CHAPTER 10

PREFERENTIAL TRADE AGREEMENTS AND REGIONAL INTEGRATION

BOTH the GATT and the GATS make explicit allowance for preferential trade agreements among a subset of members. Such agreements can be of two types: reciprocal and nonreciprocal. This chapter deals with the former; the latter are discussed in Chapter 12 as they arise in trade relations between industrialized and developing countries. Both types of preferential trade are inconsistent with the MFN principle and are therefore subject to multilateral disciplines that define minimum conditions that must be met for an agreement. The WTO also provides for multilateral scrutiny of such agreements. This chapter discusses the rationales for preferential trade agreements (PTAs) between WTO members, the WTO rules and their application in practice, and the economic literature exploring the relationship between PTAs and multilateralism (the trading system)—both theoretical and empirical. Given the steadily expanding number of PTAs, a critical question for the WTO is whether the network of PTAs create incentives to lower trade barriers on a MFN basis and thus help achieve a major objective of the drafters of the GATT.

Reciprocal trade agreements among subsets of the WTO membership have become a prominent part of the trade landscape. As of late 2007, 380 PTAs had been notified to the GATT/WTO. Of that number, 300 agreements were notified under Article XXIV of the GATT, 22 agreements involving developing countries

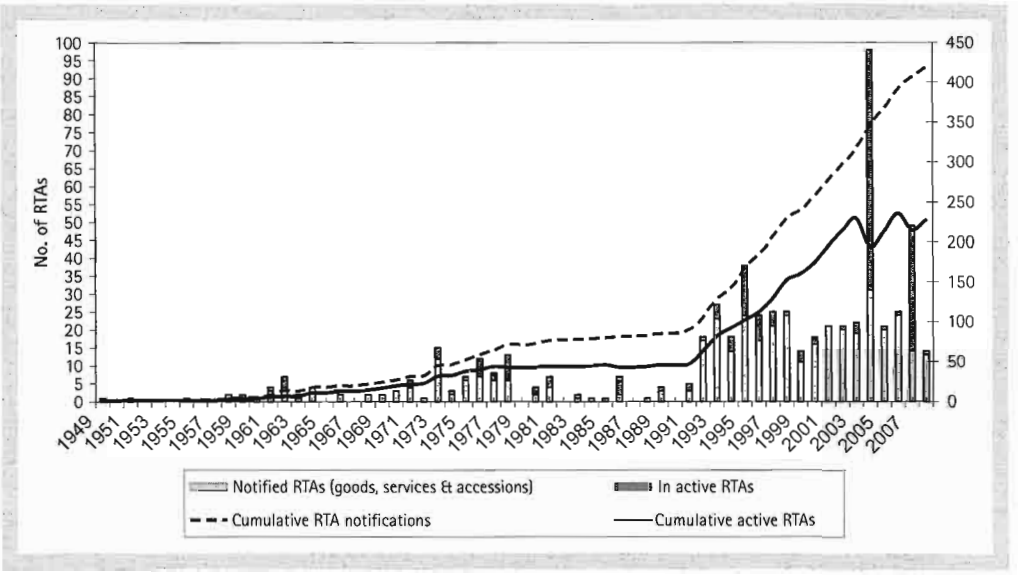

Fig. 10.1. Number of preferential trade agreements (1948–2007)

Source: WTO.

were notified under the Enabling Clause, and 58 under Article V of the GATS. Although these numbers are suggestive of a proliferation of PTAs, it is important to recognize that a large number of the PTAs notified to the GATT and WTO have involved prospective members of the EU and became irrelevant once the countries acceded to the EU.[1] About 200 PTAs were in force at the end of 2007 (Figure 10.1). Of these PTAs, customs unions account for less than 10 per cent. Many of the PTAs involve contiguous countries but many do not. In this chapter we reserve the term 'regional integration' for PTAs limited to neighbouring countries.

Since the late 1950s, the EC has been the market leader in the PTA business. European countries account for more than half of all PTAs notified to the WTO and that were still in force in 2008. The major regional grouping in Europe is the European Union, with 27 members in 2008. Other European PTAs include the European Free Trade Association (EFTA, Iceland, Lichtenstein, Norway and Switzerland).[2] and the Central European Free Trade Agreement. The EU has concluded

[1] As noted by Pomfret (2007), after the expansion of the EU by ten new members in 2004, some 65 PTAs between these countries and the EU became redundant. Note also that the numbers overstate the prevalence of PTAs because separate notifications are required under the GATT and the GATS for agreements that cover both goods and services—as many PTAs now do.

[2] At varying points in time EFTA also included Austria, Denmark, Finland, Portugal, Sweden and the UK, all of which left to become members of the EU. The European Free Trade Association has a very close economic relationship with the EU, governed by the European Economic Area Agreement.

Stability and Association Agreements with countries in south-eastern Europe, has PTAs with almost all Mediterranean nations, Chile, Mexico and South Africa, and a series of Economic Partnership Agreements (EPAs) with African, Caribbean and Pacific (ACP) countries. The EU has expressed interest in negotiating agreements with India, Korea, and the members of the Association of South-East Asian Nations (ASEAN), and the Central American Common Market, among others (WTO, 2007).

The US became a proponent of PTAs in the 1980s, starting with agreements with Israel (in 1985) and Canada (1988), followed in 1992 by the North American Free Trade Agreement (NAFTA) with Canada and Mexico. Since the mid-1990s, the US has concluded PTAs with Australia, Bahrain, Chile, four Central American countries (El Salvador, Guatemala, Honduras and Nicaragua) and the Dominican Republic (DR-CAFTA), Jordan, Morocco and Singapore. As of mid-2008, PTAs with Peru and Oman had been ratified but were pending implementation, whereas PTAs with Colombia, Panama and South Korea were awaiting approval by the Congress.

Virtually all OECD nations are now a member of one or more PTAs. Examples include the Australia and New Zealand Closer Economic Relations PTA and EFTA, as well as bilateral agreements involving members of these agreements. The long standing exception among OECD countries used to be Japan. This changed starting in 2000, with a PTA with Singapore. Since then Japan concluded bilateral deals with other trading partners in Asia (Malaysia, Indonesia, the Philippines and Thailand) as well as Chile and Mexico. An agreement with ASEAN was signed in April 2008, and talks initiated with Australia, India, Switzerland and Vietnam.

As can be inferred from the above, developing countries have been active participants in the expansion of PTAs. Often they are the '*demandeurs*' for agreements with the EU and the US, as well as using PTAs as mechanisms to create larger regional markets. The Association of South-East Asian Nations and Mercosur (the Southern Common Market) spanning Argentina, Brazil, Paraguay and Uruguay are just two examples of PTAs between developing countries.

Pomfret (2007) distinguishes between three waves of preferential trade activity. The first was in the 1960s, following the creation of the EEC, which induced other European countries to create EFTA, and many developing countries to form blocs of their own—many of which were unsuccessful and collapsed. The second was initiated in the 1980s with the decision of the US to negotiate PTAs and the move by the EU to deepen its integration process through the Single Market programme and widen its membership to 15 countries (achieved in 1995). This was complemented by the negotiation or revitalization of PTAs between developing countries that were more outward oriented—indeed, preferential liberalization went hand-in-hand with (or followed) unilateral liberalization. In the early 2000s, a third wave started, led by East Asian countries, partly as a response to the financial crisis of the late 1990s and partly a reaction to the rapid increase in China's economic power.

The specifics of each PTA vary greatly from case to case, but they all generally have one thing in common: discrimination against other countries. Preferential

trade agreements take many forms. At their simplest they lower or remove tariffs on imports originating in partner countries. Many go beyond tariffs to also cover NTMs. More recent vintage PTAs often include services trade and investment. At their most far-reaching, PTAs encompass instruments of domestic economic regulation and political cooperation, and may represent a step towards nation-building or international federalism.

The economic literature distinguishes between different types of PTAs. The most 'shallow' form is a bilateral trade agreement. This may do nothing more than grant a country MFN and national treatment. These types of agreement can be relevant for non-WTO countries—an example is a 2005 agreement between Laos and the US. More generally the term bilateral trade agreement covers deals where the parties involved do not go to free trade but only cut tariffs by a certain percentage. Such agreements are illegal under the WTO for industrialized nations but may be concluded by developing countries (Chapter 12). More far-reaching is a free trade agreement (FTA), which removes tariffs on trade between member countries, with each country retaining its own tariff structure against outsiders. A customs union is a free trade area with common external trade policies. A common market is a customs union that also integrates factor markets, allowing for the free movement of labour (workers) and capital. Finally, an economic union is a common market that includes some degree of harmonization of national economic policies of member states. In practice, PTAs are often a combination of these ideal types. Many (indeed most) tend to fail to conform to these ideal types. For example, contingent protection often remains applicable to intra-bloc trade in PTAs, implying that internal trade is not truly free. Many PTAs have not lead to internal free trade because certain sectors or industries are excluded and because differences in the external tariffs across members imply the need to enforce rules of origin to prevent arbitrage activity between low and high tariff members. Many customs unions fail to put in place a common external trade policy, resulting in continued border controls that impede trade flows between PTA members. And numerous agreements give rise to cooperation on matters of economic policy, regulation or factor markets without ever having achieved free trade between the participating members.

10.1. Motivations for Preferential Trade Agreements

In one form or another, PTAs have been around for hundreds of years. There were proposals for the provinces of France to establish a customs union in 1664.

Customs unions were precursors to the creation of new states in, for example Germany (the Zollverein) and Italy. Although nation-building objectives have been a spur to regional integration, historically, as discussed in Chapter 1, the exercise of power has been much more important in this regard. The primary driving force behind PTAs has been mercantilist: access to export markets. Negotiation of PTAs also has had a cyclical dimension. The 1930s saw great fragmentation of the world trading system as governments sought to shelter domestic industries and safeguard access to supplies. Preferential trade agreements were pursued as an instrument through which to undo some of the costs of unilateral protectionism and currency blocs. After the Second World War, regional integration contributed to the political reconstruction of Europe through the implementation of the Benelux Customs Union (1947), the creation of the European Coal and Steel Community in 1952 and the more far-reaching EEC (in 1957).

The creation of the EEC stimulated 'copy cat' regionalism among developing countries in 1960s. The associated PTAs were mostly driven by a desire to apply import-substitution industrialization strategies within a larger economic area to realize scale economies and to achieve political objectives. The agreements tended to be very protectionist (with high external barriers) and interventionist (with governments pursuing industrial policies and influencing the location of specific industries within the territory of the PTA). They were generally failures. By the late 1970s the ineffectiveness of most of these PTAs had become evident. None seemed to have contributed much to economic development, some had collapsed and the strains of the debt crisis in the 1980s made many of those that survived largely moribund.

In the mid-1980s a change in attitudes towards international trade and competition began to occur. Unilateral liberalization decisions by many governments were complemented by a new wave of more open PTAs. The EU played a major role in the resurgence of PTAs by negotiating agreements with Central and Eastern European and Mediterranean nations, as well as continuing to expand its membership, starting with Portugal and Spain in 1986. In the Americas, the Canadian–US FTA of 1988 was extended to Mexico through NAFTA in 1994, Mercosur was formed in 1991. In Asia, ASEAN members extended 25 years of political and economic cooperation with a formal free trade agreement in 1992 and admitted new members, including Vietnam (1995), Laos and Myanmar (1997) and Cambodia (1999). In Africa, a number of initiatives were launched, including the Common Market for Eastern and Southern Africa (COMESA)—which extends as far north as Egypt—and the Southern African Development Community (SADC). Members of the Arab League revitalized their long standing but stalled integration efforts by creating a Pan-Arab Free Trade Agreement (PAFTA) to remove tariffs on intra-regional trade, which was implemented ahead of schedule in 2005.

Why go preferential?

Chapter 1 has already discussed the rationales that may induce governments to negotiate trade agreements. Recall that the two major explanations offered by economists revolve around improving market access for exporters (the terms of trade argument) and credibility (time consistency)—trade agreements may be used by governments as a commitment device. In principle a multilateral agreement that includes all countries should dominate one that is limited to only one or several partner countries: better access to the world market dominates what can be offered by just a few countries, and a multilateral set of disciplines should do more for credibility than a PTA. After all, having more countries to hold the government to account would seem to be a more powerful source of discipline than one that involves only a few countries.

So why engage in PTAs? How might these arrangements do a better job in delivering market access and credibility? An extensive literature has suggested numerous possibilities. The arguments can be split along three lines, two building on the 'standard' economic rationales, and one focused more on noneconomic factors, such as foreign policy and national security motivations.

'Better' market access. The WTO does not offer free trade. A FTA with a major trading partner does and on a *preferential basis.* Such preferred access is valuable, the more so the larger the partner market and the higher the barriers the partner maintains on imports from competing suppliers. It may be easier to get if a country is small and thus not perceived as much of a threat by import-competing firms in the partner country (countries). Moreover, as tariffs fall—whether because of unilateral decisions or trade agreements—the relative importance of NTMs as barriers to trade and market integration rise (Baldwin, 1970). Preferential trade agreements may offer better instruments than the WTO for traders to get governments to *deal with market segmenting nontariff policies* that prevent the benefits of tariff removal from being (fully) realized. This can include the removal of the threat of contingent protection, on a *de facto* if not *de jure* basis (Prusa, 2006). In addition to issues that are covered by the WTO, a PTA may also offer the opportunity to negotiate disciplines on policies is areas that are not (yet) included in the WTO. Thus, PTAs offer the possibility of *deeper integration of product markets* than may be on offer in the WTO.

Preferential trade agreements may also help *put pressure on members of the WTO* to do more or to do it faster. The US moved from active hostility to a proponent of PTAs in the mid-1980s in part because of dissatisfaction with the 1982 refusal of GATT partners to initiate a MTN that covered services trade (Schott, 1989). Another factor was the end of the Cold War, which reduced American willingness to accept the *opportunity cost of free riding* by other countries in the WTO and the costs to the US of preferential liberalization elsewhere in the world. Finally, there are so-called *domino effects* (Baldwin, 1995): as major trading powers create trade

blocs, incentives for excluded countries to seek similar trading relationships increase, because the costs of being a nonmember rise. Exclusion from a major PTA market (or confronting higher barriers and costs than do 'insiders') can change the political economy equilibrium in the excluded country—increasing the incentives for exporters to mobilize and pressure their governments to seek accession or negotiate a PTA with the large blocs.

Focal points for—and greater credibility of—domestic reform. In principle a trade agreement can be used as a commitment and signalling device by governments seeking to change expectations and lobbying behaviour of firms and interest groups. As discussed in Chapter 1, the theoretical (and practical) arguments for using trade agreements as a commitment device depend on there being a credible enforcement mechanism. In the WTO context this may not exist, especially for small countries that cannot affect the terms of trade and are too small to make it worthwhile to bring to the WTO court. Preferential trade agreements may offer stronger enforcement mechanisms, especially if private interests have direct access to courts or other tribunals and mechanisms, and thus create alternatives to the WTO to bring cases (Piérola and Horlick, 2007). A few PTAs have supranational enforcement mechanisms, the EU being the primary example, which can reduce uncertainty regarding implementation of the agreement. Enforcement is obviously also relevant for the market access incentives to negotiate a PTA—if seen as more effective, there is less uncertainty associated with the PTA than there is with the WTO. This is especially the case if the available remedies are stronger. Preferential trade agreements may also allow more credible commitments to be made if proximity of member countries reduces monitoring costs and similarity with partners—in terms of per capita income, etc.—reduces implementation costs.

Political objectives, regional cooperation and club goods. In addition to these rationales there may be 'noneconomic' foreign policy and national security objectives driving PTAs. Indeed, these often predominate in public discussion and debates, with any economic costs being argued to be the price of achieving the noneconomic objectives. Some problems or issues may be shared by only a limited number of (often neighbouring) countries, and therefore call for cooperation that is limited to the countries that will benefit from cooperation. Regional infrastructure such as bridges, railways and roads, power pools and electricity grid interconnection are examples of such 'club goods'. Interest in cooperation may extend to a willingness to engage in provision of financial transfers to support the delivery of regional public goods or achieve other objectives such as regional economic development. More ambitious forms of cooperation may extend to seeking to create a larger political entity. The German Zollverein—see Chapter 1—is a prominent historical example; the EU may become another. The collapse of Soviet hegemony allowed the countries of Eastern Europe and the Baltic to embrace democracy and market-based economic systems. Accession to the EU was seen by them as a tool to counter Russia's aspirations of a regional power, cement the

transition to a market economy and revive the common European cultural heritage. Less positively, PTAs may offer convenient 'displacement activity' for governments—providing an opportunity for photo opportunities and the appearance of strengthening relationships with partner countries without doing much if anything to liberalize trade. The activity may help achieve foreign policy objectives but has no economic meaning.

Other objectives that reflect a mix of market access and political goals are *precedent-setting and first-mover* objectives. Bilateral trade negotiations give the major trading powers an opportunity to establish certain rules of the game in areas that are not (yet) covered by the WTO, especially nontrade areas of policy, or to go beyond what is feasible multilaterally. Examples include IPRs, competition and investment policies, and disciplines relating to labour standards and environmental norms. Related is the *laboratory* rationale: countries may be able to experiment in a PTA context and 'discover' what works and what does not in ways that may not be possible in the WTO.

As discussed in Chapter 1, at any point in time a certain level of market access restrictions emerges as an equilibrium outcome of interactions between different groups in the political market of a country. In seeking better market access, export-oriented firms can push for liberalization at home (which will help exports as a result of the so-called Lerner symmetry proposition: a tax on imports is a tax on exports); try to convince foreign governments to liberalize (by forming coalitions with interests in the foreign country that would benefit from opening); push for multilateral liberalization (a MTN, or a sectoral deal such as the ITA—see Chapter 6); or push for a PTA.

A PTA by definition involves substantially fewer countries than a MTN. Indeed, many PTAs involve only two countries. This may make them easier to negotiate. In part this is because obtaining agreement to exclude certain sensitive policy areas may be more feasible than in a multilateral negotiation—especially if the PTA countries share similar 'sensitivities'. In part it may be because the set of possible policy packages that could make all parties better off is larger under a PTA, encompassing issues that could not appear on the negotiating agenda of a MTN. Issue linkage or side-payments therefore may be more feasible as the negotiation set expands, facilitating agreement. The side-payments may include mechanisms to transfer income from one member to another; in the MTN-context this is rarely possible.[3] Many PTAs involve relatively similar countries. The more similar are countries in their endowments and income levels, the likelier it is that intra-industry trade will be significant. This may facilitate preferential liberalization (Box 10.1).

[3] Although some steps in this direction were taken in the Doha Round with the Aid for Trade initiative—see Chapter 12.

Box 10.1. Intra-industry trade and PTAs

Increased internationalization of markets and technological advances put pressure on firms to seek greater efficiency through larger markets and improved access to foreign technologies and investment. Preferential trade agreements can help realize these objectives by providing cheaper access to intermediate inputs and facilitating the two-way cross-border movement of parts of a product for further processing. Such intra-industry trade may give rise to less adjustment pressures than inter-industry trade because jobs lost due to customers shifting to more efficient foreign suppliers may to a large extent be offset by the job-enhancing expansion in foreign demand for similar, differentiated goods produced domestically. The political opposition to liberalization that expands inter-industry trade may be stronger because industries that are less competitive than those abroad will generally be forced to contract substantially. This is not to say that intra-industry trade will not lead to adjustment and thus pressure for protection. Specialized and relatively immobile factors of production injured by import competition can be expected to seek protection. But the injury in this case is more at the firm than at the industry level. Other firms in the industry will expand. This makes it more difficult to maintain protection, as there will be conflicting interests within industries.

The relevance of this for PTAs is that intra-industry trade tends to be high among countries with similar endowments and relatively high per capita income levels—nations that have tended to form PTAs in the post-Second World War period. Levels of intra-industry trade between the members of the most successful PTAs—the EU, EFTA, NAFTA and the Australia–New Zealand Closer Economic Relations (CER) agreement—are high, both for trade in goods and trade in services (Egger, Egger and Greenaway, 2006).

Intra-industry trade has grown rapidly since the 1980s and is part and parcel of the global production sharing and specialization that was discussed in Chapter 1. It may also generate pressures on PTA member governments to liberalize more generally to facilitate the participation of national firms to slice up their production chains. As discussed later in this chapter, over time, once tariffs are removed on a preferential basis, the net benefits of a PTA may fall for firms, especially if the rules of origin constrain (raise the costs of) global sourcing. This may lead firms to support external liberalization (Baldwin, 2006*a*).

Much attention has been devoted by researchers to identifying the impacts of PTAs on members and nonmembers. Of obvious interest is whether PTAs are detrimental to world welfare, both in the short run (their impact effects) and in the longer run (taking induced growth effects into account), and whether they are building or stumbling blocks for multilateral cooperation in the WTO. The latter question—first posed in this way by Bhagwati (1991)—has generated numerous papers by economists, and is a question to which we return in Section 10.3.

The economic impact of PTAs on member and nonmember countries will depend on the type of agreement concerned (FTA, customs union or common market) and on the degree to which intra-regional trade is liberalized. The more extensive is internal liberalization, the greater the resulting increase in competition

on regional markets. Although this is welfare-enhancing for member countries—and presumably the object of economic integration—it may also be associated with greater adjustment pressures for inefficient industries located in member countries. The latter may attempt to shift some of the adjustment burden onto third countries by seeking increases in external barriers.

Regional integration can be detrimental to members and nonmembers by inducing a shift away from the most efficient supplier of goods or services. The formation of a trading bloc can give rise to so-called trade diversion (a shift from an efficient outside supplier to a higher cost regional one, induced by the elimination of tariffs on partner country producers that are higher cost than competitors on the world market, see Annex 2). Early observers of PTAs such as Hirschman (1981) and Tumlir (1983) pointed out that a PTA is likely to require trade diversion for political reasons, as this is an effective mechanism for compensating lobbies that oppose liberalization. Nonmembers may be harmed through investment diversion as well as trade diversion if enterprises decide to invest inside PTAs and produce locally, rather than export to the PTA. The prospect of EU enlargements, for example, has resulted in substantial increases in FDI into accession countries by both EU and non-European companies that are interested in serving the EU market. Thus, the 2004 enlargement of the EU to include Hungary, Poland, the Czech Republic, Slovakia, Slovenia and the Baltic States, followed by Bulgaria and Romania in 2007, encouraged investment by non-European firms as well as EU-based multinationals in both manufacturing and services sectors in these Eastern European countries.

Such investment responses may also arise in response to deepening of PTAs. Baldwin, Forslid and Haaland (1996) studied the effect of the announcement of the Single Market programme in the EU in the late 1980s, and concluded that this had a significant negative impact on inward FDI flows into EFTA countries. Foreign direct investment only recovered after the EFTA members applied for EU membership and the remaining EFTA countries concluded the European Economic Area agreement with the EU.

10.2. WTO Disciplines

The rules of the WTO do relatively little to limit the potential for trade and investment diversion. The WTO is somewhat schizophrenic about PTAs—permitting them subject to satisfaction of specific disciplines that, in principle, make things worse for nonmembers. The major provisions are Articles XXIV GATT and V GATS, complemented by multilateral surveillance of PTAs.

GATT Article XXIV: customs unions and FTAs

Article XXIV of the GATT allows FTAs and customs unions if:

(1) trade barriers after formation of the PTA do not rise on average (Article XXIV:5);
(2) all tariffs and other regulations of commerce are removed on substantially all intra-PTA exchanges of goods within a reasonable length of time (Article XXIV:8); and
(3) they are notified to the WTO Council.

The rationales for the first and last criteria are obvious. World Trade Organization members should be told when countries decide to pursue a PTA as otherwise the PTA members open themselves to dispute settlement and retaliation for violating the nondiscrimination rules. Nonmembers will not want to allow those forming a PTA use it as a means of raising protection against them. If restrictions on imports from nonmember economies are no higher than before, the extent of possible reductions in imports from nonmembers is reduced (although clearly not zero—the higher are external barriers, the greater the likely trade diversion). A practical problem faced by the drafters of Article XXIV was that the formation of a customs union by necessity requires changes in the external tariffs of some PTA members as they adopt a common external tariff. The rule that applies to customs unions is that duties and other barriers to imports from outside the union may not be on the whole higher or more restrictive than those preceding the establishment of the customs union (Article XXIV:5*a*). The interpretation of this phrase became a source of much disagreement among GATT contracting parties in the pre-WTO period. The rule for FTAs was unambiguous, however. Duties applied by each individual PTA member may not be raised (Article XXIV:5*b*).

The second condition imposed by Article XXIV is somewhat counterintuitive in that maximum preferential liberalization in itself is likely to be more detrimental to nonmembers than partial liberalization. Requiring it, however, ensures that countries are limited in their ability to violate the MFN obligation selectively. As noted by Finger (1993*b*), the rationale behind the second condition is a public choice one: it is an attempt to ensure that participants in PTAs go all the way and not to use the PTA as a mechanism to selectively pick and choose sectors. The determination of whether PTAs satisfy Article XXIV is the responsibility of the WTO Council. In the GATT period, the Council generally created a working party to establish if the conditions were satisfied by a notified PTA. Under the WTO, a Committee on Regional Trade Agreements (CRTA) has taken over this task.

The GATT experience in testing FTAs and customs unions against Article XXIV was very discouraging. Various aspects of the rules and their application, including enforcement of the requirement that PTAs be approved by the CONTRACTING PARTIES *before* they entered into force, proved unsatisfactory. Starting with the Treaty of Rome establishing the EEC in 1957, almost no examination of

PTAs notified under Article XXIV led to a unanimous conclusion or specific endorsement that all the agreements met the GATT requirements. As noted by the chairman of the working party on the 1989 Canada–United States Free Trade Agreement, commenting on the inability to reach a consensus, 'Over fifty previous working parties on individual customs unions or free trade areas have been unable to reach unanimous conclusions on the compatibility of these agreements with the GATT—on the other hand, no such agreement has been explicitly disapproved' (*GATT Focus*, December 1991). Only four GATT-era working parties were able to agree that a PTA satisfied the requirements of Article XXIV (Schott, 1989). Thus, GATT rules largely were a dead letter, although the consultations that occurred allowed interested nonmembers to express their concerns.

The reasons underlying this impotence go back to the creation of the EEC. A conscious decision was made by GATT contracting parties in the late 1950s not to closely scrutinize the formation of the EEC. The reason was that it was made clear by the EEC member states that a finding that the Treaty of Rome was inconsistent with Article XXIV would result in their withdrawal from GATT (Snape, 1993). Whether or not a serious threat, the consensus rule of the GATT ensured that even if a working party had concluded that the EEC did not satisfy Article XXIV, the EC members could block adoption of the report. In any event, to paraphrase Finger (1993*b*), at the end of the day the other GATT signatories blinked, establishing a precedent that was often followed subsequently. Many if not most of the PTAs notified to GATT embodied many holes and loopholes (Hoekman and Leidy, 1993).

A contributing factor impeding the ability of working parties to agree was that the wording in Article XXIV is not precise. Legitimate differences of opinion could exist on how to define 'substantially all trade', how to determine whether the external trade policy of a customs union has become more restrictive on average, and what is a reasonable length of time for the transition towards full implementation of a PTA. The 'substantially all trade' test was particularly important. Does this permit the wholesale exclusion of one or more major sectors—as was the case in the EEC, where agriculture was subject to a managed trade regime (the CAP)? In the GATT period, the EC argued in favour of at least 80 per cent coverage of pre-PTA trade flows; others argued in favour of a more comprehensive coverage test. Currently, the de facto focal point appears to be somewhere around 90 per cent. This can be inferred, for example, from the view taken by the EU that in its EPA negotiations with ACP countries Article XXIV would be satisfied if the EU included all its imports from ACP partners, while the latter could limit their commitment to liberalization of 80 per cent of their imports from the EU (Stevens and Kennan, 2005).

Another ambiguous dimension of the language in Article XXIV:8 is that it is not clear what 'other regulations of commerce' covers, and in particular whether it includes the preferential rules of origin used by PTA members and instruments of contingent protection. Article XXIV is entirely silent on rules of origin, which is

rather surprising given that they have an important bearing on the effects of a PTA. The WTO membership have not been able to agree on the rules of origin question although debates in GATT working parties during the 1970s illustrate that officials were well aware that such rules can have important implications for the trade effects of PTAs (Box 10.2).

Can members of FTAs still use antidumping or safeguard measures against each other, or conversely, may they exempt each other if they apply such instruments to their external trade? Case law has concluded that contingent protection may be used by PTA members as long as the rules of the relevant WTO agreements are satisfied and as long as they do not use total imports—including from PTA partners—as the basis on which to determine if (serious) injury has occurred. This so-called parallelism requirement was confirmed by two dispute settlement cases, one involving Argentina, a member of Mercosur and one against the US (a NAFTA member): if PTA partner imports are considered in determining injury, they also need to subject a safeguard action if one is imposed.[4]

An effort was made in the Uruguay Round to make some of the criteria more precise, thereby removing one excuse for a working party not to come to a firm conclusion on a PTA. The GATT 1994 Understanding on the Interpretation of Article XXIV recognizes that the effectiveness of the role of the Council for Trade in Goods in reviewing agreements notified under Article XXIV needed to be enhanced. This was to be pursued in part by clarifying the criteria and procedures for the assessment of new or enlarged agreements, and by improving the transparency of all agreements notified to GATT under Article XXIV. The Understanding reaffirms that PTAs should facilitate trade between members—but did not define what is meant with 'substantially all trade'. Parties to a PTA are called upon 'to the greatest possible extent avoid creating adverse effects on the trade of other members' (GATT, 1994*a*: 31).

Greater specificity proved possible on the evaluation of the general incidence of the duties and other regulations of commerce applicable before and after the formation of a customs union. This must be based upon 'an overall assessment of weighted average tariff rates and of customs duties collected' by the WTO Secretariat, based on import statistics for a previous representative period (to be supplied by the customs union) on a tariff line basis broken down by WTO member country of origin.

Article XXIV:6 requires members seeking to increase bound tariff rates upon joining a customs union to enter into negotiations—under Article XXVIII (Modification of Schedules, see Chapter 9)—on compensatory adjustment. In doing this, reductions in duties on the same tariff line made by other members of the customs union must be taken into account. If such reductions are insufficient compensation, the Understanding requires the customs union to offer to reduce

[4] *Argentina—Footwear* (WT/DS/121) and *US—Wheat Gluten* (WT/DS/166).

Box 10.2. Rules of origin in free trade agreements

The extent of intra-regional trade liberalization under a FTA depends on its rules of origin. Upon the formation of a FTA, nonmember countries may not only be confronted with trade diversion due to the preferential nature of the abolition of barriers to trade, but also because of an effective increase in protection due to the choice of rules of origin. Assume an intermediate product enters a country free of duty and that this country accedes to a FTA. Industries using this input that export to FTA members may then have an incentive to shift to higher cost PTA-based producers of intermediates in order to satisfy the rules of origin for their product. In effect, the rule of origin may become the equivalent to a prohibitive tariff for the original third country suppliers of components, thus generating trade diversion (Krueger, 1997*a*). Research suggests that the tariff equivalent of NAFTA rules of origin is 3–4 per cent (Anson et al., 2005).

An important factor is whether the rule allows for cumulation. Suppose a product is imported that has been processed in at least two countries, both of which have preferential status. An origin system is cumulative if the importing country only requires that sufficient processing of the product has occurred in any of the countries to which the PTA applies. That is, it allows the exporting country of the final product to add (cumulate) the value added in other member countries to that added by itself. If the valued added criterion is 40 per cent, and 30 per cent was added in Country 1 and 20 per cent in Country 2, the product would meet the criterion under a cumulative origin system. Under a noncumulative system of origin 40 per cent would have to be added in each country. Noncumulative rules of origin are much more restrictive than rules that allow cumulation.

The more restrictive the rules of origin, the more they will reduce the extent of liberalization implied by the FTA. In an empirical analysis of trade between the EU and individual EFTA countries—each of which in principle had duty-free access to the EU—Herin (1986) found that the costs associated with satisfying the rules of origin imposed by the EU were high enough to induce 25 per cent of EFTA exports to enter the EU by paying the applicable MFN tariff.

GATT-1947 working parties were not able agree whether rules of origin are covered by Article XXIV:8. For example, in the context of the 1972 FTA between the EEC and EFTA States, the US argued that the rules of origin would generate

> ...trade diversion by raising barriers to third countries' exports of intermediate manufactured products and raw materials. This resulted from unnecessarily high requirements for value originating within the area. In certain cases... the rules disqualify goods with value originating within the area as high as 96 per cent. The rules of origin limited non-originating components to just five per cent of the value of a finished product of the same tariff heading [for] nearly one-fifth of all industrial tariff headings. In many other cases a 20 per cent rule applied. (GATT, 1974: 152–3)

Although it would appear that rules of origin are unambiguously detrimental to the welfare of participating countries, this is not necessarily the case. Duttagupta and Panagariya (2007) demonstrate that restrictive rules of origin can raise welfare by reducing the magnitude of trade diversion in trade in final goods. However, intuition suggests that in most cases, restrictive rules such as the triple transformation or yarn-forward rules used for textile products in NAFTA (products must be made from cloth embodying yarn

(*cont.*)

Box 10.2. (*Continued*)

originating within the region) will be costly to consumers (Krueger, 1999). The obvious solution is to pursue harmonization—e.g. for WTO members to agree to use the (non-preferential) WCO rules (see Chapter 5). Alternatively, they could lower external MFN tariffs to zero as this would remove the need for rules of origin! As discussed below, in practice the EU has pursued a regional harmonization strategy and has put in place a system under which there is cumulation to reduce the costs of rules of origin.

duties on other tariff lines, or to otherwise provide compensation. Where agreement on compensatory adjustment cannot be reached within a reasonable period from the initiation of negotiations, the customs union is free to modify or withdraw the concessions and affected members are free to withdraw substantially equivalent concessions (to retaliate).

The 1994 Understanding established a ten-year maximum for the transition period for implementation of an agreement, although allowance is made for exceptional circumstances (to be defended in the Council for Trade in Goods). Working parties are called upon to make appropriate recommendations concerning interim agreements—PTAs with a transitional implementation period—as regards the proposed time period and the measures required to complete the formation of the PTA. If an interim agreement does not include a plan and schedule, the working party must recommend one. Parties to a PTA may not implement it if they are not prepared to modify it in accordance with the recommendations. Implementation of the recommendations is subject to subsequent review.

Developing countries are not bound by Article XXIV as a result of the 1979 Decision on Differential and More Favorable Treatment of Developing Countries (the so-called Enabling Clause—see Chapter 12). This essentially removes the 'substantially all trade' test and allows for preferences between developing country PTA members (that is, the full removal of internal barriers—free trade—is not required). For example, Mercosur (a customs union) was notified to GATT under the Enabling Clause, not under Article XXIV.

GATS Article V: Economic Integration

The GATS is similar to the GATT in allowing for PTAs that liberalize trade in services on a discriminatory basis, subject to conditions and surveillance. The relevant provision, Article V GATS, is entitled Economic Integration, not Free Trade Areas and Customs Unions (as is Article XXIV GATT), reflecting the fact that the GATS covers not only cross-border trade in services but also the three other modes of supply. Article V GATS imposes three conditions on economic integration agreements. First, such agreements must have substantial sectoral coverage, in

terms of the number of sectors, volume of trade affected and modes of supply. Preferential trade agreements may not provide for the a priori exclusion of any mode of supply. Second, PTAs must provide for the absence or elimination of substantially all measures violating national treatment in sectors where specific commitments were made in the GATS. This must be achieved at the entry into force of the agreement or within a reasonable time frame. Third, PTAs may not result in higher trade barriers against third countries.

The substantial sectoral coverage requirement is weaker than the 'substantially all trade' criterion of Article XXIV. The same conclusion applies regarding the criteria on the magnitude of liberalization required and the external policy stance of the PTA, as the benchmark is not free trade among PTA members, but the specific commitments made under the GATS by the PTA members. As discussed in Chapter 7, most of these are far from implying free trade. Those members of the WTO engaged in economic integration efforts intending to withdraw or modify specific market access or national treatment commitments (raise external barriers) must follow the re-negotiation procedures set out in Article XXI GATS (Modification of Schedules—see Chapter 9).

There are a number of loopholes allowing for the formation of agreements that do not fully comply with multilateral disciplines. For example, Article V:2 of the GATS allows for consideration to be given to the relationship between a particular PTA and the wider process of economic integration among member countries. Article V:3 gives developing countries involved in a PTA flexibility regarding the realization of the internal liberalization requirements and allows them to give more favourable treatment to firms that originate in PTA members. That is, it allows for discrimination against firms originating in nonmembers, even if the latter are established within the area. These special and differential treatment type of provisions are unlikely to be very effective in achieving their presumed objective: attracting FDI. More importantly, they weaken the scope of multilateral disciplines, giving governments (interest groups) an opportunity to pursue agreements that are more detrimental to nonmembers.

Effectiveness of WTO disciplines

As mentioned, GATT experience in enforcing Article XXIV was disappointing, to put it mildly. Although Article XXIV is far from perfect from an economic perspective—more on this in the next section—in principle it imposes serious discipline, especially after the Uruguay Round Understanding. Despite the replacement of working parties by a single Committee on Regional Trade Agreements to review the compliance of PTAs, as under GATT 1947, agreement on whether a PTA satisfied Article XXIV and Article V proved impossible during the 1995–2007 period. The only exception was the customs union between the Czech and Slovak

Republics—not surprising given that the two countries were a federation prior to their 'velvet divorce' in January 1993. (As both countries are now EU members, this customs union no longer applies.) As under GATT, the reason for this impotence is the consensus rule.

In December 2006 the WTO General Council established a new transparency mechanism for PTAs. The mechanism—a product of the Doha Round Negotiating Group on Rules—imposes the obligation on members to inform the WTO Secretariat on newly launched negotiations as well as newly signed PTAs. Notified PTAs will be considered on the basis of a factual presentation by the WTO Secretariat, with the process to be concluded within one year of notification. Any member of the WTO may ask questions or make comments concerning factual presentations of PTAs and the implementation of the liberalization commitments relating to PTAs should be notified to the WTO Secretariat. Agreements falling under Article XXIV of GATT and GATS Article V will be considered by the Committee on Regional Trade Agreements (CRTA). Trade arrangements between developing countries falling under the Enabling Clause will be reviewed by the Committee on Trade and Development. The transparency mechanism was implemented on a provisional basis in 2007 and was expected to be reviewed and adopted on a permanent basis as part of the overall results of the Doha Round (WT/L/671, 18 December 2006).

The transparency mechanism for PTAs may help move the balance of assessments of PTAs back towards what was intended by the drafters of the GATT—*ex ante* review and engagement by the collective membership on the design of a PTA, as opposed to what gradually emerged over time: ineffectual *ex post* assessments (Mavroidis, 2007). However, the track record to date suggests that multilateral scrutiny is unlikely to be an effective source of discipline on PTAs. The Doha Round transparency mechanism does not have any teeth, and it was clear from the deliberations that preceded the creation of the mechanism that many WTO members do not intend to use it as a means of exerting greater pressure on countries to abide by the rules. The fact that they call the process 'consideration' of a PTA as opposed to 'examination' of a PTA is quite revealing in this regard.

That said, greater transparency may have an indirect effect by supporting greater scrutiny of PTAs by citizens of the countries concerned. It may also facilitate greater use of the WTO dispute settlement mechanism to contest specific aspects of PTAs. It is rather puzzling that relatively little use has been made of dispute settlement procedures to contest the operation or design of PTAs after 1995. In principle nothing constrains a WTO member from invoking Article XXIV or Article V—it is not necessary that the CRTA has come to a conclusion on whether a PTA complies with the WTO rules to contest the operation of a PTA. During the GATT period the incentive to bring a case was greatly reduced by the consensus rule that applied to both the establishment of a panel and the adoption of panel reports. But this constraint disappeared with the establishment of the WTO.

To date, there have been two disputes where conformity with Article XXIV was a factor. The first was a 1996 case brought by India against Turkey, regarding Turkey's imposition of quantitative restrictions on imports of textile and clothing products. These were imposed by Turkey as a result of a customs union agreement with the EU—the restrictions were applied by the EU so that Turkey had no choice but to apply them too. The DSB established a panel, which found that Turkey's measures were inconsistent with Articles XI and XIII GATT—the ban on quantitative restrictions. Of interest to this chapter, the panel rejected Turkey's assertion that its measures were justified by Article XXIV GATT. On appeal, the Appellate Body upheld the panel's conclusion that Article XXIV of GATT 1994 does not allow Turkey to adopt, upon the formation of a customs union with other WTO members, measures that are inconsistent with other WTO disciplines. However, the Appellate Body also concluded that the legal interpretation of Article XXIV by the panel was erroneous, and determined that a panel should *first* ascertain whether a PTA complies with Article XXIV before considering other GATT provisions. This appears to opens a rather large door to disputes.

A second case was brought in 2001 by Korea against the US, which had imposed safeguard measures on imports of circular welded carbon quality line pipe (steel products) but exempted its NAFTA partners. As in the India–Turkey dispute, the case was decided on the basis of other WTO provisions—in this case the Agreement on Safeguards—but of interest is the discussion of the Article XXIV defence invoked by the US. In *US—Line Pipes* Korea maintained that a PTA must be presumed inconsistent with Article XXIV until the CRTA makes a determination to the contrary (which of course it never has). However, the Panel held that because Korea did not refute the evidence provided by the US to the CRTA when it notified NAFTA, there was a prima facie case of consistency of NAFTA with Article XXIV. In the report on NAFTA submitted by the US to the CRTA, the US stated that duties on 97 per cent of the NAFTA-parties' tariff lines would be eliminated within 10 years. With respect to other regulations of commerce—which include safeguards—the report made a reference to the principle of national treatment, transparency and a variety of other market access rules (WT/DS34/AB/R, October 2001).

Mavroidis (2006) has argued that there is great scope to use dispute settlement to directly address instances where PTAs do not comply with Article XXIV GATT and/or Article V GATS. The additional transparency that will be generated as a result of the 2006 transparency mechanism may facilitate moving down this track. But Mavroidis also notes that the spread of PTAs is now such that virtually all WTO members are implicated and are thus likely to worry about the potential adverse consequences of using the DSU to attack a specific PTA. Fundamentally, the problem of using the DSU to deal with perceived nonconforming PTAs is the fact that disputes must be brought by a specific WTO member, i.e. the multilateral enforcement that is in principle needed given the resulting collective action problem simply does not exist.

Experience to date suggests that multilateral surveillance will only have a limited impact on the design and content of PTAs. An implication is that the payoff to efforts to strengthen specific WTO disciplines on PTAs is likely to be low, even if agreement could be attained. More fundamentally, it has been argued that efforts to devise a realistic rule that will ensure the trade policy stance of a PTA will be welfare-improving for members *and* the rest of the world are doomed to failure (Winters, 1999). There is simply no way to square the circle. As has been noted by many observers, the proliferation of PTAs extant clearly illustrates that WTO members regard PTAs as being in their interest. Thus, what matters is the economics—the incentives that are created by the proliferation of PTAs as regards the average level of MFN protection. That said, the rules are not irrelevant. In particular, they have had a major impact on the trade relations of a significant number of developing countries with the EU. As discussed in Chapter 12, an important reason why the EU concluded EPAs with ACP countries was to replace a system of unilateral preferences that violated GATT rules (and thus required a waiver) with a set of reciprocal trade agreements that satisfied Article XXIV. Interestingly, given the long history of nonenforcement of WTO rules in this area, it was the EU itself that was the 'enforcer' of Article XXIV, in that the desired trade coverage ratio of the EPAs was determined by the EU's view of what is the minimum required by Article XXIV.

10.3. Trading Blocs and the Trading System

A key factor determining the importance of the effective absence of multilateral disciplines is the extent to which PTAs have detrimental effects on nonmembers and how they react. For the trading system what matters are the dynamic forces that are created by PTAs—do they create incentives that lead to a reduction in the external barriers of PTAs and nonmembers? As is the case for trade policy more generally, the most powerful pressures for reform are almost invariably domestic, not external, although external forces can help support domestic constituencies that favour a more liberal trade regime. That suggests a focus on the economic impacts of PTAs on interest groups.

What follows briefly considers three relevant questions in this regard. First, what is the impact of PTAs on trade and welfare of members? Second, what are the consequences for nonmembers? Third, what are the incentives created by PTAs once they have been formed for both members and nonmembers in terms of their trade policy strategies?

Impacts of PTAs on members

The effects of a PTA are determined by their coverage and design and on whether and how they are implemented. If a PTA is not implemented it cannot have an effect. Many of the PTAs negotiated since the 1960s were only partially implemented, if at all, or else excluded so many industries and tariff lines that their trade effects could only be minimal. The more recent vintage PTAs are generally implemented, and as mentioned tend to have more substantial coverage of merchandise trade flows.

Empirical research assessing the magnitude of the trade impacts shows, not surprisingly, that PTAs increase trade between members. That is, after the PTA is implemented, one observes greater trade flows between members. The difficulty for researchers, however, is to establish whether there is a causal effect. In a world where countries and thus trade is growing one would expect more intra-PTA trade without a PTA. Matters are compounded by countries also undertaking unilateral liberalization at the same time or before they engage in PTAs—what then is driving increased intra-PTA trade?

This suggests that empirical evaluations of PTA impacts must compare outcomes to what would have happened absent the PTA (the counterfactual). This is very difficult if not impossible as the PTA exists after all. What can be done, however, is to control for other factors and variables that affect trade flows. The basic workhorse tool that tends to be used to assess the effects of PTAs is the gravity model. This is a model that has been shown to be very effective at explaining trade volumes between country pairs, and that is consistent with what economic theory predicts are the determinants of trade. In a nutshell it postulates that trade between two countries is a function of their size, their wealth, their distance from each other, whether they are contiguous and speak the same language, and policy variables. The latter include the existence of a PTA.

Much of the literature on this subject is summarized by Schiff and Winters (2003). Surprisingly, there is no agreement on whether PTAs lead to more intra-PTA trade—indeed some studies find a negative effect. More recent research using the gravity model by Baier and Bergstrand (2007) argues that the findings of much of this literature greatly understate the trade effects of PTAs because they ignore the political economy of trade policy, i.e. why the PTA was negotiated in the first place. Technically, what researchers often assume is that the formation of a PTA is exogenous. In practice it is not likely to be—instead the level of trade can be expected to determine whether or not to join a PTA. If account is taken of this endogeneity, the impact of PTAs on trade volumes with partner countries rises significantly. On average PTAs do have a significant effect on intra-PTA trade: according to Baier and Bergstrand (2007) on average a FTA doubles trade between two members after ten years.

Of course, more trade is not necessarily good from a welfare perspective. What matters is how much of what is observed is trade creation and how much is diversion. Not surprisingly, empirical assessments of the impact of PTAs do not come to uniform conclusions. Much depends on the structure of trade before and after the formation of the PTA, on the pattern of comparative advantage, the size and composition of the PTA, etc. That said, many PTAs have been found to generate trade diversion. For example, in a study of eight major PTAs over the 1970–92 period, Frankel, Wei and Stein (1997) found that increases in intra-bloc trade were accompanied by reductions in trade with nonmembers, i.e. generated trade diversion. Similarly, Soloaga and Winters (2001), focusing on nine major PTAs over the period 1980–96, found trade diversion in European PTAs (explained by the fact that little external liberalization occurred during this period relative to intra-PTA liberalization), whereas PTAs among developing countries saw trade increase with both members and nonmembers. The explanation for this is that trade policies were reformed by the developing countries more generally, i.e. reforms were not limited to preferential liberalization. Controlling for developments in general (MFN) trade policies, the non-European PTAs had no independent effect on intra-PTA trade flows (Schiff and Winters, 2003: 42–3).

Very detailed analysis at the HS six-digit level of disaggregation (some 5,000 products) of the impact of the FTA between Canada and the US and the subsequent NAFTA by Romalis (2005) provides clear evidence of trade diversion. He shows that the greatest increases in US imports from Mexico occurred in items on which the US imposes the highest MFN tariffs, i.e. those goods where NAFTA provides Mexico with the highest preferential tariff margins. A similar result obtains for Canada. Although overall welfare effects of NAFTA for the US are small, one reason for this is the trade diversion, which results in higher prices of protected goods. Romalis (2005) also finds that volume effects are significant: NAFTA increases trade between Mexico and both Canada and the US by almost 25 per cent. Thus, studies suggest that there may well be significant market access and terms of trade benefits for countries joining a PTA, as well as distributional effects—with consumers paying the costs of any trade diversion.

Clearly, a narrow focus on merchandise trade is inadequate to assess the effects of PTAs. As, if not more, important, are the impacts on investment and FDI, and the associated potential for the acquisition and diffusion of technology, and the extent and implications of the 'deeper integration' dimensions of PTAs. Many studies have found that 'serious' PTAs may encourage FDI inflows and that these in turn can generate positive productivity spillovers (Schiff and Winters, 2003). There is nothing automatic about such investment and spillover effects, however. The experience of some 20 developing countries between 1980 and 2000 illustrates that many PTAs have not led to significant new FDI inflows (World Bank, 2005).

Impacts of PTAs on nonmembers

From the perspective of the trading system, the impacts of PTAs on nonmembers is the relevant question. Economists sometimes argue that a necessary condition for PTAs not to be detrimental to nonmembers is that the volume of imports by member countries from the rest of the world not decline on a product-by-product basis after the implementation of the agreement (McMillan, 1993). The empirical literature suggests that the trade volume test has been met in the past. Although the intensity of intra-regional trade increased in the second half of the twentieth century, the propensity of regions to trade with the rest of the world, expressed as a percentage of their GDP, has also expanded (Anderson and Norheim, 1993). Global integration—as measured by trade flows and capital flows—does not appear to have been affected negatively by PTAs.

As pointed out by Winters (1997), the 'trade volume test' is a flawed one in that it does not guarantee that nonmembers are not hurt by a PTA. For the welfare of nonmembers what matters is the impact of a PTA on trade flows and the associated change in prices. Even if the Article XXIV conditions are met, and even if the net aggregate imports of PTA members do not contract, imports of particular products by the region may decline *ex post*, or prices received by exporters may fall, harming producers in the rest of the world.

The converse of the trade diversion discussed above is that it implies a decline in exports for nonmember countries to a PTA, and perhaps an overall decline in aggregate exports if the diverted trade cannot be redirected to other markets and sold at the same price. Schiff and Winters (2003) discuss much of the literature, which finds that nonmembers have at times experienced significant reductions in exports. In the case of NAFTA, Romalis (2005) concludes that every 1 per cent reduction in intra-NAFTA tariffs causes a decline in exports to NAFTA from the rest of the world ranging from 1.3 to 3.9 per cent. Although such findings are suggestive, a more appropriate measure of the welfare impact of a PTA on nonmembers is to focus on what happens to their export prices in PTA markets after the agreement is formed. Chang and Winters (2002) show that Brazil's membership of Mercosur was accompanied by a improvement in Brazil's external terms of trade. Exporters based in the US, EU, Japan and Korea all saw the relative prices of many of their goods on the Brazilian market fall. There is also some evidence of negative investment effects (Baldwin, Forslid and Haaland, 1996).

Limão (2006) and Karacaovali and Limão (2008) have shown that in the case of both the EU and the US, PTAs may be a force working against nonmembers: they find that both the EU and the US made fewer (shallower) multilateral (MFN) liberalization commitments in the Uruguay Round on tariff lines where there were significant preference margins for imports from their preferential trading partners. Limão (2007) hypothesizes that this may reflect the use of market access as

'payment' for concessions by PTA partners in nontrade areas. Whatever the rationale, this is evidence that PTAs can have stumbling block effects.

Thus, notwithstanding the fact that there is little evidence of large-scale negative effects of the spread of PTAs—as reflected in the steady increase in world trade and openness discussed in Chapter 1—the economic literature suggests that there is no justification for complacency regarding the effects of PTAs. Preferential trade agreements impose costs on nonmembers even if they do not raise external levels of protection. Nonmember suppliers become less competitive because they continue to pay tariffs, whereas competing producers from member countries do not. Where there are economies of scale, PTAs may help lower member country firms' costs by expanding their home market. Conversely, they may restrain the ability of firms in nonmembers from realizing economies of scale.

There are various ways through which PTAs may constrain national interest groups and thus foster a more liberal external trade policy (De Melo, Panagariya and Rodrik, 1993). A first can be called the preference-dilution effect: because the region implies a larger political community, each of the politically important interest groups in member countries will have less influence on the design of common policies. The second is a preference-asymmetry effect: because preferences on specific issues are likely to differ across member countries, the resulting need for compromises may increase the probability of more efficient outcomes. The creation of PTAs may also disrupt the formation of rent-seeking interest groups, as these have to reorganize at the regional level, establishing an institutional structure that allows them to agree on a common position. But, PTAs may also facilitate the adoption of less liberal policies. Consumer interests may be harder to defend in a PTA than at the national level, whereas producer interests are more likely to be strengthened than weakened (Tumlir, 1983). Each national producer group may face less opposition when seeking price-increasing policies, and may indeed find support from other producer groups in other countries that pursue their own interests. The need for striking compromises may then result in a less liberal regulatory regime. Moreover, it may be in the interest of national politicians to let a regional organization satisfy national pressure groups as this is less transparent for domestic voters and can be justified as being necessary to maintain the agreement.

Much will generally depend on the type of PTA that is involved, FTA, customs union or hub-and-spoke system. The first two types differ from the last in that they imply nondiscrimination between the members of the agreement: any benefit granted to member country B by member country A is also available to member country C. Under a hub-and-spoke system this is not necessarily the case: each country negotiates a separate agreement with the hub country, and perhaps with other spoke partner countries as well. A major difference between a FTA and a customs union or common market is that the latter implies a common external trade policy. Whatever the extent of internal liberalization of trade and competition, implementation of a common external trade policy can give rise to an upward bias in the level of external protection over time if import-competing industries pursue instruments of contingent protection

such as antidumping actions. Thus, there may be no net increase in external trade barriers at the formation of a customs union, but there can easily be an upward trend if contingent protection is maintained. In contrast, FTAs have a different dynamic, as members in some sense compete in their external trade policies. Although the political economy of FTAs versus customs unions is complex, on balance, FTAs are likely to be more liberal than customs unions (Box 10.3).

Box 10.3. Pressures for protection: FTAs and customs unions

Under a customs union or common market the potential returns to protection-seeking will be higher than under a FTA: the expected payoff for a unit of lobbying effort increases because the size of the protected market is bigger. Moreover, liberal-minded governments that join a customs union may find it impossible to prevent domestic industries from seeking protection or to block the imposition of protection. For example, it may be the case that certain countries did not use (or make available) contingent protection before joining a customs union. However, once a member country, any domestic firm has access to the central trade policy authority and will be able to petition for AD. Indeed, the welfare gains to liberal countries from joining a customs union that employs contingent protection are reduced, as consumers are faced with higher expected levels of protection without knowing which industries will be affected (Hoekman and Leidy, 1993).

More generally, once a common external trade policy applies, decision-making structures may be biased toward more rather than less protection because of the so-called restaurant bill problem. If a group goes to a restaurant and shares the cost of the bill, each has an incentive to order more expensive dishes than they would if they ate on their own, as to some extent the others are expected to pick up part of the cost. The same is true in the EU (Winters, 1994*b*). The costs of protection are borne by all EU consumers, and are roughly proportional to each country's GDP. Benefits accruing to producers are proportional to the share of each country in total EU production of the good concerned. This establishes an incentive for each government to pursue protection for those products where their share of total EU production exceeds their country's share of EU GDP. Thus, the Netherlands may not like the EU-wide protection for cars sought by France and Italy, but may accept it if other policies are adopted for products in which it is relatively specialized (such as agriculture). Indeed, if larger countries are able to get the Commission to propose protectionist policies in specific areas, all EU member states have an incentive to ensure that some of their producers also obtain protection.

The external trade policy bias towards protection that may arise under a customs union will be weaker in a FTA. Because there is no common external trade policy, member countries compete in their external trade policies. Industries cannot lobby for area-wide protection. Although import-competing firms in member countries may have an incentive to obtain such protection, each industry will have to approach its own government. The required coordination and cooperation may be more difficult to sustain than in a customs union where the centralization of trade policy requires firms to present a common front. In any particular instance, some member country governments will award protection, whereas others will not. If industries in member countries are all

(*cont.*)

Box 10.3. (*Continued*)

competing against third suppliers, protection by one member may benefit industries in other member states. Such free riding can result in less protection than in the absence of the FTA (Deardorff, 1994). This benefit may be offset by other aspects of FTAs, in particular the need for rules of origin, which may allow industries to limit the extent of intra-area liberalization and can be detrimental to nonmembers (see Box 10.2 above).

Some evidence is beginning to emerge that supports these theoretical considerations on the likely dynamics of FTAs versus customs unions. Rigorous empirical research on the relationship between preferential and MFN tariffs over time is sparse as a result of data constraints—information on the implementation of PTAs and the sequencing over time between unilateral and preferential tariff reduction is not available for many PTAs. In the case of Latin America, however, a study by Estevadeordal, Freund and Ornelas (2008) concludes that the preferential tariff reduction following PTA formation in Latin America promotes subsequent external tariff reduction for those PTAs that are not customs unions. Bohara, Gawande and Sanguinetti (2004), focusing on the impact of preferential trade flows from Brazil to Argentina, find that greater imports from Brazil led to lower MFN tariffs in Argentina, especially in sectors where trade diversion occurred as a result of Mercosur. As the potential for trade diversion is especially great for South-South PTAs—because developing countries tend to have relatively high external trade barriers—the associated costs provide a powerful force for multilateralization: lowering external barriers to trade will reduce such costs.

Responses by nonmembers to PTA proliferation

As stressed by Bhagwati (1991), from the perspective of the WTO a key question is whether PTAs are a stepping stone or a stumbling block for multilateral liberalization. There is no consensus on the answer. Indeed, given that PTAs differ so much, there is no reason to expect a single, simple answer, especially as this is inherently a dynamic question—the answer depends critically on how PTAs affect the incentives of pro- and antitrade forces in both PTAs and excluded countries.

The most obvious reaction of third countries to the formation of a PTA is to seek a reduction in the bloc's external trade barriers through a MTN. As noted, this arguably has been a key role of the WTO in practice, with regional integration in Europe becoming a recurrent reason for MTNs under GATT auspices. Much of the Dillon Round (1960–1; see Chapter 4), was devoted to renegotiating a balance of concessions subsequent to the implementation of the EEC's common external tariff.

The same type of objectives played a role in the Kennedy and Tokyo Rounds. At the time of the Kennedy Round, the margins of preference for EEC members had

increased substantially, as most of the internal elimination of tariffs had been achieved. 'The record leaves no doubt that a compelling factor in the decision of Congress to pass legislation authorizing a 50 per cent linear cut in tariffs [in the Kennedy Round, see Chapter 4] ... was the belief that the Common Market posed a potentially serious threat to the growth, and perhaps even maintenance of American exports' (Patterson, 1966: 176). Thus, 'the task of the Kennedy Round ... was to attempt to mitigate [the] disruptive trade effects of European economic integration' (Preeg, 1970: 29). Some success was achieved, as the Kennedy Round reportedly prevented one-third to one-half of the trade diversion that might have occurred from European integration (ibid.: 220).

The first enlargement of the EEC in 1973—to include Denmark, Ireland and the UK—was a factor behind the launching of the Tokyo Round. The CAP also played a role. A major objective of the US was to improve its market access for agricultural products and to curb the EU's use of export subsidies. Links between regional integration and the Uruguay Round included the adoption of the Single European Act (the 1992 programme), the implementation of the Canada–US FTA, the negotiations on the NAFTA, and the continuing distortions of world agricultural trade induced by the CAP. The foregoing is not to say that PTAs are good because they give countries an incentive to pursue concurrent MTN-based liberalization. Without the EEC, much more progress might have been made towards multilateral liberalization (Winters, 1994*b*).

Another policy option is to seek to join existing PTAs. The primary example here is again the EU, which expanded from six to 27 member states between 1957 and 2007. In North America, Mexico was induced to seek accession to a Canada-US FTA, with the result being a re-negotiated trilateral FTA, the NAFTA. Other nations have negotiated FTAs with each of the NAFTA members in turn. One motivation for this is market access 'insurance'. The goal is not necessarily so much to obtain duty-free access to the regional market, as average MFN tariffs are relatively low for most products, and many potential members tend to be treated preferentially in any event. More important is the elimination of uncertainty, including the threat of contingent protection. This may be complemented by a desire to enhance the credibility of recently undertaken unilateral liberalization and structural reform efforts. However, particularly important are likely to be the firms in nonmember states that see their competitors get access to an ever larger internal market, allowing them to realize economies of scale and benefit from a reduction in real trade costs. This may well give rise to the 'domino effects' that have been observed in the case of the EU and NAFTA (Baldwin, 1995). Examples of such domino effects abound, especially in the European context. As mentioned at the beginning of this chapter, the EU has numerous PTAs with third countries.

The creation of a PTA may also generate incentives for third countries to pursue PTAs in turn. This may be a defensive rationale, reflecting a desire to strengthen their bargaining position vis-à-vis major trading partners and allow them to 'better

defend themselves against discriminatory effects of other regional groups' (Patterson 1966: 147). The formation of EFTA is an example. It was established in 1960 in reaction to the formation of the EEC. Its membership consisted of European countries that did not want to join the EEC because of concerns relating to the supranational aspects of the EEC and the likely level of the common external tariff (most EFTA countries tended to be relatively liberal). The EFTA reaction to the formation of the EEC was not unique. Japan informally proposed a Pacific Free Trade Area with the US, Canada, Australia and New Zealand in the mid-1960s for the same reason (De Melo and Panagariya, 1993). More recently, Pacific nations agreed to pursue regional free trade under auspices of the Asian-Pacific Economic Cooperation (APEC) framework. This is an example of so-called open regionalism, where PTAs are used as a focal point for concerted liberalization. Essentially this involves the formation of a privileged group (see Chapter 4). Free riding problems can be expected to be important in such efforts—it is unlikely that APEC will realize the stated goal of free trade by 2020 given the shift by many of the East Asian economies towards the negotiation of formal PTAs. However, as discussed by Baldwin (2006*b*), the resulting 'noodle bowl' may generate incentives for firms in the region to push more actively for regional, if not global, free trade.

Arguments suggesting PTAs may be detrimental to the trading system often revolve around some variant of the optimal tariff argument. As trade blocs expand, so does their market power and, at least in principle, their ability to influence the terms of trade in their favour. If successful, this is detrimental to the rest of the world. Although possibly true in some cases, it is not a well-founded generalization that PTAs will have an incentive to increase their tariffs against the rest of the world. For one thing, there are big differences between FTAs and customs unions. As already discussed, members of FTAs may have good reasons for lowering tariffs on nonmembers, as this reduces trade diversion.

Baldwin (2006*a*, *b*) argues that there may well be positive incentive dynamics resulting from hub-and-spoke PTAs. The domino effects noted previously may move more countries to lower trade barriers, as over time the 'balance of power' between export- and import-competing interests shifts in favour of those benefitting from a more open trade regime. Baldwin's theory of how this may play out starts from the premise that at a given point in time export interests see benefit in expanding access to locations where they can undertake parts of their product process and can get their government to negotiate a PTA. Another important part of the story is that the major players are the big markets—such as the EU—so that one result of the process is a hub-and-spoke system of PTAs. This essentially consists of a set of bilateral trade agreements. Because a hub-and-spoke system involves separate agreements between the hub (e.g. the EU) and the spoke countries, there is much scope to exclude 'sensitive sectors' from the coverage of each bilateral agreement (Snape, Adams and Morgan, 1993). Each spoke is likely to have comparative advantage in a somewhat different set of such sectors. If each country

maintains contingent protection options (AD, safeguards) against member countries, powerful import-competing industries in the hub country will have an interest in including wide-ranging safeguard clauses and relatively stringent rules of origin.

This was the case in the Association Agreements negotiated between the EU and various Eastern neighbours in 1992 (Winters, 1995). By allowing bilateral deals regarding sectoral coverage and the depth of the agreement, vested interests could be assuaged through specific rules of origin as well as safeguard provisions. The nature of these types of PTAs were such that they allowed for significant internalization of benefits by producers who wanted to restrict as much as possible the (new) regional market for themselves. Such groups also opposed broader multilateral reform. As summarized by Bhagwati (1993), such groups took the view 'the region is our market', and that 'our markets are large enough'.

In the event, however, the political economy equilibrium that underpinned the hub-and-spoke model of major 'systems' of PTAs began to break down. In part as a result of continuous technological change and in part as a result from increasing competition by China firms in Europe began to see an interest in further reducing the cost of production. One way this could be achieved was through reduction of the administrative costs of the hub-and-spoke system, in particular the associated rules of origin. One result was the adoption of the pan-European Cumulation System in 1997, under which any inputs sourced from *any* of the spokes or the EU member states counts for purposes of determining origin, and thus eligibility for duty-free treatment. This is an example of how regionalism may generate forces through which the objective of a multilateral, nondiscriminatory, trade regime might emerge endogenously.

The Baldwin (2006*a*) story is comforting for those who have been worrying about the systemic implications of PTAs. Saggi and Yildiz (2008) offer another, more theoretical, argument for the positive systemic effects of PTAs. They note that the voluminous literature on PTAs and regional integration ignores a key question: would the WTO serve the cause of global free trade more effectively if it did not include the exception to MFN provided by Article XXIV? Would global free trade be easier to achieve if all WTO members were somehow constrained to pursue trade liberalization on only a multilateral basis? The relationship between preferential and multilateral liberalization has been the subject of much theoretical analysis, but Saggi and Yildiz are among the first to treat both bilateral and multilateral liberalization as endogenous and to allow for the fact that countries are not symmetric in size. A central result is that bilateralism can provide an impetus to multilateral trade liberalization. The insight underlying this result is that a country that is choosing whether or not to participate in global free trade must consider how it would fare under the agreement that would emerge in the absence of its participation. Their model has the reasonable feature that a nonparticipating country is worse off under a bilateral trade agreement than under a

multilateral agreement—this is because a preferential agreement discriminates against the outsider whereas a multilateral agreement does not. As a result, a country's incentive to opt for free trade is stronger when the alternative to free trade is a bilateral agreement between the other two countries as opposed to a multilateral one. It is noteworthy that this result obtains in their model despite the fact that the formation of a PTA induces its members to impose lower tariffs on the nonmember relative to their Nash equilibrium tariffs, a result referred to as the tariff complementarity effect in the literature. In fact, even though a preferential agreement leads to more trade liberalization than a multilateral agreement in which all countries do not participate, it harms the outsider relatively more precisely due to the discrimination that is inherent to it.

The analysis of Saggi and Yildiz makes two additional points. First, they demonstrate that the debate regarding preferential versus multilateral liberalization is moot in the absence of some type of asymmetry across countries, which in their model implies that the gains generated by a shift from the status quo of noncooperative tariffs to free trade are unequally split across countries.[5] Indeed, they show that under sufficient symmetry, both the preferential and the multilateral route lead to global free trade. A second important insight provided by their analysis is that to properly address the issue of how preferential trade liberalization interacts with multilateral trade liberalization, we need to better understand when and why countries choose to pursue the preferential route given that the multilateral route is available. Only a model in which both types of liberalization are fully endogenous can shed real light on this question.

Deeper integration and PTAs

So-called deep integration has become an increasingly important feature of PTAs over the last decade and a half as border barriers decline. This spans many aspects of product and market regulation, including standards, government procurement, services, investment, competition, labour and environmental policies, as well as IPRs and protection of other intangible and tangible assets.

As is illustrated by Figures 10.2 and 10.3, there is enormous variance across recent PTAs in the scope and depth of these policy areas, with PTAs that involve the United States generally having the broadest coverage. There is evidence that PTAs include more service sectors than countries have scheduled at the WTO, but the available research also suggests that their clauses do not move much beyond those in the WTO (Fink and Molinuevo, 2007; Roy, Marchetti and Lim, 2007). This

[5] In a model with repeated interaction, Saggi (2006) had shown that when countries are asymmetric, the exogenous formation of a preferential trade agreement may facilitate multilateral tariff cooperation whereas such a result does not obtain under symmetry.

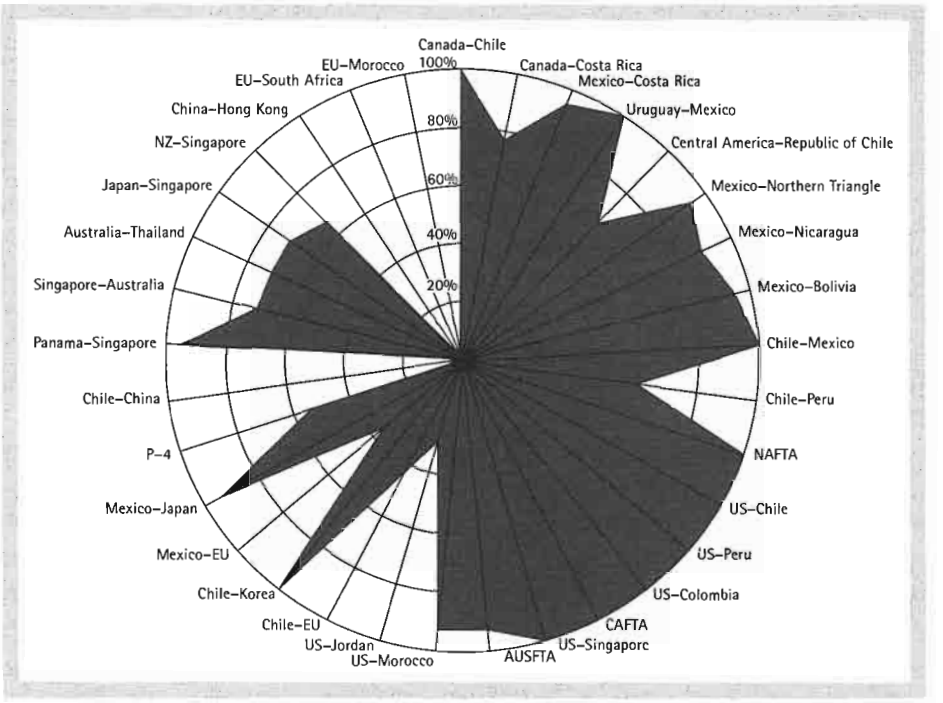

Fig. 10.2. Coverage of 17 investment provisions in selected PTAs

Source: Estavadeordal, Shearer and Suominen 2007.

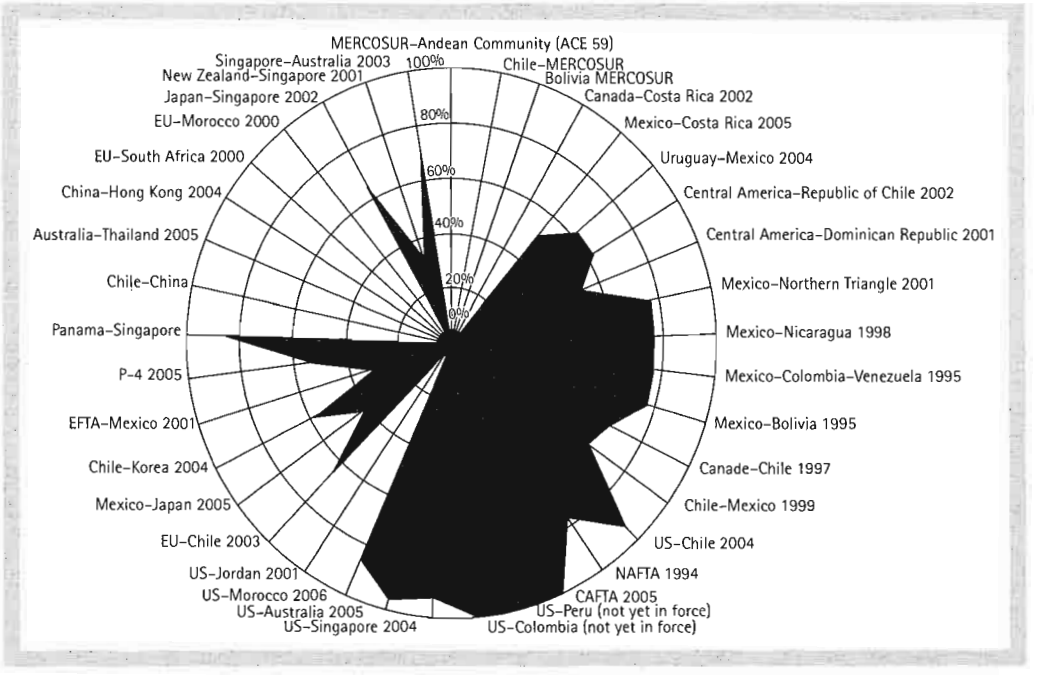

Fig. 10.3. Coverage of 29 services provisions in selected PTAs

Source: Estavadeordal, Shearer and Suominen (2007).

suggests that PTAs may broaden the coverage of commitments to lock-in service liberalization, most of which has been implemented autonomously, but they do not deepen it. It is also the case that many PTAs now cover FDI, whereas the WTO does not. But here again one can question the extent to which disciplines are additional. For example, there are already over 3,000 bilateral investment treaties in place. Moreover, in areas of key importance to the developing countries such as mobility of labour or constraints on the ability of OECD partner countries to offer incentives to investors the bilateral routes have not achieved more than what was possible in the WTO.

How much further than the WTO do PTAs go? Horn, Mavroidis and Sapir (2008) assess the coverage of 14 US and 14 EU PTAs, focusing on the prevalence of WTO+ commitments (provisions that address matters covered by the WTO but that go beyond the WTO in terms of extent of disciplines imposed) and what they call WTO-X commitments: provisions on matters that are not dealt with in the WTO. They conclude that the EU and the US both negotiate WTO+ disciplines in many areas, with the US taking the lead. With the exception of the (early) agreement with Israel and the PTA with Jordan, US PTAs have additional disciplines in almost all the categories distinguished by the authors: industrial market access; agriculture; customs; export taxes; SPS; TBT; STEs: AD; CVDs: subsidies; procurement; TRIMs; services and IPRs. Moreover, these WTO+ disciplines are legally binding (enforceable). In the case of the EU PTAs, the areas with WTO+ disciplines are fewer, in particular for export taxes, SPS, TRIMs and services.

The EU PTAs in contrast have much greater coverage of WTO-X provisions, with some PTAs covering over 30 policy areas that are not included in the WTO. Examples include competition policy, environmental laws, investment, IPRs, capital movement, consumer and data protection, cultural cooperation, education, energy, health, human rights, illegal immigration, illicit drugs, money laundering, R&D, SMEs, social matters, statistics, taxation, and visa and asylum policies. The US PTAs are much less focused on WTO-X policy areas, but at the same time there is also much less variance in the subjects that are covered. They are limited to anticorruption (not found in any EU PTA), competition policy, environmental laws, investment, IPRs, labour market regulation and capital movement. Although the EU is clearly the 'market leader' when it comes to WTO-X provisions, about three-quarters of the relevant articles in the EU PTAs do not impose binding disciplines. Instead, they tend to constitute soft law—technical assistance and cooperation. In contrast, the disciplines in WTO-X areas found in US agreements are generally legally binding (that is, enforceable). This reflects a distinct difference in the strategies and approaches pursued by the EU and the US. The US relies more on binding agreements and legal enforcement, the EU has tended to put the emphasis on embedding technical assistance and other forms of cooperation in its PTAs, and supplementing this with financial aid and policy/political dialogue.

This 'softer' approach is a basic premise of the EU's European Neighbourhood Policy, for example (Hoekman, 2007). Which approach will prove to be more influential in setting standards remains to be seen.

From the perspective of members a key question is whether 'deeper integration' provisions are beneficial in and of themselves, or simply a cost that must be paid to obtain preferential access to major markets. We return to that question in Chapter 13 as it is one that obtains in the WTO context as well. From the perspective of nonmembers and the trading system the basic question is the same as discussed above: what is the impact on them of deeper integration among subsets of countries in PTAs? This depends both on the extent of discrimination that is implied and on the scope for cooperation in a PTA context to increase the likelihood that the norms set by PTAs become the de facto standard or the focal point for subsequent multilateral cooperation in the WTO (or elsewhere).

The prospects for the 'multilateralization' of PTA commitments in these areas may be significant. In many cases regulation is quite naturally applied in a nondiscriminatory fashion, treating domestic and all overseas suppliers or firms equally—where 'domesticity' is defined more frequently in terms of location of production than ownership. This is quite different from tariffs and NTBs affecting trade in goods, where domestic/foreign and intra-foreign discrimination is the objective. From the perspective of achieving regulatory objectives, nationality often will (and should) not matter. But even if regulation applies to all sources of supply, it can still have the effect of segmenting markets and reducing competition.

If liberalization—defined as taking actions to enhance the contestability of a market—is more likely to be nondiscriminatory for regulations than for merchandise trade barriers, it is, equally, less likely to come about at all. This is because it is inherently more far-reaching and because it is simultaneously necessary and very difficult to distinguish between regulations that are genuinely needed for the achievement of domestic objectives and those that are oriented towards segmenting markets and protecting domestic incumbents. In practice it is certainly not inevitable that regulations are applied on a nationality-blind basis—insofar as protectionism is an objective of policymakers, regulation can be (and is) used to achieve this. One reason is that the legitimate, nonprotectionist class of regulation frequently requires the acquiescence of domestic firms if it is to be implemented effectively and almost always entails consulting those firms about any reforms. With the complex and subtle nature of many regulations, incumbents (and national regulators) will have a great deal of influence over regulatory structures and details, and may well have veto power over policymakers.

For cooperation on product market regulation and domestic policies in PTAs one can envisage three different processes of multilateralization (Hoekman and Winters, 2009). First, hegemonic multilateralization: a hegemonic economic power is essentially able to impose its own model on its partners, not necessarily

coercively but by the force of its market size. As different partners adopt the hegemon's approach over their own local one, a degree of multilateralism is achieved. And it is possible that as the partners enter further bilateral or regional arrangements with other partners the model is extended. As Schiff and Winters (2003) observe, the accretion of two different groups of supporters around two different models—say a US and a EU model—could make the final multilateral step (harmonization or recognition of equivalence) less rather than more likely. But, if a high degree of similarity or consistency is achieved, goods and services designed for one market can be sold elsewhere, greatly increasing the contestability of markets. Examples of the hegemonic model abound in 'deep integration'. The US requires partners in Bilateral Investment Treaties (BITs) to conform to an identical template and imposes its own intellectual property right (IPR) protection provisions in its PTAs—World Bank (2005). Another example is the EU interest in extending its system of geographical indications through its PTAs (see Chapter 8).

The second route to multilateralism is a convergence route. This operates within a PTA where the erosion of barriers to trade increases the pressure to harmonize regulations because they start to have greater impact on trade patterns, competitiveness and profitability. This is essentially the 'competition between rules' that featured in the EU's Single Market programme, which applies equally to goods and services. It depended, in the case of goods, not only on the removal of traditional barriers to intra-EU trade (tariffs, quotas, etc.) but on the aggressive policy of the European Commission and European Court of Justice towards other limitations on the freedom of movement of goods such as product standards (where the principle of mutual recognition was applied). In services the political sensitivity of the convergence route is evident in the constrained liberalization of cross-border services espoused by the recent Services Directive in the EU and the difficulties that have affected efforts by the EU and the US to make progress in moving towards accepting each other's regulatory norms for specific services as being effectively 'equivalent'. Note that the convergence route also spans a frequently mentioned rationale for PTA-based cooperation in 'non-WTO' areas: PTAs may be a useful forum for experimentation and learning. Successful examples of cooperation between members of a PTA may be adopted in other PTAs (or unilaterally), thus promoting multilateralization over time.

In general, the larger the region or the more important it is as a trading partner, the greater the incentives for a country to adopt the regulatory standards of the PTA. There will often be a link, implicit or explicit, between harmonization of regulatory regimes and the threat of contingent protection. One factor driving harmonization is to reduce the possibility of being confronted with contingent protection. As PTAs increasingly are instruments for such regulatory harmonization, or for the adoption of mutual recognition procedures, the potential cause for concern on the part of nonmembers is obvious. As discussed in Chapter 13, one size fits all is not necessarily optimal.

The third route to multilateralism is the one identified by Baldwin (2006*a*) for trade in goods—what could be called a political evolution route. Here, changes in the political weight of different parties or in the relative importance of different costs change the political economy so that groups which once sought to segment markets now seek to integrate them. This route is, of course, premised on policies being applied in a discriminatory manner vis-à-vis nonmembers of a PTA and thus will not apply (and will not be needed) if policies are applied on an MFN basis. One difference for deep integration is that, compared with restrictions on goods trade, regulations are complex and require greater complicity from the relevant industry. The strong position of incumbents may make liberalization more difficult; in particular, it is difficult to envisage incumbents in a sector seeking the liberalization of that sector. However, offsetting this is that downstream (using) sectors may have stronger incentives to oppose policies that raise the costs of services than is the case with goods.

Hoekman and Winters (2009) argue that when it comes to 'deeper integration' in PTAs, to date most reform is unilateral. There is very little direct evidence that PTAs do a lot to drive reform. One problem is to determine the direction of causality. One cannot infer from the spread of specific PTA disciplines ('templates') that PTAs are driving reforms beyond what governments had already decided was beneficial autonomously. Most research in this area focuses on legal texts, not on extent to which PTAs imply/require or result in changes in national legislation. It may well be that the source for reform has primarily been knowledge and information—the demonstration effects of successful countries or the general focus of academia and the international community on the benefits of deregulation, competition, etc. Maybe the IMF, World Bank, OECD, APEC, etc., which have been advocating for better policies and more transparency for years are the key: the World Bank's *Doing Business* report may well have been a more potent driver of reforms than the PTAs of which countries were members.

Even if developing countries are adopting disciplines in PTAs and applying them on a MFN basis, this does not imply that the norms concerned are beneficial for them. The content of the norms that are included in PTAs obviously matter. Whether these are autonomously decided or externally imposed, they need to benefit the countries that adopt them and the countries affected by them (the nonmembers). From this perspective, another important priority is the establishment of institutions or other means to help developing countries take an informed view of what they are asked to do in PTA negotiations and how neighbours' PTAs impact upon them. The pressure by high-income countries for developing country PTA partners to adopt TRIPS-plus disciplines is an example.

Finally, the opportunity costs of PTAs in terms of taking up scarce administrative capacity of developing country governments needs to be recognized. One can question whether the negotiation of PTAs is the best use of the limited policy attention and human resources that are available in many low-income countries (Schiff and Winters, 2003). Much depends on the content of a PTA and the strategic

use that is made of it. Preferential trade agreements that are not implemented or not used as instruments to realize substantial economic benefits can have a significant opportunity cost in terms of diversion of policy attention and capacity.

10.4. Conclusion

Always a central element of the trade policy strategy of European countries, 'regionalism' has become an important form of international cooperation on trade policy for virtually all the members of the WTO, developed and developing. Although subject to conditions contained in Articles XXIV GATT and V GATS, multilateral disciplines are not enforced. On a number of dimensions they are also weak. An example is the absence of any disciplines with respect to preferential rules of origin in the WTO (Box 10.3). Another is the absence of a requirement that PTAs be open to new members (Bhagwati, 1993). Multilateral surveillance is limited—even if the CRTA were effective, the focus is on WTO tests and not on the economic effects of PTAs. The WTO Secretariat has no mandate to monitor the trade value or terms-of-trade effects of PTAs. Developing countries can opt out of WTO disciplines on PTAs altogether by invoking the Enabling Clause.

Most of the economic literature on PTAs has been theoretical or policy-focused. Rigorous empirical research on the effects of PTAs has been limited until recently. The weight of the empirical analyses of PTAs suggests that if one abstracts from the many PTAs that were never implemented or were designed to have no effect in opening up economies, overall the benefits outweigh the costs, especially if the focus of attention is on the dynamic effects over time. Preferential trade agreements may lead or retard nondiscriminatory reductions in tariffs, but the evidence that is now emerging suggests that PTAs have complemented multilateralism in the sense of promoting lower barriers against the rest of the world.

The proliferation of PTAs has been accompanied by steadily declining barriers to trade generally and high growth rates in world trade. The uniform tariff equivalent of all applied most-favoured-nation tariffs of high-income OECD countries in 2005 was 4.8 per cent. Excluding agricultural products, the figure drops to 2.7 per cent (Kee, Nicita and Olarreaga, 2008). For the developing countries, applied MFN tariffs have also fallen substantially; Kee and colleagues estimate that the median average overall trade restrictiveness index was 7.5 per cent for the 57 countries for which data are available in 2005, compared with 12.3 per cent ten years earlier. Much of this liberalization trend is not due to PTAs. The fact that MFN tariffs have fallen significantly in almost all countries, whatever their participation in PTAs, suggests that unilateral decisions to liberalize have been paramount (Martin and Messerlin, 2007).