its ex-factory price if and only if the demand curve is not too convex. The condition for a terms-of-trade gain in the policy active country is $R \equiv mp''(m)/p'(m) > -1$, where $p(m)$ is the home inverse demand for imports m. If this condition is satisfied—or equivalently, if the inverse demand curve is flatter than the inverse marginal revenue curve—then a small tariff improves the home country's terms of trade and the country's optimal tariff is positive. Much like the tariff that arises from unilateral welfare maximization in a competitive setting, the optimal rent-extracting policy is a beggar-thy-neighbor policy. The revenues captured by the home government come entirely at the expense of the foreign producer's profits. Since global efficiency requires an output greater than what the monopolist sells under free trade and since the home country's tariff reduces the monopolist's output, the optimal tariff in fact reduces global welfare.

Brander and Spencer (1984b) extend the analysis by considering a market in which a pair of firms with different national origins compete as duopolists. The firms compete in segmented markets by simultaneously choosing their deliveries to the two destinations. In this setting with Cournot competition, not only does a unilateral tariff by some country have the potential to extract rents from the foreign firm, to the benefit of the domestic

treasury, but it also changes the outcome of the strategic competition between the two firms, to the benefit of the domestic producer. Accordingly, if a tariff induces the foreign firm to reduce the ex-factory price of output destined for the home market, this is sufficient to ensure a welfare gain for the home country. But even if the ex-factory price rises, home welfare might rise due to the profit shifting that results from the changes in market shares.

Brander and Spencer go on to consider the Nash equilibrium of a noncooperative game between two welfare-maximizing governments. In the Nash equilibrium, each government sets a positive tariff on imports from the other's country's monopolist. In so doing, it captures revenue and shifts profits towards its domestic producer. But the profit-shifting effects in the two markets offset one another and, taken together, the pair of Nash tariffs reduce world output. The noncooperative tariffs exceed those that maximize world welfare. In this setting, the governments have an incentive to negotiate mutual trade liberalization and improved market access for their own national firms in the other country's market.

Brander and Spencer (1985) provide the cleanest analysis of the international externalities that arise from profit-shifting trade policies. They assume that there are single suppliers in each of two countries that produce a common good only for export to a third market. The governments can subsidize their local firm's exports in anticipation of the Cournot competition for third-country sales. Consider the unilateral incentives for the use of trade policy in one of the exporting countries. There is no consumption there, so no concern about consumer surplus. The subsidy payment represents a dollar-for-dollar transfer from the government to the domestic firm, which is neutral from the point of view of aggregate welfare, if tax revenues and firm profits are weighted equally. What remains, then, is only the strategic effect of the export subsidy on the outcome of the oligopolistic competition. When a government offers a subsidy, the local producer has a greater incentive to export than otherwise. For any given quantity of its rival's sales the local producer sells more than it would without the subsidy. In other words, the firm's best response function shifts outward in the space of the two output levels. If the foreign best-response function slopes downward—as is commonly assumed for Cournot competition—then the subsidy induces the rival producer to reduce its exports to the third market. This strategic response increases the market share for the subsidized firm and increases its export price relative to what it would be without the rival's retreat from the market. The price in the third-country market typically will be lower than what it would be absent the export subsidy—which implies a terms-of-trade loss for the subsidizing country. But the extra profits captured by the domestic firm at the expense of its foreign rival more than compensate for this. When both best-response functions slope downward, each government has a unilateral incentive to subsidize exports in a grab for oligopoly profits.

Again, we recognize an international externality. The profit gains for one producer come at the expense of the other. And the subsidy causes the price in the third country to

fall, which means that joint profits for the two exporters also fall. If both governments were to pursue their unilateral incentives, the resulting pair of subsidies would roughly neutralize one another, leaving market shares about where they would have been without the interventions. The firms may benefit from their governments' largesse, but welfare inclusive of the subsidy costs must fall in at least one of the exporting countries. Therefore, the two governments of the exporting countries have a shared incentive to negotiate a trade agreement that limits the use of such strategic subsidies. Notice that the purpose of such an agreement would not be to limit manipulation of the terms of trade (since the noncooperative policies are subsidies that, in fact, worsen the exporters' terms of trade), but rather to prevent the unilateral pursuit of profit shifting.[w]

As Brander and Spencer (1985) point out, and Bagwell and Staiger (2012a, Section 5) further emphasize, the externality associated with profit shifting can explain a trade agreement that limits or prohibits export subsidies among a pair (or group) of exporting countries, such as the Boeing-Airbus pact between the United States and Europe that limited the countries' use of credit subsidies on foreign sales of large passenger aircraft. But the profit-shifting externality cannot explain a limitation on export subsidies in the context of a multilateral trade negotiation in which all countries participate. The benefit that exporting countries would capture from an agreement that restricts export subsidies comes at the expense of higher prices and reduced consumer surplus in the importing country or countries. Bagwell and Staiger examine a Nash equilibrium in a three-country model with symmetric exporting firms in two of the countries and all consumption confined to the third country, much as in Brander and Spencer (1985). The Nash equilibrium in the three-country policy game involves profit-shifting export subsidies in each of the two exporting countries and a rent-extracting import tariff in the importing country. An agreement that achieves efficiency for the three countries together does not pin down the tariff in the importing country or the subsidies in the exporting countries; efficiency only determines the magnitude of the combined policy wedge. However, it can be shown that the volume of trade is inefficiently low in the Nash equilibrium as compared to what is required for efficiency. Therefore, a negotiated agreement should either lower the tariff in the importing country or raise the subsidies in the exporting countries. Under

[w] Eaton and Grossman (1986) have pointed out that the optimality of subsidies in an export duopoly rests heavily on the assumption of Cournot competition. If, instead, the rival exporters engage in Bertrand (price) competition, the Nash equilibrium policies typical involve a pair of export taxes. A unilaterally imposed tax serves to temper competition in the third market and to generate a more collusive outcome. In a Bertrand competition—where a tax on one firm's exports typically induces its rival to raise its price—the strategic response can generate revenues for the taxing government that exceeds the loss in its firm's profits. Nonetheless, an international externality arises from unilateral policies, because the governments do not consider the profit gain that the foreign firm enjoys as a result of their export taxes. In fact, the joint welfare of the two exporting countries would be greater if the export taxes were raised above the levels in the Nash equilibrium.

an efficient agreement, the trade policies should be set so that the firms in the two exporting countries receive the same cum-subsidy price as one another and the consumers in the importing country pay the same price as they would under perfect competition with free trade.

Bagwell and Staiger (2012a) revisit the question of what is the "fundamental purpose" of a trade agreement in settings with monopoly or oligopoly profits. In other words, they ask, What is the nature of the international externality that a trade agreement seeks to correct? Brander and Spencer might concede that rent extraction is a form of terms-of-trade manipulation inasmuch as the government captures revenue from its import tariff by inducing a foreign monopolist or oligopolist to lower its ex-factory price. But what about strategic subsidies that steal profits from other exporters and generate welfare gains despite causing a deterioration of the exporters terms-of-trade? Bagwell and Staiger conclude that the identification of a profit-shifting motive for a trade agreement arises only when the governments lack full sets of trade policy instruments to tax or subsidize imports and exports and when the importing countries are, for some reason, left out of the calculus. They show that, in a variety of settings with excess profits—albeit all with quasilinear utility—a country's welfare can be written in reduced form as a function $G(p,p^*,p_w)$ of local prices p at home and p^* abroad, and a world price, p_w.[x] They define again a "political optimum" as the policies that the governments would choose in a Nash equilibrium if, for some reason, they were to ignore the welfare enhancing effects of a terms-of-trade improvement, given local prices at home and abroad. That is, their benchmark arises when all governments choose policies to satisfy their first-order conditions, but in the process act as if $G_{p_w}(p,p^*,p_w) \equiv 0$. They show that the efficient agreement coincides with this benchmark, which leads them to conclude once again that the fundamental externality operates through the terms of trade.

It is not clear to me why their benchmark is appropriate, especially in a setting such as this one where the prices in each market are chosen by active players rather than resulting from market clearing. What does it mean—aside from the formal definition in terms of the first-order condition—that the government acts "as if the terms of trade do not affect aggregate welfare given the domestic prices?" And why is it so important that we pin a name to the externality, be it "profit shifting" or "terms of trade"? Be that as it may, an interesting insight emerges from Bagwell and Staiger's analysis. Namely, in a setting in which each traded good is subject both to an export tax or subsidy imposed by the exporting country and an import subsidy or tax imposed by the importing country, and when utility is quasilinear so that the traded good in question is not subject to any income effects, the policies in the importing and exporting countries become perfect substitutes in terms of their effects on quantities, consumer surplus, and profits. When one

[x] The competitive setting considered in Bagwell and Staiger (1999, 2002) is a special case in which the foreign price p^* does not enter the reduced-form government objective function.

government sets its policy at a level that the other takes as given, the second government can always "undo" the effects of this policy on local prices in each market by an appropriate choice of its own instrument. In so doing, it will impact its trade tax revenues or subsidy outlays, but it can achieve its chosen targets for consumer surplus and local producer profits. This means that the governments can each "get what they want" in the Nash equilibrium, except for the implied revenues and fiscal costs.[y]

3.3 Profit-Extracting Externalities in International Outsourcing Relationships

A recent paper by Antràs and Staiger (2012) addresses the inefficiencies that result from noncooperative policy setting in an environment with international outsourcing. In such circumstances, governments may use their trade policies both to correct allocative inefficiencies that result from incomplete contracting and to extract rents from foreign producers. The two governments share the former objective but not the latter, and a trade agreement may be needed to help them to achieve their common goals while avoiding the inefficiencies that result from conflict.

The economic environment has customized intermediate inputs, two-sided buyer-supplier relationships, and incomplete contracts that give rise to hold-up problems. Two countries, Foreign and Home, are small in relation to the world market for some final good. They take the world price of the final good as given. Foreign alone among the two can produce intermediate inputs, with a unit continuum of potential suppliers. Home alone among the pair can produce the final good, with a unit mass of potential buyers. Each potential supplier in Foreign is matched randomly with a potential buyer in Home and the two engage in bilateral bargaining in order to work out an outsourcing arrangement. If the bargaining breaks down, no final good can be produced and any customized inputs produced for the relationship become worthless. Either country can import the final good from the rest of the world at the fixed world price and Home can export the final good to the world market at this same price.

Imperfect contracting manifests in that the specifications of the input cannot be stipulated before production takes place. Instead, the supplier must produce some quantity of the intermediate good and then negotiate to sell it to its downstream partner. Given that the outside options are zero at the bargaining stage, the negotiation always results in a transaction, with a division of *ex post* surplus dictated by exogenous parameters. But the supplier in any relationship anticipates that it will bear the full cost of production while capturing only a fraction of the surplus; this hold-up problem generates

[y] As Maggi (2014) notes, this would not be true if the allocation of revenues had an effect on market demands, as they would in a demand system with income effects. It would also not be true if the international distribution of revenues and outlays had implications for efficiency, for example, if raising revenue entails deadweight losses from distorting taxes.

underinvestment in the intermediate input and an inefficiently low level of offshoring in a free-trade equilibrium.

Now consider the use of trade policies. The government of Foreign can encourage greater investment by subsidizing exports of the intermediate good. The government of Home can do likewise by subsidizing imports of the input or by subsidizing exports or taxing imports of the final product. But the combination of these three policies has implications for the distribution of rents among the two firms and the two governments.[z] And the governments care about national welfare, not global welfare; if they act unilaterally in a noncooperative equilibrium, they will take actions on the margin that benefit their local producers and their treasury at the expense of profits and revenues abroad.

Consider first the incentives facing the government of Home. If its goal were to maximize global welfare given the foreign export subsidy (or tax), it would allow free trade in the final good and subsidize imports of the intermediate good to induce the efficient level of input production. Such a policy would solve the hold-up problem without generating any by-product distortion of consumption decisions. But, by taxing exports or subsidizing imports of the final good, the government of Home can engineer a bargaining outcome that is more favorable to its local firms at the expense of their foreign suppliers. The best response of the government of Home to any policy in Foreign is to combine such a tax on exports or subsidy to imports with an import subsidy for intermediates that pushes the volume of input trade closer to the efficient level, but not fully there. The government of Foreign, in turn, finds it optimal to respond to this pair of policies with a tax on exports of the intermediate input. Two considerations explain why a tax is optimal. First, the foreign firms choose their output levels to maximize their own profits, which coincides with the goal that the foreign government has for its local industry. Second, the tax generates revenue for the foreign government and part of the tax burden is passed on to home firms in the bargaining process. As Antràs and Staiger (2012) show, the Nash equilibrium involves an inefficiently low volume of input trade and an inefficiently low price of the final good in Home, which creates the motivation for a trade agreement.[aa]

The requisite trade agreement in these circumstances is, however, rather subtle. First, efficiency requires intervention in the input market, not free trade. The governments

[z] The government of Foreign has no reason to place a tariff on imports of the final good, inasmuch as this has no effect on the efficiency of the outsourcing relationships and it cannot alter its terms of trade vis-à-vis the rest of the world.

[aa] Antràs and Staiger proceed to introduce political-economic concerns in the form of a potential extra weight that the government attaches to profits relative to consumer surplus or tax revenues. They then ask whether the so-called "political optimum" in this setting, defined as elsewhere, achieves (political) efficiency for the two governments. They conclude in the affirmative if the governments place no extra weight on profits, but not otherwise. Accordingly, they describe the trade agreement that arises with political-economic concerns as addressing more than just a terms-of-trade externality.

must jointly subsidize the trade in intermediates to overcome the underinvestment associated with the hold-up problem. But, second, it is not enough that they agree to an appropriate level of joint subsidy to intermediate trade. If such a joint subsidy were agreed and the government of Home were left free to set its own trade policy for final goods—perhaps because Foreign does not sell this good to Home and it can always import this good at a fixed terms of trade from the rest of the world—then the government of Home would subsidize exports of the final good in order to tilt the bargaining between supplier and buyer in favor of the latter. In fact, an efficient agreement must constraint not only the policies that directly affect input trade, but also those that affect the outsourcing relationship in other ways.

3.4 International Agreements to Protect Intellectual Property

Most of the literature on the purpose of international agreements focuses on contracts aimed at limiting the opportunistic use of trade policies. But the externalities approach that I have outlined here can also be applied to international agreements that may arise in other policy areas. To illustrate, I will briefly discuss the purpose of international agreements to protect intellectual property, such as the TRIPS Agreement in the WTO. A similar approach has also been applied, for example, to externalities that arise from governments' choices of environmental policies; see the survey by Barrett (2005) for a review of this research.

Grossman and Lai (2004) consider two countries, North and South, that differ in the sizes of their populations and in their endowments of human capital. The countries use labor alone to produce varieties of a horizontally differentiated product and a homogeneous good. They use labor together with human capital to develop new varieties of the differentiated product. North has a larger endowment of human capital. Preferences are quasilinear, with each differentiated product generating some consumer surplus for households. The differentiated products are sold at monopoly prices by their inventors until imitation takes place due to imperfect patent protection or until the patent runs out. Every product becomes obsolete after a fixed period of time. There is free entry into product development, so the dynamic equilibrium is characterized by equality between the expected profits from a new product over the course of its economic life and the cost of developing such a product. In the steady state, the rate of invention of new products exactly matches the rate of obsolescence of old products.

The welfare maximizing governments in North and South choose policies that determine the degree of their protection of intellectual property. Grossman and Lai define a pair of policy variables as $\Omega_N \equiv \omega_N(1 - e^{-\rho\tau_N})/\rho$ and $\Omega_S \equiv \omega_S(1 - e^{-\rho\tau_S})/\rho$, where ω_J is the instantaneous probability that a patent used in country J will be violated due to lack of sufficient enforcement, τ_J is the duration of patent protection in country J, and ρ is the discount rate. Thus, Ω_J is a combination of the length and strength of patent protection in

country *J*. By assumption, these policies are applied by each government with *national treatment*; ie, local and foreign producers are treated similarly by the patent enforcement authorities of each country.

There are two international externalities that arise in this setting. First, governments that maximize national welfare ignore the surplus that consumers in the other country derive from a new invention over the course of that product's subsequent economic life. Second, governments that are concerned only with national welfare neglect the loss of producer surplus (profits) that foreign monopolies suffer when imitation occurs or patents expire. Both of these externalities point in the same direction: patents are too short and too weakly enforced in a noncooperative regime of patent policies compared to the efficient level of protection of intellectual property. The externalities create the opportunity for a Pareto-improving patent agreement.

If the governments behave noncooperatively, their best-response functions are downward sloping curves in (Ω_N, Ω_S)-space, because the policies implemented by the two governments are strategic substitutes; when one country affords greater protection of intellectual property, this induces greater innovation by firms in both countries and reduces the incentive that the other country has to offer its own inducement for R&D. The Nash equilibrium occurs at the unique intersection of these downward sloping curves. Grossman and Lai show that, in a Nash equilibrium, if the size of the consumer population in North is at least as large as that in South and the endowment of human capital in North is larger that in the South, then $\Omega_N > \Omega_S$; ie, North provides greater protection for intellectual property than South. In such circumstances, North has greater incentive to protect intellectual property, because it has more consumers who can enjoy the surplus from new products and it has more monopolies that stand to lose profits by patent infringement or patent termination. Notice that these results do not rely on discrimination, because national treatment precludes any discrimination in the application of patent rules.

Grossman and Lai compare the patent policies that emerge in a Nash equilibrium with those that would be stipulated by an efficient agreement. The efficiency frontier lies uniformly outside the two best-response functions, because the two externalities imply that, given any policy of the other government, each government provides less protection of intellectual property than would maximize global welfare. Therefore, an international agreement must strengthen patent protection in at least one country and provide greater incentives for innovation worldwide. An efficient agreement need not strengthen patent protection in both countries in order to generate Pareto welfare gains and "harmonization" of national patent policies is not necessary for global efficiency. In fact, a continuum of combinations of Ω_N and Ω_S can be used to achieve efficiency; all that matters for efficiency is the aggregate protection of intellectual property in the world economy, whereas the policies required of each country under an efficient agreement govern the division of welfare between the two.

4. TRADE AGREEMENTS AS COMMITMENT DEVICES

In Sections 2 and 3, I reviewed a literature that treats trade agreements as a means to overcome international externalities. The externalities may arise from the incentives governments have to manipulate their terms of trade or from other externalities, such as the incentive to relocate firms to the local market or to shift profits to local oligopolists. Another strand of literature offers an alternative explanation for trade agreements, namely that they provide a means for governments to tie their own hands and resist the temptation to give in to local special interests that advocate polices inimical to the general good. I prefer to think of the commitment motive as explaining why a country might sign an existing trade agreement, rather than a reason for two countries to get together to negotiate an agreement *de novo*. It is not clear to me why governments would prefer to design a trade agreement to achieve commitment vis-à-vis their own special interests rather than to self commit by some other, simpler means. Negotiating a new agreement is a complicated process that involves many compromises; wouldn't it be easier to do so unilaterally by, for example, passing a constitutional amendment that restricts the use of trade policy instruments? However, if an agreement already exists in some form, a country may choose to take advantage of its existence by acceding to its terms. Be that as it may, the literature offers an interesting answer to the question, Why might a government be willing to sign a trade agreement?

Pre-commitment is desirable when a policy that the government regards as beneficial before the private sector takes some irreversible (or costly to reverse) action no longer is so afterwards. This situation creates a "time-inconsistency" problem, as famously described by Kydland and Prescott (1977). The literature offers several examples of economic environments where the trade policy that is optimal *ex ante* no longer is so *ex post*. I will describe informally a setting akin to that in Staiger and Tabellini (1987) as an example.

Suppose the world price of some import good falls, which depresses incomes for those who work in the import-competing sector. This creates an incentive for workers to move from the adversely impacted sector to others. Imagine that workers must make a decision at some point in time whether to move to a new job or not. Moving requires them to incur a sunk cost that cannot be recouped subsequently. Let there be a distribution of such costs among workers in the import-competing sector, so that all those with a personal cost below some critical level move, while the remainder stay in their original jobs, albeit at lower pay. Finally, suppose the government values high national income, but also has a preference for an equal distribution of that income.

Once the trade shock has occurred and workers have made their irreversible choices to stay or move, the government will see higher incomes in the export sector than in the import-competing sector. If resource allocation has been fully determined by this point, the policy makers may find it attractive to use trade policy as a means to redistribute

income. The government can impose an import tariff that somewhat restores wages in the import-competing sector, thereby narrowing the income gap that results from the terms-of-trade shock.

However, if the government is free to use trade policy as a redistributive tool, the workers may anticipate such interventions when they make their decisions whether to move jobs or not. If the workers understand that the government's optimal policy will involve intervention that partly restores wages in the import-competing sector, fewer of them will move to the export sector than would be the case without such expectations. The government's *ex post* response to potential inequality will come at a high cost in terms of national income, because there will be less movement of workers than what efficiency requires. Herein lies the potential benefit from precommitment: a trade agreement that precludes tariff hikes in response to terms-of-trade shocks can increase allocative efficiency to such an extent as to more than offset the perceived social welfare cost from accepting greater inequality.[ab]

Maggi and Rodriguez-Clare (1998) extend a similar logic to a setting with political-economic forces at work. Consider a small country that faces given world prices and that has a fixed stock of capital. Initially, the capital is malleable and can be allocated to either of the economy's two, nonnumeraire sectors. Once allocated, however, the capital is specialized such that it no longer can be used to produce the other good. The timing is as follows. First, capital owners allocate their capital to one industry or the other. Then interest groups form (exogenously) and the organized groups lobby for trade policy.[ac] In a small departure from Grossman and Helpman (1994), suppose that the lobbying proceeds by Nash bargaining between the policy maker and the organized groups, with an exogenous fraction σ of the surplus accruing to the former. Finally, the government implements trade policies, competitive firms hire capital and mobile labor to produce output, and households devote their after-tax incomes (reflecting payments made or rebates received to balance the government's budget) to consumption of the three goods. The policy maker has an objective function of the form $G(W,C) = aW + C$, while the lobbies, which represent a small fraction of the total population, seek to maximize

[ab] One might reasonably ask, In what sense does a trade agreement preclude a tariff hike? Mightn't a government that has signed a trade agreement decide anyway to raise its tariffs? Presumably, a government that is tempted to use trade policy to redistribute income might be dissuaded from doing so by its participation in a trade agreement only if it fears some sort of retaliation from its agreement partners. Moreover, the country's trading partners will not have any incentive to retaliate against an unauthorized tariff hike unless there are external effects of trade policies. In this general sense, the commitment power of a trade agreement also relies on the presence of international trade policy externalities.

[ac] Mitra (2002) disallows intersectoral capital mobility both ex ante and ex post but introduces a fixed cost of political organization such that lobbies form endogenously. In his setting, as well, a government may wish to sign a trade agreement in order to precommit to free trade in anticipation of how the political economy will play out in the absence of commitment.

their industry's capital income in view of the owners' negligible stake in aggregate consumer surplus and in the government budget.

Suppose there is a multilateral trade agreement in effect, but that the government of the small country has not acceded to it, thereby retaining its sovereign right to set whatever trade policies it likes. Once the capital has been allocated and the lobbies have been formed, the organized groups bargain with the government over trade policies. At this stage, the fallback positions are zero contributions from any organized group and free trade; the latter policy maximizes aggregate welfare and therefore the government's objective G in the absence of any meeting of the minds about contributions and policies. Let $\Pi_i[(1+t_i)p_i^*, K_i]$ be the payments to capital in industry i when the world price is p_i^*, the ad valorem tariff or export subsidy applied to this good is t_i, and the capital that has previously been allocated to the sector is K_i. The joint surplus of the organized lobbies and the policy maker is given by

$$J(\mathbf{K}) = \max_{[\mathbf{t}]} aW(\mathbf{t},\mathbf{K}) + \sum_i I_i \Pi_i\left[(1+t_i)p_i^*, K_i\right] - aW(\mathbf{0},\mathbf{K}) - \sum_i I_i \Pi_i\left[p_i^*, K_i\right],$$

where $\mathbf{t}$ is the vector of trade policies applied to the two nonnumeraire industries, $\mathbf{K}$ is the vector of capital allocations to the two nonnumeraire industries, $W(\mathbf{t},\mathbf{K})$ is aggregate welfare when the trade policies are $\mathbf{t}$ and the capital allocations are $\mathbf{K}$, $W(\mathbf{0},\mathbf{K})$ is aggregate welfare under free trade when the capital allocations are $\mathbf{K}$, and I_i is an indicator variable that takes on a value of one if industry i is politically organized and zero otherwise. In these circumstances, the government achieves the political welfare $G = aW(\mathbf{0},\mathbf{K}) + \sigma J(\mathbf{K})$, in view of the fact that it receives its fallback level of welfare plus the fraction σ of the surplus in the lobbying relationships. The net payoffs for the organized groups can be calculated similarly.

Still assuming that the government has not acceded to the multilateral trade agreement, we can solve for the capital allocations in a rational-expectations equilibrium. The equilibrium allocations, $\widetilde{\mathbf{K}}$, are those that equalize the expected net incomes for capital allocated to the alternative uses, considering the trade policies that are anticipated as well as the contributions (if any) that are expected to be made.

Now suppose that the government has the opportunity at the outset to sign a trade agreement that commits the country to free trade. By doing so, it foregoes the surplus from its political relationships with the lobbies. Although the government sees free trade as a desirable outcome for aggregate welfare, given any $\mathbf{K}$ it prefers to enact protectionist or export-promoting policies in exchange for the valuable contributions it can extract from the lobbies. So why would the government potentially wish to precommit to trade freely? The answer, as before, has to do with the *ex ante* allocation of capital. With a trade agreement in place, the capital owners recognize that lobbying for protection or export subsidies will be futile. Accordingly, they anticipate earnings of $\Pi_i[p_i^*, K_i]$ in industry i.

The allocation that equates the returns in the two industries, $\mathbf{K}_{FT}$, is of course the one that is *ex ante* most efficient. Accordingly, $W(\mathbf{0},\mathbf{K}_{FT}) \geq W(\mathbf{0},\widetilde{\mathbf{K}})$, with strict inequality whenever $\widetilde{\mathbf{K}} \neq K_{FT}$. In deciding whether to accede to the trade agreement, the government compares $aW(\mathbf{0},\mathbf{K}_{FT})$, its expected political welfare when it joins the agreement, to $aW(\mathbf{0},\widetilde{\mathbf{K}}) + \sigma J(\widetilde{\mathbf{K}})$, the net payoff including contributions that it achieves when it opts not to join.

Maggi and Rodriguez-Clare (1998) observe that joining the trade agreement will be attractive to the government if $\sigma = 0$, but not so if $\sigma = 1$. If the policy maker captures none of the surplus in its relationship with the lobbies, then the comparison hinges on the aggregate welfare that results from the initial allocation of capital, and $aW(\mathbf{0},\mathbf{K}_{FT}) > aW(\mathbf{0},\widetilde{\mathbf{K}})$. On the other hand, if the policy maker captures most of this surplus, then the capital owners will anticipate a net return after lobbying that differs little from what they would earn under government's fallback position of free trade; accordingly, the allocation $\widetilde{\mathbf{K}}$ will be very close to $\mathbf{K}_{FT}$. Then $aW(\mathbf{0},\widetilde{\mathbf{K}}) + \sigma J(\widetilde{\mathbf{K}}) \approx aW(\mathbf{0},\mathbf{K}_{FT}) + \sigma J(\widetilde{\mathbf{K}}) > aW(\mathbf{0},\mathbf{K}_{FT})$. In short, the government prefers to tie its own hands when its bargaining position vis-à-vis domestic interest groups is weak, but not when it is strong; in the latter case, it can use the flexibility to implement trade policies to attract contributions from the lobbies that exceed (in its political assessment) the losses that it policies impose on the general public.

Brou and Ruta (2013) extend the model of Maggi and Rodriguez-Clare (1998) to allow for domestic subsidies. The government can transfer income to special interests in the import-competing sector either by affording protection or by providing production subsidies that are financed by distortionary taxation. An agreement that limits only the use of trade policies will not be very attractive to the government, because the value of precommitting the use of one instrument is limited when the interest group knows that the government can readily substitute another. Brou and Ruta use the model to analyze the Subsidies and Countervailing Measures Agreement in the WTO system; a government that wishes to tie its hands vis-à-vis domestic lobbies will be more inclined to accede to an international agreement that limits the use of production subsidies alongside the use of tariffs (or export subsidies) than one that restricts only trade policies.

Whereas an agreement that eliminates tariffs and export subsidies may not be attractive for commitment purposes if it leaves the government free to use good substitutes for trade policies like production subsidies as alternative means to redistribute income, a government may be willing to sign an agreement that does not constrain the use of more inefficient means of income transfer. Limão (2011) study a government's willingness to constrain the use of tariff policies when nontariff barriers (NTBs) are available as substitutes. They assume that NTBs, like tariffs, transfer rents to domestic special interests in the import-competing sectors. But these policies dissipate some of these rents, so they are strictly less efficient as tools of redistribution. The government may be willing to sign a

trade agreement that constraints the use of tariffs for reasons akin to those offered by Maggi and Rodriguez-Clare while recognizing that the costliness of NTBs provides some assurance that it will not succumb to that temptation, or at least not do so to any great extent. Limão (2011) offer their analysis as an explanation for why the WTO system binds tariffs and prohibits production subsidies, but does not constraint the use of a variety of less efficient policies that can serve as (imperfect) substitutes for the tariffs and subsidies.

If governments have incentives to negotiate a trade agreement in order to mitigate international externalities, then the benefits from precommitment vis-à-vis domestic special interests can provide an added inducement for doing so. Maggi and Rodriguez-Clare (2007) incorporate the two distinct benefits of a trade agreement in a single model, which generates some interesting further insights. There are two symmetrically different countries and three goods. In the home country, capital endowment K_1 can be used to produce either the numeraire good or good 1. Capital endowment K_2 can only be used to produce good 2. In the foreign country, capital endowment K_1^* can be used only to produce good 1. Capital endowment K_2^* can be used either to produce good 2 or the numeraire good. Endowments and demands are such that the home country imports good 1 and exports good 2. Production technologies are linear in capital and only the import-competing sector in each country is politically organized. The governments have a political motive for providing protection and a terms-of-trade motive, as in Grossman and Helpman (1995b). The international externality lends value to a trade agreement, as does the capital misallocation that results from anticipated protection.

Let the initial allocations of capital and the initial tariff rates emerge from the Nash equilibrium of a noncooperative game, as in Grossman and Helpman's trade war. The resulting tariffs are larger than the welfare-maximizing rates, and so there is overinvestment in the import-competing industry. Initially there is no thought of a trade agreement and the allocations do not anticipate one being negotiated. But suddenly, that possibility arises.[ad] The lobbies and the governments negotiate over a campaign contribution and the terms of an agreement, which takes the form of an (endogenous) cap on tariff rates. The agreement maximizes the joint surplus of the two governments and the two lobbies. After an agreement is signed—if that happens—each owner of a unit of K_1 in the home country or of type K_2^* in the foreign country has the opportunity to move that capital to the numeraire sector with some exogenous probability z. The parameter z is meant to capture the degree of capital mobility in the import-competing industry, from complete specificity ($z = 0$) to perfect mobility ($z = 1$). After any capital reallocation takes place, the lobby in each country and its government negotiate again about the actual level of the

[ad] Maggi and Rodriguez-Clare (2007) show that their insights carry over to a version of the model in which the opportunity to negotiate a trade agreement is not a surprise, but rather is perfectly anticipated when the initial capital allocation takes place.

tariff, but this time subject to the constraints imposed by the international agreement. Finally, political contributions are paid and production, trade and consumption take place.

Maggi and Rodriguez-Clare show first that the governments and the lobbies (weakly) prefer an agreement that imposes caps on tariffs to one that explicitly determines their levels. If an agreement must stipulate exact levels of the policy instruments, a jointly efficient agreement will reduce tariffs from their noncooperative levels, but not to zero, and there will remain distortions in the allocation of capital and in consumption. Now suppose that there is an option instead to set that same tariff as a maximum, rather than as a requirement. Such an agreement would leave discretion for each government to set the actual tariff below the agreed ceiling, and so the lobbies would have to offer contributions to avoid such an outcome. Inasmuch as a tariff ceiling imposes an additional burden of contributions on the capital owners, it reduces the overinvestment in politically organized sectors. Joint surplus is raised by an agreement that reduces such distortions, and all parties (governments and lobbies) can share in the gains by appropriate adjustment of contributions in the initial round of lobbying. Accordingly, the model predicts that a trade agreement will designate tariff ceilings ("bindings") rather than tariff levels, if such a contract is possible.

The model links the size of tariffs cuts (from the initial Nash levels to ultimate policies that are set subject to the constraints of the agreement) to the degree of capital mobility, as captured by the parameter z. In the extreme, if $z = 1$, the owners of capital in the import-competing sector can always earn the return promised in the numeraire sector. There are no rents to be captured in the *ex post* stage of lobbying, hence the lobbies are not willing to pay anything to their governments to compensate for long-run distortions associated with protection. Accordingly, the trade negotiation cuts tariffs to zero. At the opposite extreme, if $z = 0$, there is no possibility to undo the misallocation of capital. The equilibrium agreement eliminates the terms-of-trade component of the Nash tariff—as in Grossman and Helpman's trade talks—but it cannot reduce the domestic-commitment problem. Maggi and Rodriguez-Clare show that the tariff cut relative to the initial, Nash equilibrium level is monotonically increasing in capital mobility, z. If we interpret this result as a cross-sectional prediction, it says that tariff cuts should be deeper in those industries where capital specificity is less and outward reallocation is easier.

Finally, the authors consider how the extent of trade liberalization varies with the extent of the governments' concern for aggregate welfare relative to campaign contributions. In a setting without precommitment considerations, such as Grossman and Helpman (1995b), the Nash equilibrium tariffs tend to be higher when the governments' concern for aggregate welfare is small. High initial tariffs limit the volume of trade and thus weaken the terms-of-trade externalities. We might expect, therefore that tariff cuts will be smaller when the government has a greater taste for contributions and when it places less weight on aggregate welfare. But in the model developed by Maggi and

Rodriguez (2007), the opposite can be true when capital is sufficiently mobile. In this setting, high initial tariffs in the Nash equilibrium (that emerge when the government has less concern for aggregate welfare) imply a large departure from allocative efficiency. If z is large, there is much to be gained by committing to low tariffs and thereby inducing a substantial reallocation of capital. Accordingly, the agreement should call for deep tariff cuts when a small government weight on welfare induces a large initial distortion and a high degree of capital mobility implies that the costs can be reduced greatly by a commitment to freer trade.

5. INCENTIVES TO FORM REGIONAL OR PREFERENTIAL TRADE AGREEMENTS

Until now, I have mostly discussed the purpose of trade agreements in the context of a two-country world economy. One exception concerned the incentives that exporting countries might have to limit their use of strategic export subsidies to third-country markets. But bilateral and plurilateral agreements take place in many other contexts. Indeed, the number of bilateral, regional, or other preferential trade pacts has been growing in leaps and bounds, giving rise to what Bhagwati (1995) has called the "spaghetti bowl" of international agreements. Baldwin and Venables (1995), Panagariya (2000), and Krishna (2005) have written excellent surveys of the theoretical literature on the economic effects of preferential trade agreements (PTAs), while Limão (2016) reviews the empirical research in Chapter 14. In this section, I will limit myself to those few articles that address the reasons that governments might choose to negotiate exclusive agreements—rather than, or in addition to, multilateral agreements—in a world economy with more than two countries.

5.1 The Ohyama-Kemp-Wan Theorem

Any discussion of the incentives for trade agreements among a limited set of countries should begin with the renowned Ohyama-Kemp-Wan theorem (see Ohyama, 1972; Kemp and Wan, 1976). These authors proved a striking result: If lump-sum transfers are feasible within a union, any group of countries can form a customs union and set a common external tariff in such a way that all member countries benefit and no excluded country is harmed. The logic of the argument is simple. Let the union choose an external tariff that leaves its members' aggregate vector of trades with the rest of the world unchanged. (We know this always is possible based on results about the existence of market-clearing prices in a competitive equilibrium.) Then markets will clear in the rest of the world at the prices that prevailed before the union. With the same prices and the same trades, these countries are exactly as well-off as before. As for the union members, we can treat their vector of trades as if it were an endowment vector. Efficiency within the union requires equalization of marginal rates of transformation and of marginal rates

of substitution across member countries. This is achieved by a common vector of prices, which in turn is guaranteed by internal free trade. All that remains is to share the efficiency gains, which can be accomplished costlessly when the countries have access to lump-sum transfers between members.[ae]

Countries have an incentive to form customs union in order to achieve allocative efficiency. The same is true if the governments have noneconomic production targets (Krishna and Bhagwati, 1997). The result also extends to environments in which the policy makers have political objectives besides aggregate welfare, provided they have access to efficient instruments to redistribute income to favorite interest groups. However, as Richardson (1995) cautions, the Ohyama-Kemp-Wan theorem should be interpreted with care. It cannot be taken to imply that a group of countries can form a mutually beneficial customs union *no matter what is the response in the rest of the world*. The proof assumes that the rest of the world responds to the customs union by making the same vector of trades at the same prices; ie, the aggregate offer curve of the rest of the world is not affected by the formation of the union. If nonmembers can respond by, for example, setting a new vector of "optimal" or "politically guided" tariffs, then gains for union members are not assured.

5.2 Terms-of-Trade Gains

Just as a single country can gain at the expense of its trading partners by exploiting its monopoly power in world markets, so too can a group of large countries benefit by cooperating to exploit their joint market power in trade. In fact, two large countries stand to gain by forming a free-trade area (FTA) even if they do not alter their external tariffs vis-à-vis nonmember countries.

Consider a three-country world in which countries A and B form an FTA with distinct external tariffs and country C represents the excluded rest of the world. Suppose first that A imports some good from B and C, both of which have upward-sloping, competitive supply curves. Country A has an external tariff of t^A that applies initially to imports from all sources. Once it forms an FTA with B, the tariff applies only to imports from C. Let p^C represent the initial price received by exporters in both B and C for sales in country A, so that $p^C\left(1+t^A\right)$ is the pre-FTA domestic price in A. When country A eliminates its tariff on imports from B, suppliers there can sell in A's market at the prevailing domestic price. There is excess supply in the world market at the original prices, as firms in B

[ae] Dixit and Norman (1986) show that lump-sum transfers are not necessarily to share the gains from trade, if countries have access to a full set of consumption and factor taxes and subsidies. Panagariya and Krishna (2002) extend the Ohyama-Kemp-Wan result to include free trade areas in which member countries maintain separate external tariffs vis-à-vis imports from the rest of the world but trade freely within the area. In this case, member tariffs are chosen to preserve the initial vector of trades by each area country and resulting internal prices are not the same in these countries.

produce more at the higher delivered price. The price of imports from the rest of the world must fall to clear the world market. The fall in p^C benefits A while harming B, but since A's imports are larger than B's exports, the net effect must be positive. The members of the FTA capture a terms-of-trade gain as a result of trade diversion.

Now suppose instead that A and C both import from B. The initial price received by exporters in B is p^B and the domestic price in A is $p^B\left(1+t^A\right)$. With the elimination of the barrier to internal trade, the price in country A falls. This creates excess demand. The supply price from B must rise to clear the world market. The increase in p^B benefits B while harming A, but since B's exports are larger than A's imports, again the gains outweigh the losses. In this case, trade creation generates a positive terms-of-trade effect for the members of the FTA.

Countries that forge a regional or preferential trading arrangement can gain even more by adjusting their external tariff or tariffs. The incentive for doing so is analogous to that for merger among competing oligopolists; whereas each country can exploit market power on its own, the joint influence over world prices is greater than for any one alone. The potential gains are evident as a corollary to the Ohyama-Kemp-Wan theorem (or the Panagariya-Krishna extension to FTAs): If a group of countries can benefit by forming a customs union with an external tariff that leaves the terms of trade the same as before, then they can benefit even more by adjusting their external tariff optimally.

Kennan and Riezman (1990) investigated whether countries can gain by forming a customs union, once the tariff response by nonmember countries is taken into account. They examined an endowment economy with three goods and three countries that are symmetric up to a relabeling of the endowment goods, with a linear expenditure system in all countries. In this setting, they compared the outcome in a Nash equilibrium in which two of the countries allow internal free trade while jointly choosing an optimal external tariff to the outcome in a Nash equilibrium without any cooperation, and the outcome in an equilibrium with global free trade. Whenever each country's endowments of its export good is not too large relative to the total world endowment, any pair of countries fares better in a customs union equilibrium than in one with global free trade. Therefore, the possibility of forming customs unions undermines the prospects for a multilateral trade agreement

Kennan and Riezman also examined how size affects the incentives that countries have to form a customs union and, in particular, whether a pair of smaller countries can gain by joining forces to enhance their collective market power once retaliation is taken into account. In their examples, welfare rises in each of a pair of smaller countries when they form a customs union compared to the outcome with no cooperation, but each fares less well than it would in an equilibrium with universal free trade. In contrast, larger countries fair better in a Nash equilibrium in which they are partners in a customs union compared to both the Nash equilibrium without any cooperation and the equilibrium with global free trade.

Of course, the motivation to form a customs union or FTA in order to exert collective market power relies on the same beggar-thy-neighbor calculus as does the unilateral imposition of optimal tariffs in a setting without trade agreements. The gains to the member countries come entirely at the expense of countries on the outside. Krugman (1991a) began a small literature that addresses the welfare implications of having a trading system with multiple, nonoverlapping blocs in which each bloc allows internal free trade but behaves noncooperatively vis-à-vis the others. He considers a world with a large number of symmetric countries divided into a smaller number of symmetric blocs. Each country produces a unique good and all such goods are CES substitutes for one another. He takes the number of such blocs as exogenous. Each bloc sets an external tariff that maximizes the joint welfare of its members, given the tariffs set by other blocs. In these circumstances, the height of each country's tariff grows with the size of the typical bloc. Consolidation of the world into larger blocs has offsetting effects on welfare in the typical country; between-bloc trade distortions grow monotonically larger with bloc size, but so does the fraction of world trade that takes place within blocs. Welfare is highest when the entire world comprises one bloc, but also is high when the world has small blocs that have little monopoly power and therefore impose low tariffs. In between, welfare is a nonmonotonic function of bloc size. Krugman notoriously found that, for many values of the elasticity of substitution, welfare is minimal when the trading systems comprises three symmetric blocs.

The findings in Krugman (1991a) rely heavily on the assumption that countries are symmetric and form their blocs arbitrarily. In a follow-up paper, Krugman (1991b) discusses informally a case with "natural trading blocs." In this setting, geography or other considerations give certain groups of countries a greater predilection to trade with one another than with those outside the group. If blocs form "naturally" among groups of countries that trade intensively, the free movement of goods within blocs will cover a majority of world trade and the external trade barriers will apply to a small volume of trade. In the limit, with very high costs of trade outside a natural grouping, the formation of trading blocs must be raise welfare for all involved.[af]

5.3 Political Incentives for Regional or Preferential Agreements

In addition to the potential improvements to allocative efficiency and to external terms of trade that can motivate countries to form preferential trade agreements such as FTAs and customs unions, there are political forces that can explain this outcome.

Grossman and Helpman (1995a) examine the political conflict between supporters and opponents of an FTA in a model with industry campaign contributions. There are two small countries that initially have the tariffs predicted by the lobbying model

[af] In another extension of Krugman (1991a), Bond and Syropoulos (1996) examine more fully the relationship between bloc size and the market power exerted by trading blocs, by allowing for blocs of different sizes and by allowing for alternative endowment structures.

of Grossman and Helpman (1994). Each can trade with the rest of the world at fixed terms of trade. The countries have an opportunity to eliminate tariffs on their internal trade while retaining their initial external tariffs for imports from the rest of the world. Export industries offer contributions to encourage the government conclude an agreement in order to expand market access in the partner country. Import-competing interests offer contributions to discourage the agreement, in order to protect their local markets. Each government chooses a stance that maximizes a weighted sum of aggregate welfare and contributions. A FTA is politically viable if and only if it is favored by both governments. The question is, Under what conditions is an FTA most likely to form?

An FTA has no effect on industries in which both countries export to the rest of the world. So, Grossman and Helpman focus on industries in which at least one of the member countries has positive imports in the initial equilibrium. They distinguish two possible outcomes for such an industry once an FTA comes into being. First, the importer might expand its imports from its partner while continuing to import from the rest of the world. In this case, the internal price in the importing country remains equal to the world price augmented by the MFN tariff. There is no expansion of its total imports, only diversion of trade from the rest of world to the FTA partner. Second, the importing country might cease its imports from the rest of the world and import only from its FTA partner. In this case, the price falls in the importing country and total imports expand. The internal price in the exporting member of the FTA need not rise very much and might not rise at all. Grossman and Helpman refer to the former outcome as one with *enhanced protection*, which comes hand in hand with trade diversion. The latter outcome is one with *reduced protection*, together with trade creation.

Now consider the politics. In the case of enhanced protection, the industry in the exporting country gains producer surplus from preferential access to the higher internal prices in the partner country. Meanwhile, the import-competing industry suffers no losses, as the internal price there remains as before. On net, such industries contribute to political viability; the export interests lobby in favor of the agreement and the import-competing interests have no reason to oppose. In contrast, with reduced protection, the exporting interests stand to gain little or nothing, while the import-competing interests in the potential partner face the prospect of falling local prices. The existence of such industries bolster opposition to a potential agreement in one country with only modest (or no) support generated in the other. Overall, an FTA will be viable if most industries have the potential for enhanced protection and not so if most face the prospect of reduced protection. But trade creation enhances economic efficiency in the small countries, whereas trade diversion detracts from an efficient allocation of resources. This leads Grossman and Helpman to conclude that an FTA is most likely to be politically viable when it is also economically harmful. Krishna (1998) comes to much the same conclusion in a model with oligopolistic trade in which trade politics are driven entirely by effects on firms' profits.

However, Ornelas (2005a) provides a counterargument. He considers an economy much like the one in Krishna (1998), with quasilinear utility, constant returns production of a numeraire good, and oligopolistic competition in a second sector, with fixed numbers of firms in each country and segmented markets. But whereas Krishna assumes that MFN tariffs remain the same after any agreement is signed, Ornelas allows the FTA members to adjust their external tariffs. He shows that, in this setting, they have incentive to *reduce* their tariffs vis-à-vis nonmembers, for two reasons.[ag] First, a government's incentive to set high tariffs in order to extract profits from firms in nonmember countries is mitigated by the FTA, because firms in partner countries with improved access to the local market capture more of the profits that are extracted from the outsiders. Second, the political support for high tariffs vis-à-vis nonmember countries is reduced by the formation of an FTA, because politically active firms lose sales in their home market to rivals in partner countries, and so have less incentive to lobby for protection from outsiders. In the economic environment considered by Ornelas, the fall in external tariffs that results from an FTA always is deep enough to generate net trade creation with nonmember countries. Moreover, in his model, an FTA can be politically viable only if it is efficiency enhancing. Since the FTA reduces political contributions by local firms, it can be attractive to the policy makers only if aggregate welfare rises to more than compensate.

When FTAs cause trade diversion from nonmembers, the politics of preferential trade can be self-reinforcing, a phenomenon that Baldwin (1993) termed "domino regionalism." He considers a world with many potential members of a customs union. Each government's stance toward joining the union reflects an internal trade-off between consumer welfare and industry interests, but with an additional term that reflects the country's exogenous "resistance" (positive or negative) to being a member of the club. Countries differ in this regard, and those that are least resistant are the first to join. The economies produce a homogeneous good and a fixed set of differentiated varieties. The former good is traded freely, whereas the latter goods are traded subject to iceberg trading costs. These costs are higher for firms outside the region than for those that are potential members of the customs union. The political contributors in this setting are the owners of the differentiated varieties.

Baldwin imagines an initial equilibrium in which all countries in the region with resistance below some critical level are members of the union and those with greater resistance are not. Then, there is an exogenous shock that reduces trade costs within the region. The decline in within-group trade costs causes additional governments to apply for membership. But, as membership expands, firms in nonmember countries intensify their lobbying, as the potential profits for insiders expand at the expense of profits for outsiders. An initial round of union expansion alters the political balance in nonmember countries,

[ag] Richardson (1993) shows that countries also have an incentive to reduce their external tariffs after forming an FTA when markets are perfectly competitive.

such that further expansion follows. The ultimate growth in regionalism, then, is a multiple of what one might expect from the initial shock. In short, Baldwin argues that the trade diversion associated with preferential trade agreements can be politically contagious.[ah]

But once again, the conclusion can be different if an FTA results in a reduction of external tariffs that promotes trade with nonmember countries. Ornelas (2005b) considers one such setting, using the same economic model as in Krishna (1998) and Ornelas (2005a). The fact that firms in nonmember countries benefit from the reduction in external tariffs of the FTA partners without having to reduce their own tariffs means that they may have less incentive than before to enter into agreements with the member countries, either as partners to a PTA or as signatories to a multilateral agreement.

5.4 PTAs as Stepping Stones to Multilateral Free Trade

Another potential purpose of preferential trade agreements is to facilitate a process of multilateral trade liberalization. The literature identifies circumstances under which PTAs serve as "building blocks" for global free trade; that is, a multilateral agreement that implements free trade becomes achievable as an equilibrium outcome in a dynamic game when PTAs are negotiated along the way, when such an outcome would not be possible in a negotiation game that precludes preferential agreements. Of course, PTAs can also represent "stumbling blocks" in some circumstances; that is, they can impede or prevent the achievement of global free trade in situations where a multilateral agreement would emerge as an equilibrium if discriminatory trade were prohibited.[ai] Indeed, the Ornelas (2005b) paper mentioned above provides one such example. But, in such cases, one would not typically regard the "purpose" of the PTA as being to interfere with multilateral negotiations; rather, the impediment to free trade would be seen as an unintended consequence. Since this chapter deals with the purpose of trade agreements and not their unintended consequences, I will not review further the research that describes PTAs as stumbling blocks.

[ah] Chen and Joshi (2010) examine how the existence of FTAs affects the incentives that countries have to form new agreements. They consider a model with three countries and two goods. The numeraire good is competitively produced while the other good is produced by one firm in each country. Preferences are quasilinear, demands for the oligoply good are linear, and the competition features Cournot behavior with segmented markets. In this setting, if one country in a pair has an existing FTA with the third country but the other does not, then the existing FTA strengthens the incentive for the member to form another FTA but weakens the incentive for the nonmember country to do so, compared to a benchmark with no pre-existing FTA. However, if both potential members of a new FTA participate in pre-existing agreements with the third country, the incentives that both have to conclude a new agreement are inevitably strengthened.

[ai] The terminology of building blocks and stumbling blocks was first introduced by Bhagwati (1991).

One way in which a PTA may facilitate the achievement of global free trade is by raising the cost of being left out. Saggi and Yildiz (2010) illustrate this idea in a three-country, three-good, endowment model with linear demands. Each country is endowed with two goods that it exports to the other two. It has no endowment of the third good, which it imports from its two trade partners. In a dynamic game that allows for "bilateralism," each country announces in a first stage the names of any trade partners with whom it is willing to engage in mutual free trade. In the second stage, the countries set external tariffs for any and all countries with whom they have not entered into an agreement in the first stage. In a game that precludes bilateralism, the countries can only agree to liberalize trade on a multilateral basis. If any country declines to do so, then the noncooperative tariffs ensue.[aj]

Saggi and Yildiz consider first a situation in which the countries are symmetric with respect to their endowment levels. In such circumstances, global free trade is the unique stable equilibrium under both bilateralism and multilateralism. Accordingly, with symmetry, PTAs have no role to play in the achievement of global free trade in their model; a multilateral agreement is reached even if bilateral agreements are not allowed. But the conclusion can be different with asymmetric countries. Then, global free trade emerges as a stable equilibrium for a wider set of parameter values with bilateralism than with multilateralism. In other words, for some endowment combinations, a multilateral agreement can be reached only if bilateral agreements represent a permissible alternative.

The finding is readily understood. Consider a country that fares better in a Johnson-like Nash equilibrium than under global free trade. In a multilateral process, such a country would decline to name any of its partners in the negotiation stage. By submitting a blank sheet, it could effectively block an agreement and achieve its first best. Suppose instead that bilateral agreements are possible. The country that fares best in the noncooperative equilibrium might not have this outcome as a viable option. If the other two countries can gain by forming a bilateral block compared to universal noncooperation, then the relevant comparison for the third country is between a multilateral agreement and a world with an PTA from which it is excluded. Since trade diversion within the member countries would hurt the excluded country, it might well prefer the multilateral agreement to being the one country on the outside.[ak]

[aj] Saggi et al. (2013) compare bilateralism and multilateralism in a model in which countries are free to form customs unions but not FTAs. They show that customs unions, unlike FTAs, can prevent the achievement of global free trade, because two countries may prefer to exclude the third country from mutual free trade in order to exploit their joint market power. The incentive for exclusion is stronger under a customs union than under a FTA, so the former can be a stumbling block but not a building block for multilateral trade liberalization in their setting.

[ak] Saggi and Yildiz (2011) reach much the same conclusion in a model with a different economic setting. There they consider a world economy with Cournot competition in segmented markets and one firm per country. In a symmetric setting in which all three firms have the same costs, global free trade emerges as the unique stable equilibrium with bilateral agreements are possible or not. However, the possibility of bilateral agreements is necessary for the achievement of global free trade in a setting in which one firm has much higher production costs than the others.

In Saggi and Yildiz (2010), the possibility of a PTA facilitates the conclusion of a multilateral agreement, but no PTAs need to form for global free trade to be realized. This is an inevitable consequence of the game structure, wherein either a bilateral agreement or a multilateral agreement is signed, but not both. Aghion et al. (2007) introduce a sequential structure in which PTAs may actually form along the equilibrium path. They consider a trade-negotiation process in reduced form, with the details of the economic environment suppressed. In particular, they specify payoffs for each of three countries under all possible coalition structures: with each country alone; with all combinations of bilateral agreements; and with a multilateral agreement. They allow for utility transfers between coalition members, so the relevant payoffs are those that accrue jointly to the parties to any trade agreement. The bargaining protocol features an agenda setter that makes take-it-or-leave offers to the other two countries. In a multilateral negotiation, the offers to engage in free trade are made simultaneously to the others and they can accept or reject. With bilateralism, the agenda setter makes the offers sequentially, in whatever order it prefers. If the first to receive an offer accepts, a bilateral agreement is formed. If the second also accepts, the agreement becomes multilateral. Aghion et al. (2007) ask, when does the agenda setter choose the sequential process? And, when is a sequential process necessary to achieve global free trade?

The authors distinguish several cases. If the payoffs exhibit "grand coalition superadditivity," then the sum of the payoffs under global free trade exceeds the sum of the payoffs under any other coalition structure. Grand coalition superadditivity necessarily applies in a neoclassical economy with welfare-maximizing governments. But it need not apply if there is imperfect competition in world markets, or if the governments have political motivations. With or without grand coalition superadditivity, there can be positive or negative coalition externalities. Payoffs are characterized by positive coalition externalities when the welfare of a country that is excluded from a PTA is higher than it would be without the bilateral agreement. Negative coalition externalities imply just the opposite.

With grand coalition superadditivity and transferable utility, the agenda setter always prefers a multilateral agreement to any alternative. Moreover, it can achieve such an outcome by offering payoffs to the others that make them indifferent between joining and not. The question remains, How do the structure of coalition externalities affect the agenda sender's choice between simultaneous and sequential negotiation? Aghion et al. show that, if at least one coalition externality is negative, the agenda setter will opt to negotiate sequentially. First it will form an agreement with the country (if any) that faces a positive externality. Then it will invite the participation of the remaining country. By the time the second country receives its offer, its fallback position is the lower level of welfare that it would suffer if excluded from a bilateral pact. So, the agenda setter can extract surplus from this country by confronting it with an inferior default option.

Even if the payoffs do not satisfy grand coalition superadditivity, a sequential negotiation process might lead to global free trade. A multilateral agreement can be reached

even at a cost in terms of the collective (economic or political) welfare of the three countries under conditions that the authors describe. The outcome becomes possible, because the agenda setter's option to form a bilateral agreement first with one country allows it to extract enough surplus from the other that it is willing to begin these negotiations in preference to the *status quo*. More specifically, let *C* represent the agenda setter and *A* and *B* represent the other two countries. Suppose *C* prefers the *status quo* with no trade agreements to global free trade. In a simultaneous, multilateral negotiation, *C* must offer *A* and *B* their *status quo* payoffs. If the sum of the three status quo payoffs exceeds the joint welfare under a multilateral trade agreement, then *C* will not make any offers under multilateral bargaining, and the *status quo* will prevail. But suppose *C* has already negotiated a bilateral agreement with *A* and that the coalition externalities for *B* are negative. Then *C* will not need to offer *B* as much as its status quo level of welfare in order to induce it to join the existing agreement. With the lesser payment that is needed, *C* might prefer to have *B* join once an agreement with *A* has been established. Moreover, *C* might prefer this outcome (with the smaller payment to *B*) than what it can attain in the *status quo*. In these circumstances, *C* will approach *A* first and then *B*, and the bilateral agreement with *A* serves as a building block for a multilateral agreement with both *A* and *B*.[al] The authors develop in their appendix an example of an economic setting with the requisite payoff structure.

6. CONCLUSION

This chapter has reviewed research that asks about the purpose of trade agreements. Why do governments willingly give up their sovereign rights to set trade policies and enter into agreements that restrict their choices? A unifying theme in much of the literature is that they do so in order to help internalize international externalities. No matter what the governments' objectives, be they the welfare of the aggregate polities or of particular constituent groups, or even the interests and well-being of the politicians themselves, interdependence in the trading system implies that any government's actions affect outcomes abroad. Each government would like others to take its concerns into account when setting policy. The only way to secure such an outcome is to display a willingness to take account of other governments' concerns when setting ones own policies.

The literature has usefully identified a number of international externalities that can arise in different market settings and with different political institutions. Less fruitful, in my opinion, have been the efforts to pin labels on these externalities. Is the fundamental purpose of the agreement to eliminate manipulation of the terms of trade or to ensure that

[al] An additional condition is needed to ensure that *C* prefers to approach *A* for a bilateral agreement before negotiating with *b*. Of course, there are analogous conditions on the payoffs for which multilateral free trade is achieved after an initial bilateral agreement between *C* and *B*.

domestic exporters are granted satisfactory access to foreign markets? It is not clear to me why this distinction is important, so long as we understand that noncooperative policy setting gives rise to inferior outcomes. We do want to understand why trade agreements have the features they do and what provisions must be added or modified to generate efficient outcomes and mutual gain. In my opinion, future research effort would be more productively spent understanding and improving the design of trade agreements (ie, extending the literature reviewed in Chapter 8) than in worrying about the words that best describe the purpose of trade agreements.

ACKNOWLEDGMENTS

I am grateful to Daniel Goetz for excellent research assistance and to Kyle Bagwell and Bob Staiger for comments on an earlier draft. I am also grateful to Henrik Horn, with whom I have discussed many of these issues before (see Grossman and Horn, 2013).

REFERENCES

Aghion, P., Antràs, P., Helpman, E., 2007. Negotiating free trade. J. Int. Econ. 73 (1), 1–30.
Antràs, P., Staiger, R.W., 2012. Offshoring and the role of trade agreements. Am. Econ. Rev. 102 (7), 3140–3483.
Ashley, P., 1904. Modern Tariff History: Germany, United States, France. John Murrary, London.
Aumann, R.J., 1985. Repeated games. In: Feiwel, G.R. (Ed.), Issues in Contemporary Microeconomics and Welfare. State University of New York Press, Albany, NY.
Bagwell, K., Staiger, R.W., 1999. An economic theory of GATT. Am. Econ. Rev. 89 (1), 215–248.
Bagwell, K., Staiger, R.W., 2001. Domestic policies, national sovereignty and international economic institutions. Q. J. Econ. 116 (2), 519–562.
Bagwell, K., Staiger, R.W., 2002. The Economics of the World Trading System. MIT Press, Cambridge, MA.
Bagwell, K., Staiger, R.W., 2006. Will international rules on subsidies disrupt the world trading system? Am. Econ. Rev. 96 (3), 877–895.
Bagwell, K., Staiger, R.W., 2012a. Profit shifting and trade agreements in imperfectly competitive markets. Int. Econ. Rev. 53 (4), 1067–1104.
Bagwell, K., Staiger, R.W., 2012b. The economics of trade agreements in the linear Cournot delocation model. J. Int. Econ. 88 (1), 32–46.
Bagwell, K., Staiger, R.W., 2015. Delocation and trade agreements in imperfectly competitive markets. Res. Econ. 69 (2), 132–156.
Bagwell, K., Staiger, R.W., 2016. The design of trade agreements. In: Bagwell, K., Staiger, R. (Eds.), Handbook of Commercial Policy, vol. 1A. North Holland, Amsterdam, pp. 435–529.
Baldwin, R.E., 1993. A Domino theory of regionalism. National Bureau of Economic Research, Cambridge, MA. NBER Working Paper 4465.
Baldwin, R.E., Venables, A.J., 1995. Regional economic integration. In: Grossman, G.M., Rogoff, K. (Eds.), Handbook of International Economics, vol. 3. North-Holland, Amsterdam.
Barrett, S., 2005. The theory of international environmental agreements. In: Mäler, K.G., Vincent, J.R. (Eds.), Handbook of Environmental Economics, vol. 3. North Holland, Amsterdam.
Bhagwati, J., 1991. The World Trading System at Risk. Princeton University Press, Princeton, NJ.
Bhagwati, J., 1995. U.S. trade policy: the infatuation with free trade areas. In: Bhagwati, J.N., Krueger, A.O. (Eds.), The Dangerous Drift to Preferential Trade Agreements, The AEI Press, Washington, D.C.
Bond, E.W., 1990. The optimal tariff structure in higher dimensions. Int. Econ. Rev. 31 (1), 103–116.

Bond, E.W., Park, J.H., 2002. Gradualism in trade agreements with asymmetric countries. Rev. Econ. Stud. 69 (2), 379–406.
Bond, E.W., Syropoulos, C., 1996. The size of trading blocks: market power and world welfare effects. J. Int. Econ. 40, 412–437.
Brander, J., Spencer, B., 1984a. Trade warfare: tariffs and cartels. J. Int. Econ. 16, 227–242.
Brander, J., Spencer, B., 1984b. Tariff protection and imperfect competition. In: Kierzkowski, H. (Ed.), Monopolistic Competition and International Trade, Clarendon Press, Oxford.
Brander, J., Spencer, B., 1985. Export subsidies and market share rivalry. J. Int. Econ. 18, 83–100.
Brou, D., Ruta, M., 2013. A commitment theory of subsidy agreements. The B.E. J. Econ. Anal. Policy 13 (1), 239–270.
Chen, M.X., Joshi, S., 2010. Third-country effects on the formation of free trade agreements. J. Int. Econ. 82 (2), 238–248.
Destler, I.M., 2005. American Trade Politics, fourth ed. Institute for International Economics, Washington, D.C.
Dixit, A.K., 1985. Tax policy in open economies. In: Auerbach, A., Feldstein, M.J. (Eds.), Handbook of Public Economics, vol. 1. North Holland, Amsterdam, pp. 313–374.
Dixit, A.K., 1987. Strategic aspects of trade policy. In: Bewley, T.F. (Ed.), Advances in Economic Theory: Fifth World Congress, Cambridge University Press, Cambridge, UK.
Dixit, A.K., Norman, V.D., 1986. Gains from trade without lump-sum compensation. J. Int. Econ. 21, 111–122.
Eaton, J., Grossman, G.M., 1986. Optimal trade and industrial policy under oligopoly. Q. J. Econ. 101 (2), 383–406.
Ethier, W.J., 2007. The theory of trade policy and trade agreements: a critique. Eur. J. Polit. Econ. 23, 605–623.
Ethier, W.J., Kim, N., 2013. The trade-agreement embarrassment. J. East Asian Econ. Integr. 17, 243–260.
Fukushima, K., Kim, N., 1989. Welfare improving tariff changes: a case of many goods and many countries. J. Int. Econ. 26, 383–388.
Graaf, J., 1949. On optimum tariff structures. Rev. Econ. Stud. 17, 47–59.
Grossman, G.M., Helpman, E., 1994. Protection for sale. Am. Econ. Rev. 84 (4), 833–850.
Grossman, G.M., Helpman, E., 1995a. The politics of free-trade agreements. Am. Econ. Rev. 85 (4), 667–690.
Grossman, G.M., Helpman, E., 1995b. Trade wars and trade talks. J. Polit. Econ. 103 (4), 675–708.
Grossman, G.M., Helpman, E., 1996. Electoral competition and special interest politics. Rev. Econ. Stud. 63 (2), 265–286.
Grossman, G.M., Horn, H., 2013. Why the WTO? An introduction to the economics of trade agreements. In: Horn, H., Mavroidis, P.C. (Eds.), Legal and Economic Principles of World Trade Law, Cambridge University Press, Cambridge, UK.
Grossman, G.M., Lai, E.L.-C., 2004. International protection of intellectual property. Am. Econ. Rev. 94 (5), 1635–1653.
Hillman, A.L., Moser, P., 1995. Trade liberalization as politically optimal exchange of market access. In: Canzoneri, M., Ethier, W.J., Grilli, V. (Eds.), The New Transatlantic Economy, Cambridge University Press, Cambridge, UK.
Hillman, A.L., Long, N.V., Moser, P., 1995. Modeling reciprocal trade liberalization: the political-economy and national welfare perspectives. Swiss J. Econ. Stat. 131, 503–515.
Hudec, R.E., 1993. 'Circumventing' democracy: the political morality of trade negotiations. Int. Law Polit. 25, 311–322.
Johnson, H.G., 1953. Optimum tariffs and retaliation. Rev. Econ. Stud. 21 (2), 142–153.
Katrik, H., 1977. Multi-national monopolies and commercial policy. Oxf. Econ. Pap. 29 (2), 283–291.
Kemp, M., 1967. Notes on the theory of optimal tariffs. Econ. Record 43, 395–403.
Kemp, M., Wan Jr., H., 1976. An elementary proposition concerning the formation of customs unions. In: Kemp, M. (Ed.), Three Topics in the Theory of International Trade: Distribution, Welfare and Uncertainty, North Holland, Amsterdam.
Kennan, J., Riezman, R., 1990. Optimum tariff equilibria with customs unions. Can. J. Econ. 23 (1), 70–93.

Krishna, P., 1998. Regionalism versus multilateralism: a political economy approach. Q. J. Econ. 113 (1), 227–251.
Krishna, P., 2005. The economics of preferential trade agreements. In: Choi, E.K., Harrigan, J. (Eds.), Handbook of International TradeBlackwell, Oxford.
Krishna, P., Bhagwati, J., 1997. Necessarily welfare improving customs unions with industrialization constraints: the Cooper-Massell-Johnson-Bhagwati proposition. Jpn. World Econ. 9, 441–446.
Krugman, P.R., 1980. Scale economies, product differentiation, and the pattern of trade. Am. Econ. Rev. 70 (5), 950–959.
Krugman, P.R., 1991a. Is bilateralism bad? In: Helpman, E., Razin, A. (Eds.), International Trade and Trade Policy, MIT Press, Cambridge, MA, pp. 9–23.
Krugman, P.R., 1991b. The move toward free trade zones. In: Policy Implications of Trade and Currency Zones: A Symposium Sponsored by the Federal Reserve Bank of Kansas CityFederal Reserve Bank of Kansas City, Kansas City, MO.
Kydland, F.E., Prescott, E.C., 1977. Rules rather than discretion: the inconsistency of optimal plans. J. Polit. Econ. 85 (3), 473–492.
Limão, N., 2016. Preferential trade agreements. In: Bagwell, K., Staiger, R. (Eds.), Handbook of Commercial Policy, vol. 1B. North Holland, Amsterdam, pp. 279–367.
Limão, N., Tovar, P., 2011. Policy choice: theory and evidence from commitment via international trade agreements. J. Int. Econ. 85 (2), 186–205.
Maggi, G., 1999. The role of multilateral institutions in international trade cooperation. Am. Econ. Rev. 89 (1), 190–2014.
Maggi, G., 2014. International trade agreements. In: Gopinath, G., Helpman, E., Rogoff, K. (Eds.), Handbook of International Economics, vol. 4. North Holland, Amsterdam.
Maggi, G., Rodriguez-Clare, A., 1998. The value of trade agreements in the presence of political pressures. J. Polit. Econ. 106 (3), 574–601.
Maggi, G., Rodriguez-Clare, A., 2007. A political-economy theory of trade agreements. Am. Econ. Rev. 97 (4), 1374–1406.
Mayer, W., 1981. Theoretical considerations on negotiated tariff adjustments. Oxf. Econ. Pap. 33 (1), 135–153.
Mitra, D., 1999. Endogenous lobby formation and endogenous protection: a long-run model of trade policy determination. Am. Econ. Rev. 89 (5), 1116–1134.
Mitra, D., 2002. Endogenous political organization and the value of trade agreements. J. Int. Econ. 57 (2), 473–485.
Ohyama, M., 1972. Trade and welfare in general equilibrium. Keio Econ. Stud. 9, 37–73.
Ornelas, E., 2005a. Endogenous free trade agreements and the multilateral system. J. Int. Econ. 67 (2), 471–497.
Ornelas, E., 2005b. Trade creating free trade areas and the undermining of multilateralism. Eur. Econ. Rev. 49 (7), 1717–1735.
Ossa, R., 2011. A new trade theory of GATT/WTO negotiations. J. Polit. Econ. 119 (1), 122–152.
Panagariya, A., 2000. Preferential trade liberalization: the traditional theory and new developments. J. Econ. Lit. 38 (2), 287–331.
Panagariya, A., Krishna, P., 2002. On necessarily welfare-enhancing free trade areas. J. Int. Econ. 57, 353–367.
Park, J.H., 2011. Enforcing international trade agreements with imperfect private monitoring. Rev. Econ. Stud. 78 (3), 1102–1134.
Pauwelyn, J., 2008. New trade politics for the 21st century. J. Int. Econ. Law 11 (3), 559–573.
Regan, D.H., 2006. What are trade agreements for? Two conflicting stories told by economists, with a lesson for lawyers. J. Int. Econ. Law 9, 951–988.
Regan, D.H., 2015. Explaining trade agreements: the practitioners' story and the standard model. World Trade Rev. 14 (3), 391–417.
Richardson, M., 1993. Endogenous protection and trade diversion. J. Int. Econ. 34 (3-4), 309–324.
Richardson, M., 1995. On the interpretation of the Kemp/Wan theorem. Oxf. Econ. Pap. 47 (4), 696–703.
Saggi, K., Woodland, A., Yildiz, H.M., 2013. On the relationship between preferential and multilateral trade liberalization: the case of customs unions. Am. Econ. J. Microecon. 5 (1), 63–99.

Saggi, K., Yildiz, H.M., 2010. Bilateralism, multilateralism, and the quest for global free trade. J. Int. Econ. 81, 26–37.
Saggi, K., Yildiz, H.M., 2011. Bilateral trade agreements and the feasibility of multilateral free trade. Rev. Int. Econ. 19 (2), 356–373.
Staiger, R.W., Tabellini, 1987. Discretionary trade policy and excessive protection. Am. Econ. Rev. 77 (5), 823–837.
Svedberg, P., 1979. Optimal tariff policy on imports from multinationals. Econ. Record 55 (1), 64–67.
Turunen-Red, A.H., Woodland, A.D., 1991. Strict Pareto-improving reform of tariffs. Econometrica 59, 1127–1152.
Venables, A.J., 1985. Trade and trade policy with imperfect competition: the case of identical products and free entry. J. Int. Econ. 19, 1–20.
Venables, A.J., 1987. Trade and trade policy with differentiated products: a Chamberlinian-Ricardian model. Econ. J. 97, 700–717.