

## E. The multilateral trading system and PTAs

A perennial policy question is how the multilateral trading system is affected by the rise of preferential trading agreements (PTAs). Is multilateral trade cooperation compromised by burgeoning regionalism? Should we see these different approaches as complementing or competing with each other? Are there synergies, or inevitable conflicts? Building on the analysis of the report so far, this final section examines these questions.

## Contents

1. Systemic effects of preferential tariff liberalization	166
2. Deep PTA provisions and the multilateral trading system	168
3. Regionalism and the WTO: historical perspective	182
4. The relationship between PTAs and the WTO	187

### Some key facts and findings

- Deep integration is often non-discriminatory in nature.
- Global production networks can result in PTAs with tariff and non-tariff measures that are more consistent with the principles of the multilateral trading system.
- A large number of disputes between PTA members are brought to the WTO dispute settlement system. On average, about 30 per cent of WTO disputes are between members who are parties to the same PTA.
- A critical-mass approach to decision-making in the WTO may be required, at least in the short term, to move forward on an agenda that creates greater coherence between PTAs and the multilateral trading system.

## 1. Systemic effects of preferential tariff liberalization

In the late 1980s and early 1990s, a series of events led analysts to focus on the systemic effects of regional integration (Baldwin, 2009).<sup>1</sup> Regionalism rose in North America, where the Canada-United States PTA was followed by the North American Free Trade Agreement (NAFTA) negotiations. It also reignited in Europe with the Single European Market initiative and the disintegration of the Soviet Union. At the same time, the prospects for a prompt and comprehensive completion of the Uruguay Round were shrouded in uncertainty.

The possibility of a causal link between the expansion of regionalism and difficulties in coming to closure in multilateral negotiations could not be ignored. This turned the regionalism debate into a systemic discussion. This section provides a short overview of the literature in this area, drawing on several surveys that have been published recently: Baldwin (2009), Freund and Ornelas (2010) and Winters (2011).

The broad concern of this literature is the relation between discriminatory and non-discriminatory tariff liberalization. The standard approach is to study whether preferential tariff cuts lead to a reduction or to an increase in the most-favoured nation (MFN) tariff, which is applied by WTO members on a non-discriminatory basis. As discussed in Section C, the evidence so far is not conclusive. However, there are some studies that focus on the effect of preferential tariff liberalization on non-discriminatory tariff liberalization.<sup>2</sup> Due to the paucity of adequate data, opportunities for convincing empirical work are limited. The literature is therefore mostly theoretical, and its predictions are often supported only by anecdotal evidence.

### (a) Do PTAs foster or hinder multilateral tariff reductions?

A number of different mechanisms have been identified through which PTAs could foster or hinder multilateral trade opening.

As discussed in Section C, the Kemp-Wan theorem is a theoretical benchmark showing that PTAs need not have adverse effects on multilateral tariff reductions. Starting from a situation where all countries have MFN tariffs, groups of nations can always raise their collective welfare by forming a trade bloc. A piecemeal enlargement of the bloc will raise bloc members' welfare, and the highest welfare will be reached when all nations are part of the bloc (Kemp and Wan, 1976). This theoretical result rests on two strong assumptions. First, PTA members must set external tariffs at levels that freeze their trade flows with the rest of the world. Secondly, lump-sum transfers between members ensure that they all gain from the PTA.<sup>3</sup>

The fear of preference erosion is an important aspect of the relationship between preferential and multilateral tariff opening.<sup>4</sup> In a world where more open trade would be in the interest of all nations but where individual nations fearing erosion of their preferences would veto it, regionalism can help achieve global trade opening. Baldwin (2009) illustrates the argument with an example where Home country signs separate PTAs with Partner 1 and with Partner 2, thereby forming a so-called hub and spoke system. This system puts Home in a favourable position as it combines opening trade on the import side with preferential tariffs on the export side. Home, the hub, is likely to oppose WTO talks aimed at achieving more open trade for fear of preference erosion. Despite this, Home and its two partners could reach global trade opening, not through multilateral negotiations, but rather through a PTA between the two spokes. As Baldwin shows, the two partners would always prefer global trade opening to the hub-and-spoke situation.

The fear of preference erosion can, however, constitute a potent force of resistance to multilateral tariff reductions. The economic literature has shown that two or more nations can form a PTA which increases their joint welfare at the expense of third nations. Such a PTA will hinder multilateral trade opening because its removal will be resisted by member countries precisely to avoid preference erosion. This can be true not only if PTA members increase their external MFN tariffs, but also when external tariffs are frozen. Baldwin (2009) provides an example in which at a sufficiently low initial tariff, the gains of maintaining a PTA that reduces third-country welfare are worth more than the standard gains of global trade opening.<sup>5</sup>

Developing countries that were granted non-reciprocal preferential access to developed countries' markets are particularly concerned by preference erosion, particularly where reduced advantages from preferential tariffs are not offset by the gains in market access due to tariff cuts on goods that do not receive preferences.<sup>6</sup>

Political economy factors can also affect the pace at which preferential tariffs are extended to non-members on a MFN basis. If PTAs are trade-creating, they will increase the size of export sectors and reduce the size of import-competing sectors. If political power is proportional to the size of the sector, the PTA will increase support for trade opening.<sup>7</sup> In particular, it can make it politically optimal for governments to cut MFN tariffs to levels that would have been undesirable without the PTA.<sup>8</sup>

Along the same lines, if workers have imperfect information on how they will be affected by more open trade, they may initially oppose global trade opening but accept a PTA, which is an intermediate form of trade barrier reduction (Frankel et al., 1995). A PTA

may inform workers on how they will be affected by global trade opening and make an MFN approach politically feasible.

The political economy models discussed in Section C (Grossman and Helpman, 1995; Krishna, 1998), however, offer some insights as to why PTAs might inhibit multilateral tariff reductions. In such models, interest groups might seek primarily trade-diverting PTAs, i.e. agreements that provide enhanced protection.<sup>9</sup> In Krishna's model the extent of trade diversion determines the degree of political opposition to a multilateral agreement that would find support in the absence of the PTA. Intuitively, if there is little or no trade diversion, firms from each member country obtain higher market shares (and profits) in the other member's market but lose domestic profits, with an overall small effect on net profits. However, if the PTA allows bloc firms to displace those from excluded countries in each other's markets, it surely enhances profits for all firms, at the expense of outsiders (Freund and Ornelas, 2010).<sup>10</sup>

The result that specific interest groups might oppose multilateral trade opening that would be supported in the absence of a PTA is also obtained in a median-voter setting by Levy (1997). He shows that a bilateral PTA might offer disproportionately large gains to key agents in a country, making them unwilling to support a multilateral agreement, which would therefore be blocked. This might be the case, for instance, if the two countries have similar factor endowments, so that a lot of trade within the PTA is intra-industry trade, with limited redistributive effects. A move towards multilateral opening would alter domestic factor prices, creating winners and losers and adding only modest gains from increased variety or specialization based on comparative advantage. In this case, the median voter would oppose such a move, and the PTA acts as an obstacle to multilateral trade opening.

Some PTAs may be concluded partly in pursuit of non-economic objectives, such as understanding and reconciliation between former enemies (e.g. France and Germany), or between nations with former colonial links (Schiff and Winters, 1998). As discussed in Section C, some authors have argued that these non-economic objectives might lead member countries to oppose further multilateral trade opening. In a model by Limão (2007), PTAs allow partner countries to extract mutual cooperation on the non-trade issue, using preferential tariffs as bargaining chips. The prospect of dissipating this possibility via multilateral trade opening might make countries less likely to favour a global approach.<sup>11</sup>

PTAs may also increase the adjustment costs associated with multilateral trade opening when firms have to make sunk, sector-specific investments to produce. As shown by McLaren (2002), in such a situation the *ex post* gains from multilateral reductions

can be reduced relative to those from preferential trade opening, and the latter emerges in equilibrium. The reason is the following: if firms expect global trade opening to arise, they will invest in sectors of comparative advantage, so every country will become highly specialized. In this situation, the *ex ante* gains of multilateral trade opening materialize, and such opening is likely to occur. If, however, firms expect a PTA to be signed, they will invest in goods in which excluded countries have a comparative advantage, because external tariffs will render these goods expensive. For similar reasons, firms from excluded countries will invest in goods where PTA members have a comparative advantage. As PTA countries become specialized relative to each other, and less specialized relative to outsiders, the gains from global trade opening will be reduced. As McLaren (2002) explains, the resulting regionalism is "insidious" because it is an inferior outcome for all participants, and it emerges only because it prompts sunk investments that reduce the value of multilateral trade opening.

Finally, opposition to further multilateral tariff opening by PTA members might come from excluded countries. The logic is as follows: if PTA members reduce their external tariffs for political economy reasons after signing an agreement, this might result in pure trade creation. As argued by Ornelas (2005b), non-members benefit from such PTAs by obtaining increased market access to member countries without having to reduce their own tariffs, as would be required under a multilateral agreement. Therefore, non-members may turn against multilateral trade opening that they would support in the absence of the PTA.<sup>12</sup>

The overview of the literature thus suggests that the effect of regionalism on the prospects of multilateral trade opening will depend on a number of factors. The results depend on how much members and non-members stand to gain from a PTA, and how much they would lose from multilateral trade opening, on the importance of political economy considerations in policy formation, and on the extent of lock-in effects of preferential trade opening. Moreover, results depend on whether regionalism is open or not (Yi, 1996); on the presence of dissimilarities in endowments or costs (Saggi and Yildiz, 2009); on the rules of the multilateral trade system (Bagwell and Staiger, 1999; Saggi and Yildiz, 2009); as well as on the formal enforcement constraints (Bagwell and Staiger, 1997a; 1997b).

## (b) Evidence on the systemic effects of regionalism

When the theory is inconclusive, the most natural thing to do is to turn to empirical evidence. A first strand of literature tests whether MFN and preferential tariffs are complements or substitutes.<sup>13</sup> As discussed in Section C, different results emerge for developing and developed countries. While in the former group of

countries preferential trade agreements appear to reduce external tariffs, in the latter group of countries they seem to increase them. Most of the contributions do not distinguish between MFN tariffs that have been negotiated at the multilateral level and unilateral tariff reductions.<sup>14</sup> The notable exceptions are Limão (2006) and Karacaovali and Limão (2008), who explicitly consider the effect of preferential trade opening on multilateral trade opening at the Uruguay Round in the United States and the European Union, respectively.<sup>15</sup>

A second strand of literature investigates the correlation between PTA formation and multilateralism. One often-used example of regionalism promoting multilateral trade opening is when the United States, which for many years had been advocating multilateralism, converted to regionalism in the 1990s and thereby revived the Uruguay Round negotiations (Bergsten and Schott, 1997).<sup>16</sup> Mansfield and Reinhardt (2003) observe that more PTAs are formed during multilateral negotiations than at other times. They interpret this result as evidence consistent with multilateralism promoting PTAs as devices to obtain bargaining leverage within the multilateral regime (pressuring outsiders to open their markets or escaping from free-riders).

A general problem with the approach of linking PTAs with multilateral trade rounds is that the latter are rare events. Moreover, the practice of multilateral trade rounds is to negotiate multilateral opening with more or less ambitious scenarios of trade opening, rather than opting for full or no multilateral opening. Therefore, a direct test of whether PTAs decrease the likelihood of signing multilateral trade opening agreements is impossible (World Trade Organization (WTO), 2007).

Anecdotal evidence can be found in support both of the view that PTAs facilitate further multilateral trade opening and of the view that they hinder it.<sup>17</sup> On the one hand, there is anecdotal evidence that PTAs increase excluded countries' incentive to move on the multilateral front to avoid trade diversion. A related argument is that the last three rounds of multilateral trade negotiations have started in tandem with major moves towards regional integration, which is sometimes taken as evidence of the building block relationship between the two processes. Furthermore, the cost from overlapping PTAs can trigger a rationalization of the system – as in the case of the Pan-European Cumulation System – or a recourse to the multilateral system – as in the case of the WTO Information Technology Agreement.<sup>18</sup>

On the other hand, it has been argued that the concern for preference erosion has contributed to the stalling of multilateral negotiations and has actually been reflected in less multilateral trade opening, see for instance Curtis and Vastine (1971). Furthermore, there is also evidence that the engagement in regional

negotiations may stall the process of multilateral trade opening by absorbing resources away from the multilateral negotiations (World Trade Organization (WTO), 2007).

## 2. Deep PTA provisions and the multilateral trading system

While the literature on the systemic effects of preferential tariffs is rich and very active, so far there has not been much research on the systemic effects of other, “deep” integration, provisions. Available results suggest that in some deep integration areas, such as technical barriers to trade (TBT), multilateral regulation may not be economically optimal or politically feasible. Because deep integration is often MFN in nature, however, such regulation may also be less necessary. Indeed, the literature has identified a number of mechanisms through which deep integration “automatically” supports further opening, or at least does not entail negative static effects on the multilateral trading system.

### (a) Deep integration is often non-discriminatory in nature

By their very nature, some deep integration provisions are *de facto* extended to non-members because they are embedded in broader regulatory frameworks that apply to all. An example is provided by services trade opening. Barriers to trade in services are generally behind-the-border, regulatory measures. Even though some services barriers could in practice be applied in a differentiated manner depending on the suppliers' country of origin (e.g. restrictions on the movement of persons, foreign equity restrictions, or foreign direct investment screening), one expects that barriers removed or relaxed as a result of a PTA be extended *de facto* to non-parties. This also makes most economic sense, and may limit any economic distortion resulting from services PTAs.<sup>19</sup>

Evidence suggests that in certain cases, preferential treatment was granted to PTA parties, but proper analysis of this is made difficult by the absence of comprehensive information on the treatment applied by countries to services and suppliers of their trading partners. This is compounded by the fact that analysis of non-discriminatory treatment in services would need to consider not only treatment specified in laws and regulations, but also *de facto* treatment – for example, which suppliers receive operating licences, which are sometimes limited in number. Furthermore, given the importance of first-mover advantage for suppliers in a number of services sectors,<sup>20</sup> what matters is whether non-preferential treatment is available for all suppliers of different origins from the moment trade opening takes place. While this may well be the situation most of the time, information is lacking.

The fact that services commitments in PTAs can be non-discriminatory also suggests that any technical or economic obstacle to the multilateral extension of such PTA commitments as part of the Doha Round would be limited. It can be hoped that preferential commitments made by several WTO members make their way into these members' conditional offers and inject momentum in the Doha services negotiations. This has not happened in offers currently on the table – which for the most part were submitted in 2005 – therefore suggesting that other factors are at play, either within the Doha negotiations or domestically. One such factor may be that, in the context of the growing number of preferential trade agreements in recent years, a number of countries may wish to keep leverage for their PTA negotiations, where commitments that go beyond the General Agreement on Trade in Services (GATS+ commitments) are exchanged as part of the overall trade-off between parties (e.g. against preferential goods access), even though the resulting overall outcome is less economically significant than what the Doha Round can produce, including for these PTA parties.

Another factor to consider is that rules of origin (RoOs) for services do not carry the same potential for distortion as they do for goods trade. RoOs in services PTAs are usually liberal, along the lines of GATS Article V(6),<sup>21</sup> although there are certain exceptions.<sup>22</sup> This reduces the extent of the spaghetti bowl effect (see Section C).

For mode 1 (cross-border supply), PTAs generally focus on the territorial presence of the provider rather than on its nationality or the origin of the service, according to origin status to the services provided by entities located in a PTA partner nation. For mode 2 (consumption abroad), the supplier's nationality is unimportant as well; the focus is on the territory in which the service is supplied and consumed. For mode 3 (commercial presence), RoOs typically accord origin status to firms with “substantive business operations” within the PTA region, irrespective of the nationality of business owners. In other words, the only requirement is to establish a legal presence and a certain level of commercial activity in one of the PTA members.<sup>23</sup>

In other areas, such as mutual recognition agreements (MRAs) on testing, RoOs are absent. If two nations (for example, the United States and Singapore) sign an agreement whereby the United States accepts products tested in Singapore laboratories, independently of their origin, Singapore can become a regional hub for testing and conformity assessment. Neighbouring countries can ship their products there to be certified before being exported to the United States. The lack of RoOs automatically multilateralizes the bilateral testing MRA, reducing the spaghetti bowl effect (Baldwin et al., 2009).

Competition policy provisions in PTAs are also mostly characterized by non-discrimination (Teh, 2009; Dawar and Holmes, 2010). Competition disciplines usually operate through the use of domestic regulations. While it is not impossible for these regulations to be tailored to favour enterprises originating from PTA partners, it may be costly to do so and becomes even more difficult as the number of PTAs to which a country is a signatory increases. Transparency and in particular the obligation to publish laws promoting competition will provide information that becomes (simultaneously) available to PTA and non-PTA members alike.

The substantive obligations in the competition policy chapters of PTAs generally involve applying competition law or setting up a competition authority. To the extent that enforcement of competition law in a country reduces the market power of domestic incumbents, the prospects of foreign enterprises, whether they are from a PTA member or not, are improved. Carrying out the competition obligations also opens up opportunities for new foreign entrants (either from PTA or non-PTA members) to challenge domestic incumbents.

Moreover, there are positive effects from competition provisions, particularly if they are contained in regional agreements (Dawar and Holmes, 2010). There can be economies of scale from the creation of a regional competition authority. Even if no centralized authority is established, benefits can come from information-sharing and cooperation among enforcement authorities. There could be demonstration effects to other jurisdictions when a competition authority in one PTA member takes action against anti-competitive behaviour. Eventually, more common competition norms and practices within the PTA will prevent regulatory arbitrage, where enterprises locate themselves in a jurisdiction in the PTA with relatively lax competition policy.

Finally, PTAs may directly refer to WTO rules. Lesser (2007) argues that the majority of technical barriers to trade (TBT) provisions in PTAs signed after 1995 reaffirm the parties' rights and obligations under the WTO TBT Agreement and make reference to its objectives.

Furthermore, most transparency commitments included in PTAs are similar in nature to the ones included in the WTO TBT Agreement. Finally, provisions that require parties to provide an explanation in case of non-recognition of standard-related measures and mechanisms supporting further cooperation among parties (e.g. technical assistance, joint standardization) can in fact support and enhance the implementation of the WTO TBT Agreement, supporting the multilateral trading system.

### Box E.1: Investment provisions in international agreements: is there a potential for third-party discrimination?

The process of gradual opening of foreign direct investment (FDI) has been the outcome of a multi-layered process combining autonomous MFN investment opening, commitments made in the context of bilateral investment treaties (more than 2,700 to date),<sup>24</sup> and only more recently commitments made in PTAs. Despite the progress in investment provisions in PTAs, investment remains overwhelmingly regulated by bilateral investment treaties (BITs).

Investment provisions are typically included in PTAs to foster investment flows between member countries. Some provisions are clearly aimed at protecting investors, without increasing barriers to investment from third countries (Baccini and Dür, 2010). The investment chapters of PTAs normally include absolute standards of treatment providing a minimum level of protection for investors. In many cases, they reflect the actual state of domestic legislation concerning FDI and the level of commitment achieved in earlier BITs. The provisions regarding investment protection are either directly included in the text of the agreement, such as in the agreements signed by the United States, or they are indirectly referred to in agreements providing that investors should be treated in accordance with customary international law (Kotschwar, 2009).

It has been noted, however, that the creation of a PTA may be a source of investment discrimination, whereby potential investors from excluded countries are put at a disadvantage vis-à-vis investors from member countries. This can occur through two channels: one direct and the other one indirect (Baccini and Dür, 2010). First, investment discrimination can result directly from the inclusion of provisions that open up certain sectors for investment only on a preferential basis. All PTAs include relative standards of treatment, namely MFN and national treatment (NT).<sup>25</sup> Most recent PTAs, including the ones signed by the United States and the ones among Asian countries, tend to provide both MFN and NT during all phases of the investment (pre- and post-establishment).<sup>26</sup> Relative standards of treatment can provide a competitive advantage to investors from member countries vis-à-vis investors from non-member countries, especially in the services sector. For instance, the PTA between Australia and the United States relaxes the requirements for government screening of FDI for US companies investing in Australia (Baccini and Dür, 2010).

Secondly, investment discrimination can result indirectly from discriminatory tariff reductions. Assume firms from countries A and B are engaged in market-seeking FDI in country C. They source inputs domestically, and import them into C at the MFN tariff  $\tau_C$ . A PTA between A and C, that eliminates tariffs on intermediary inputs from A, creates investment discrimination by putting investors from country B at a competitive disadvantage. However, there is very little empirical evidence on the actual incidence of such discrimination.

The extent of potential investment discrimination also depends on the RoOs included in the PTA. Liberal RoOs in the services sector, for instance, reduce the discriminatory aspects of investment provisions for services providers. There is, however, considerable variation in the strictness of rules of origin for investment across PTAs (Baccini and Dür, 2010). Moreover, one should consider the relation between the provisions of PTAs and the ones contained in BITs.

BITs are traditionally about the protection of investment that is already established in the host countries (DiMascio and Pauwelyn, 2008), guaranteeing compensation in cases of expropriation and repatriation of profits. In the early BITs, what mattered for host country governments was the flexibility to differentiate between national and foreign governments, not so much among foreign investors. Nonetheless, a host country could wish to exercise selective screening over the admission of foreign investors and the terms of their admission as part of its policies to promote national investments. For example, it could wish to offer investment incentives only to certain foreign investors on a discriminatory basis. Despite an improvement in absolute standards of treatment in recent BITs, most of them still do not cover pre-establishment or entry of investments, according NT and/or MFN only once investments are in the country. For this reason, and also because they do not cover tariff reductions, Baccini and Dür (2010) argue that BITs are not very likely to lower PTAs' potential for investment discrimination.

It should be noted that investment discrimination need not imply a reduction in FDI flows from excluded countries into member countries. Tariff discrimination may lead to tariff-jumping FDI (i.e. the establishment of a production facility in a member country, through FDI, in order to avoid a tariff). Studies finding that PTAs attract FDI from third countries, such as te Velde and Bezemer (2006), do not, therefore, provide evidence against PTA-driven investment discrimination.

(b) Several mechanisms supporting further liberalization are found in PTAs

First, PTAs may include “non-party” MFN clauses. These clauses stipulate the extension to current PTA partners of preferences or concessions that member countries may have granted in the past or may grant in the future to third nations.<sup>27</sup> In the case of services and government procurement for instance, such provisions ensure that future and more advantageous commitments with other non-member partners should be granted to PTA partners as well (Fink and Molinuevo, 2008). Many PTA procurement provisions require third-party MFN guarantees so as to limit the extent to which preferential procurement is undermined by subsequent PTAs (Baldwin et al., 2009).<sup>28</sup>

Secondly, there is a tendency to replicate trade-opening rules in PTAs because template approaches are often used for PTAs. The spread of the NAFTA-style telecommunication competition provision is an example. Baldwin et al. (2009) argue that the large number of countries that have included this provision in PTAs suggests that it is progressively becoming a norm. They further argue that harmonization to a single regulatory regime, including a common set of rules that governments apply to private firms in many nations, tends to foster competition and trade and it cannot be considered preferential.

Another example is provided by NAFTA's investment provisions, in particular performance requirements. These provisions have spread in Latin America and beyond. Fifteen countries have agreed never to apply performance requirements against foreign investors from any jurisdiction. Another 36 countries have committed to forgo the application of such requirements, however only against Canadian and US investors (Baldwin et al., 2009).

Along similar lines, as argued by Anderson et al. (2010), “the government procurement provisions of RTAs have made feasible a significant further expansion of the membership of the Government Procurement Agreement (GPA), in the event that parties decide to take this step.”

Thirdly, domino effects (Baldwin, 1993) pointing in the direction of progressive extension of preferential market access might be at play also for deep integration provisions. Consider the example of the GPA. With the EU enlargement from 15 to 25 members, non-EU GPA members started facing more competition in government procurement both in the 15 EU incumbents (from the ten newcomers) and in the ten EU newcomers (from the 15 incumbents). As a reaction to this form of trade diversion, the non-EU GPA members started pressuring the new EU members to join the GPA.<sup>29</sup> Similar domino effects can be discerned in all cases in which countries excluded from a PTA find themselves in

a position to adopt similar provisions to the ones adopted by member countries to avoid trade diversion. The implementation by countries in the European Free Trade Association (EFTA) of competition policy norms that mimic the ones of EU countries can be interpreted as a way of ensuring that firms in EFTA countries do not find themselves at competitive disadvantage vis-à-vis firms in the European Union (Baldwin et al., 2009).

(c) The effects of global production sharing

The presence of international fragmentation of production can alter political-economy forces in favour of the adoption of tariff and non-tariff measures that are less discriminatory, and more consistent with the principles of the multilateral trading system. The underlying logic can be explained with the example of the Pan-European Cumulation System (PECS) of rules of origin (Baldwin et al., 2009).

Firms from EU countries started to relocate labour-intensive stages of production in low-wage neighbouring nations from the 1990s. At the same time, the European Union engaged in bilateral agreements with a number of countries both from Central and Eastern Europe and from the Southern Mediterranean. These agreements contained non-harmonized rules of origin, giving rise to a spaghetti bowl effect that restricted firms' ability to source intermediate goods from the cheapest source (Gasiorek et al., 2009).

Moreover, the downsizing of production in the European Union, also due to competition from emerging Asian countries such as China, reduced the number and political influence of EU-based producers of intermediate inputs which benefited from the protectionist effects of the spaghetti bowl. The political economy forces thus turned in favour of harmonizing rules of origin across PTAs, to avoid the cost of different administrative requirements, and permitting diagonal cumulation (i.e. allowing EU final good producers to source inputs from a wider set of countries without fear of losing origin status). This was accomplished with the signing of the PECS in 1997.<sup>30</sup>

International fragmentation of production may also be a driver of deep integration, and of the multilateral extension of deep provisions. Examples can be found in the field of technical barriers to trade (TBTs), the opening of markets for trade in services and the presence of contingency measures within trade commitments (Baldwin et al., 2009). In TBTs, unbundling of production may help explain the adoption of international standards, at least in parts and components, in industries characterized by global sourcing (e.g. electronics). Concerning the opening of markets for trade in services, offshoring is likely to create an incentive for nations to apply international standards to improve the competitiveness of their own exporters and to make their own services markets more attractive to foreign investors.



**Box E.2: Making rules of origin more compatible with the multilateral trading system**

It has been argued in this report that rules of origin (RoOs) are likely to strengthen the “spaghetti bowl” effect of PTAs. In view of this adverse effect, various commentators have argued in favour of reforming RoOs, making them more transparent and compatible with the principles of the multilateral trading system (see for instance Cadot and de Melo, 2007).<sup>31</sup> This box discusses the system of “cap and convergence” proposed by Estevadeordal et al. (2009a) and supported by Baldwin and Thornton (2008), based on the two concepts of “multilateralization” and “convergence”.

“Multilateralization” of RoOs refers to the establishment of multilateral rules that limit the restrictiveness and complexity of RoOs in PTAs (Estevadeordal et al., 2009a). According to the authors, such rules would ensure that “at least the qualifying production methods in a given sector remain relatively similar across export markets”. They claim that multilateralization should ideally be coupled with “convergence”, which is the “unification of multiple overlapping existing RTAs into a single cumulation zone with a new, single list of rules of origin”, like in the European PECS.

The proposed system of “cap and convergence” would increase transparency (one of the key principles of the multilateral trading system). Moreover, it could be subject to WTO discipline. Estevadeordal et al. (2009a) suggest that the non-preferential RoOs currently negotiated at the WTO could serve as the global benchmark with which to compare the overall restrictiveness of RoOs of a given PTA. This would be analogous to the General Agreement on Tariffs and Trade (GATT) Article XXIV restriction on a customs union's external tariff, which caps it at the average of the tariffs previously charged by the members (Baldwin and Thornton, 2008). This provides another strong reason for concluding the long-standing negotiations on non-preferential rules of origin at the WTO.

The rationale for coupling convergence with capping is the following: larger cumulation zones increase trade, especially among the current spoke countries (see Section C). However, observed restrictiveness of RoOs is positively correlated with the size of the cumulation zone, measured as the combined GDP of members (Estevadeordal et al., 2009b). Larger cumulation zones could therefore end up with highly restrictive RoOs that would serve to isolate production within each zone, increasing trade diversion and reducing global efficiency. Trade diversion for third nations justifies involvement of the WTO through multilateralization efforts aimed at limiting the overall restrictiveness of RoOs within a given cumulation zone.

Finally, unbundling of production may create greater support for new multilateral rules on contingency measures, such as safeguards, anti-dumping and countervailing measures, in trade commitments. When firms engage in outsourcing, they prefer measures discouraging the imposition of contingency measures in as many bilateral trading relationships as possible, rather than in any one bilateral trade relationship. This underlies the producer support for the spread of a common or similar set of rules on the application of contingency measures (Baldwin et al., 2009).

#### (d) Relationship between the WTO and PTA dispute settlement systems

As noted in Section D, the vast majority of PTAs establish some kind of dispute settlement mechanism. Porges (2010) presents a survey of dispute settlement mechanisms in PTAs. She describes these mechanisms as generally falling into the following three types: (i) diplomatic or political mechanisms (such as the Latin American Integration Association, ALADI); (ii) standing tribunals (such as the European Union and the Andean Community); and (iii) referral to ad hoc panels (such as NAFTA and other US FTAs, EU FTAs with Chile, the Republic of Korea and Mexico, the Association of Southeast Asian Nations Enhanced

Dispute Settlement Mechanism, and the Southern Common Market – MERCOSUR). The survey indicates that referral to ad hoc panels is the dominant model for PTA dispute settlement mechanisms. A slightly different classification is used in Ramirez Robles (2006), which classifies the mechanisms as: (i) diplomatic; (ii) quasi-judicative; and (iii) “hybrid”, (i.e. mechanisms that have features of both models).

The relationship between the WTO and PTA dispute settlement mechanisms has received considerable attention in the trade literature and some commentators have cautioned about potential risks from the coexistence of dispute settlement mechanisms at different levels (multilateral, regional and bilateral) that may have overlapping jurisdictions. In this subsection, we first describe how the jurisdictions of the WTO and PTA dispute settlement systems may overlap. We then discuss the concerns that have been raised and the recommendations that have been made to reduce the risks of conflict. This is followed by a review of the handful of WTO disputes in which the relationship of the WTO dispute settlement system and a PTA dispute settlement mechanism has been raised as an issue. Finally, we present data on the use of the WTO dispute settlement system by members who are partners in a PTA.

*(i) Overlapping jurisdictions*

Article 23.1 of the WTO's Dispute Settlement Understanding (DSU) provides that "(w)hen Members seek the redress of a violation of obligations or other nullification or impairment of benefits under the covered agreements or an impediment to the attainment of any objective of the covered agreements, they shall have recourse to, and abide by, the rules and procedures of this Understanding." The Appellate Body has explained that "Article 23.1 lays down a fundamental obligation of WTO Members to have recourse to the rules and procedures of the DSU when seeking redress of a violation of the covered agreements" and "establishes the WTO dispute settlement system as the exclusive forum for the resolution of such disputes"<sup>32</sup> (Appellate Body Report, *US / Canada – Continued Suspension*, para. 371).

Recourse to the WTO dispute settlement system may be had where a WTO member considers that any benefits accruing to it directly or indirectly under the WTO agreements are being impaired by measures taken by another member. Thus, in principle, a WTO member may not have recourse to the WTO dispute settlement system to prosecute an alleged violation of a PTA obligation.<sup>33</sup> The potential for overlapping jurisdiction arises where an issue is regulated both under the WTO and the PTA. Porges (2010) observes that "(a)lmost all PTAs overlap with the WTO Agreement, as both PTAs and the WTO require national treatment and ban quantitative restrictions on trade. Indeed, many PTAs simply incorporate GATT Articles III and XI by reference".

PTAs take different approaches to how they regulate the relationship between their own dispute settlement mechanism and that of the WTO. Porges (2010) identifies the following four approaches. Most PTAs use the "fork-in-the-road" approach which allows the party initiating the dispute to choose between the multilateral or the PTA fora. However, once it has initiated the dispute in one forum, the other option (be it the PTA mechanism or multilateral one) is no longer available to it. (See, for example, the NAFTA and the Colombia-EU PTA.) The NAFTA has a provision (Article 2005(4)) under which the respondent party may require an environmental dispute to be addressed at the regional level, even if the complaining party has initially chosen the multilateral fora. This provision is the subject of a pending dispute between the United States and Mexico (discussed further below). A third approach, which has been used in far fewer PTAs, is to establish the PTA dispute settlement mechanism as the exclusive forum where the matter is one regulated under the PTA. The EU-Mexico and EU-Chile PTAs take the opposite approach, requiring disputes involving a breach of a PTA obligation that are equivalent in substance to a WTO obligation to be brought to the WTO (Porges, 2010).

There are many factors that can influence a country's decision to bring a dispute to one forum over the other where the choice is available to it. Horlick and Piérola (2007) examine a list of factors that may be relevant, including: the type of measure that is being challenged, the applicable law, issues of standing, the time-frame of the proceedings, the remedies available, and the possibility of other countries participating in the dispute as third parties. According to Horlick and Piérola (2007), "the cautious decision-making process to choose the appropriate forum requires weighing and balancing of all these factors in accordance with the ultimate needs and objectives of the complainant".

*(ii) Concerns over the coexistence of the WTO dispute settlement system and PTA dispute settlement mechanisms*

The concerns raised about the coexistence of the WTO dispute settlement system and the increasing number of dispute settlement mechanisms of PTAs revolve around two sets of issues. The first set of issues derive from the view that the proliferation of PTA dispute settlement mechanisms could undermine the WTO dispute settlement system's status as a public good. Those who hold this view consider that the WTO dispute settlement system has positive externalities for members that are not parties to a particular dispute.

Drahos (2005), for example, notes that the interpretation of the WTO agreements provides greater certainty to WTO rules. He also observes that when a respondent member brings an infringing measure into conformity with its WTO obligations, this will be of benefit to the membership at large because of the MFN principle. Thus, Drahos (2005) proposes that where a dispute concerns a matter regulated under both the WTO and the PTA, it be brought to the WTO. Davey and Sapir (2009) take a different approach and propose that the WTO should require members that do not belong to a PTA to be allowed to participate in the PTA dispute settlement forum as third parties.

The other set of concerns relates to the possibility that a dispute is brought under both the WTO and PTA dispute settlement mechanisms. Here there is concern over the inefficiency of litigating similar matters twice and more importantly about fairness to the respondent party that would have to defend itself in two fora (see Kwak and Marceau, 2006). There is also concern about the more extreme situation in which the WTO and PTA fora issue parallel or consecutive conflicting decisions. One way of reducing the risks of this happening is through stricter jurisdictional clauses in PTAs that preclude a dispute from going to both fora or foreclose bringing a dispute to the WTO over a matter regulated under the PTA (Marceau and Wyatt, 2010). This raises, however, the question of the extent to which such clauses would bind WTO adjudicatory bodies.

At the other extreme, there is the risk that the jurisdiction of the WTO could be gradually “carved out”. For the moment, it appears that few PTAs completely close off access to the WTO dispute settlement, but rather leave the choice of forum to the complaining party. The data discussed below show that an important number of disputes between members that are partners in a PTA continue to be brought to the WTO dispute settlement system. Some could also conceive of making changes to the WTO's Dispute Settlement Understanding to regulate the relationship with dispute settlement fora of PTAs. This approach, however, has not been taken up by WTO members in the negotiations to improve the Dispute Settlement Understanding currently under way.

The academic literature discusses other more complex arrangements that could minimize the risks of conflicts and promote more coherence between the multilateral dispute settlement system and the dispute settlement systems of PTAs. For example, there has been discussion of making exhaustion of PTA dispute resolution procedures a prerequisite to initiation of WTO dispute settlement (see Kwak and Marceau, 2006). Another suggestion is to create a system of preliminary references from the dispute settlement systems of PTAs to the WTO dispute settlement system where the issue concerns the interpretation of provisions of the WTO agreements (Kuijper, 2010).

Commentators have also referred to several international law doctrines that could be used to avoid or resolve conflicts between overlapping jurisdictions.<sup>34</sup> The doctrine of *res judicata* or finality refers to situations where a matter has been decided by a competent adjudicative body barring its relitigation in subsequent proceedings. *Lis Alibi Pendens*, for its part, refers to parallel proceedings and is a principle pursuant to which once a dispute is pending in one forum, it cannot be brought before another forum. However, for these doctrines to apply, there must be an “inextricable link” between the proceedings, which usually is understood as an identity of the parties and of the issues (Shany, 2005). Thus, application of the doctrines can be avoided in certain circumstances.<sup>35</sup>

Under the principle of comity or *forum non conveniens*, an adjudicative body could seek to avoid exercising jurisdiction over a dispute if it considers that it would be more appropriate for another tribunal to exercise jurisdiction. There is considerable debate as to the applicability of these principles to resolve a potential conflict of jurisdiction involving the WTO dispute settlement system and a PTA dispute settlement mechanism (see Kwak and Marceau, 2006). The WTO dispute settlement system is available to WTO members as of right; they do not have to seek leave to start the process under the current rules. Thus, some would consider that applying these prerequisites could only be effected through a change in the rules.

As discussed below, questions about the relationship between the WTO dispute settlement system and PTA dispute settlement mechanisms have come up in only a handful of WTO disputes. It should be noted that so far concerns over potential conflicts have not materialized to the extent that some had feared.<sup>36</sup> This is not to say that it is not important to think through issues arising from the coexistence of the multilateral and PTA settlement systems.

### (iii) *Issues relating to PTA dispute settlement raised in WTO disputes*

As noted earlier, issues touching on the relationship of the WTO dispute settlement system and PTA dispute settlement mechanisms have come up in a handful of WTO disputes. In *Argentina – Poultry*, Argentina argued that Brazil was “estopped” from pursuing the dispute at the WTO because Brazil had first challenged the anti-dumping measures in the MERCOSUR forum. The panel rejected Argentina's argument, noting that there was “no evidence on the record that Brazil made an express statement that it would not bring WTO dispute settlement proceedings in respect of measures previously challenged through MERCOSUR”. Moreover, the panel found that:

“In particular, the fact that Brazil chose not to invoke its WTO dispute settlement rights after previous MERCOSUR dispute settlement proceedings does not, in our view, mean that Brazil implicitly waived its rights under the DSU. This is especially because the Protocol of Brasilia, under which previous MERCOSUR cases had been brought by Brazil, imposes no restrictions on Brazil's right to bring subsequent WTO dispute settlement proceedings in respect of the same measure. We note that Brazil signed the Protocol of Olivos in February 2002. Article 1 of the Protocol of Olivos provides that once a party decides to bring a case under either the MERCOSUR or WTO dispute settlement forums, that party may not bring a subsequent case regarding the same subject-matter in the other forum. The Protocol of Olivos, however, does not change our assessment, since that Protocol has not yet entered into force, and in any event it does not apply in respect of disputes already decided in accordance with the MERCOSUR Protocol of Brasilia. Indeed, the fact that parties to MERCOSUR saw the need to introduce the Protocol of Olivos suggests to us that they recognised that (in the absence of such Protocol) a MERCOSUR dispute settlement proceeding could be followed by a WTO dispute settlement proceeding in respect of the same measure.” (Panel Report, *Argentina–Poultry*, para. 7.38)

Alternatively, Argentina argued that if Brazil were entitled to bring the dispute to the WTO, “then the Panel is bound by the earlier MERCOSUR ruling on the measure at issue in this case” as “the earlier MERCOSUR ruling is part of the normative framework to be applied by the Panel as a result of Article 31.3(c) of the *Vienna Convention*”. This argument was also rejected by the panel, which explained its reasons as follows:

“Rather than concerning itself with the interpretation of the WTO agreements, Argentina actually argues that the earlier MERCOSUR Tribunal ruling requires us to rule in a particular way. In other words, Argentina would have us apply the relevant WTO provisions in a particular way, rather than interpret them in a particular way. However, there is no basis in Article 3.2 of the *DSU*, or any other provision, to suggest that we are bound to rule in a particular way, or apply the relevant WTO provisions in a particular way. We note that we are not even bound to follow rulings contained in adopted WTO panel reports, so we see no reason at all why we should be bound by the rulings of non-WTO dispute settlement bodies.” (Panel Report, *Argentina – Poultry*, para. 7.41)

The panel report in that case was not appealed.

The issue also arose in *Mexico – Taxes on Soft Drinks*, where the United States was challenging certain tax measures and book-keeping requirements imposed by Mexico on soft drinks and other beverages that used sweeteners other than cane sugar. Mexico argued that the WTO dispute was “inextricably linked to a broader dispute regarding access of Mexican sugar to the United States’ market under the NAFTA.” Mexico requested the panel to decline jurisdiction over the dispute. According to Mexico, WTO panels have “implied jurisdictional powers” and these include “the power to refrain from exercising substantive jurisdiction in circumstances where ‘the underlying or predominant elements of a dispute derive from rules of international law under which claims cannot be judicially enforced in the WTO, such as the NAFTA provisions’ or ‘when one of the disputing parties refuses to take the matter to the appropriate forum!’”

The Appellate Body affirmed the panel’s finding that, under the DSU, it had no discretion to decline to exercise its jurisdiction in that case. Before reaching this finding, however, the Appellate Body noted that Mexico had not argued that the subject matter nor the respective positions of the parties were identical in the NAFTA and WTO disputes and Mexico had not identified a legal basis that would allow it to raise, in a WTO dispute settlement proceeding, the market access claims Mexico was pursuing under NAFTA. Furthermore, it was undisputed that no NAFTA panel

had yet decided the “broader dispute” to which Mexico had alluded and Mexico had acknowledged that the “exclusion clause” of Article 2005(6) of NAFTA had not been exercised. Thus, the Appellate Body did not “express any view on whether a legal impediment to the exercise of a panel’s jurisdiction would exist in the event that features such as those mentioned above were present.” (Appellate Body Report, *Mexico – Taxes on Soft Drinks*, paras. 44-57)

Another case that has been discussed in the literature is a dispute between Canada and the United States over the imposition by the latter of anti-dumping and countervailing duties on imports of softwood lumber from the former. Various aspects of this dispute were the subject of litigation in both the WTO and NAFTA. At one point an injury determination made by the US investigating authority was found to be lacking by a NAFTA panel, while a WTO panel upheld it. The conflict nevertheless was eventually resolved when the decision of the WTO panel was eventually overturned upon review by the Appellate Body (Hillman, 2009).<sup>37</sup>

The relationship between the dispute settlement mechanisms of NAFTA and the WTO has surfaced again in a more recent dispute between Mexico and the United States. In 2009, Mexico requested that a WTO panel examine the consistency of certain requirements concerning the labelling in the United States of tuna products as “dolphin safe” (WT/DS381/4). In response, the United States invoked Article 2005(4) of NAFTA, which it considers to require that in certain types of disputes, if the defending party makes such a request, NAFTA rather than any other forum should be the sole venue of the dispute. The United States initiated a dispute under NAFTA challenging Mexico’s decision not to move the dispute from the WTO to NAFTA, as requested by the United States (United States Trade Representative (USTR), 2010). Both proceedings are presently ongoing.

#### (iv) *WTO disputes between WTO members that are partners in a PTA*

In this subsection, we examine data on WTO disputes between WTO members who are partners in a PTA. Data on the number of disputes refer to requests for consultations, which is the first step under the WTO dispute settlement procedures. The data concern participation by WTO members (who are PTA partners) as complainants and respondents, and does not include participation as third parties. Moreover, the exercise looks only at WTO dispute settlement and does not examine whether the disputes could have been brought under the PTA dispute settlement mechanism. Certainly a more complete analysis would require looking at whether the disputes could have been taken to the PTA dispute settlement mechanism. Notwithstanding this limitation, the data provide some useful insights.

First, the data show that WTO members that are partners in a PTA continue to have frequent recourse to the WTO dispute settlement system to resolve trade disputes (the methodology employed in Tables E.1 to E.3 and Figure E.1 is explained in Box E.3). As illustrated in Table E.1, 82 of the 443 disputes brought to the WTO up to 2010 were between complainant and respondent members who at the time were partners in a PTA. Disputes between PTA partners represent 19 per cent of all disputes. The ratio is higher where the complainant is a developing country (28 per cent) than when it is a developed country (13 per cent). This is probably explained by the fact that the United States, the European Union, Japan and China do not have PTAs between them, and they have been parties in an important number of disputes.

The largest share of the disputes between PTA partners brought to the WTO is made up of disputes between parties to NAFTA, but there also have been WTO disputes between WTO members that are partners in other PTAs, as illustrated in Figure E.1.

As depicted in Table E.2, the share of WTO disputes between PTA partners increased steadily since 1995, reaching a peak of 50 per cent in 2005. Since then, the share has remained around 30 per cent, although it was significantly below this number in 2009. The steady increase in the share of disputes between PTA partners may be partly a reflection of the negotiation of new PTAs, but is more likely a reflection of the diversification of parties making use of the WTO dispute settlement system. An interesting point that

**Table E.1: Frequency of requests for consultations, by development level and existence of PTAs in force between the parties, 1995-2010** (Total number of pairs of members/pairs with a PTA in force)

		COMPLAINANT			
		Developed	Developing	LDC	TOTAL
DEFENDANT	Developed	154 / 24	115 / 10	0 / 0	269 / 34
	Developing	102 / 8	71 / 39	1 / 1	174 / 48
	LDC	0 / 0	0 / 0	0 / 0	0 / 0
	TOTAL	256 / 32	186 / 49	1 / 1	443 / 82

Source: WTO Secretariat based on Legal Division's and RTA's databases. The table takes account of 419 requests for consultations under the WT/DS document series as of 31 December 2010, which account for a total of 443 pairs of members (i.e. complainant-defendant). See Box E.3.

**Figure E.1: PTAs in force at the time of the request for consultations, 1995-2010**

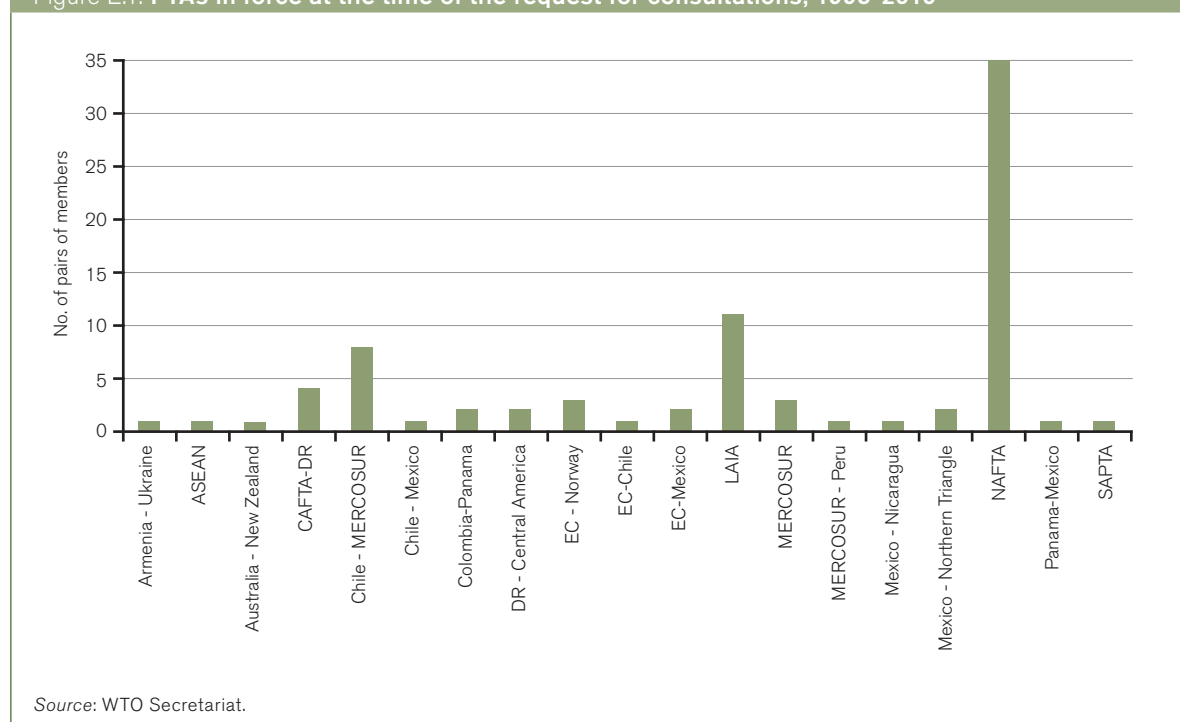


Table E.2: Requests for consultations, by year and subsequent procedures, 1995-2010

Year of request for consultations	Request for consultations				With a panel established			
	Total requests for consultations	Total pairs of members	Pairs w/ a PTA in force		Total panels established	Total pairs of members	Pairs w/ a PTA in force	
			No.	Share (%)			No.	Share (%)
1995	22	25	1	4.0	12	12	0	0.0
1996	42	50	3	6.0	19	24	1	4.2
1997	47	47	2	4.3	20	20	1	5.0
1998	43	43	3	7.0	15	15	1	6.7
1999	31	35	4	11.4	17	17	1	5.9
2000	30	30	7	23.3	11	11	3	27.3
2001	27	36	12	33.3	11	20	7	35.0
2002	34	34	7	20.6	23	23	5	21.7
2003	28	28	9	32.1	16	16	4	25.0
2004	20	20	5	25.0	9	9	1	11.1
2005	12	12	6	50.0	5	5	1	20.0
2006	18	18	6	33.3	13	13	4	30.8
2007	15	15	5	33.3	7	7	4	57.1
2008	17	17	4	23.5	10	10	4	40.0
2009 <sup>1</sup>	16	16	2	12.5	n.a.	n.a.	n.a.	n.a.
2010 <sup>1</sup>	17	17	6	37.5	n.a.	n.a.	n.a.	n.a.
<b>TOTAL</b>	<b>419</b>	<b>443</b>	<b>82</b>	<b>18.5</b>	<b>188</b>	<b>202</b>	<b>37</b>	<b>18.3</b>

Note: The numbers for each row were calculated for the year in which the request for consultations was made (i.e. they always refer to the same group of requests for consultations made in that year and not to the number of panels established during a particular year).

<sup>1</sup> The figures relating to the number of panels established for the period 2009-2010 were not included because they are not comparable (i.e. due to ongoing procedures).

Source: WTO Secretariat based on Legal Division's and RTA's databases. See Box E.3.

comes out of Table E.2 is that the share of disputes between PTA partners that advance to the panel stage (45 per cent) is very close to the overall average, indicating that a dispute between PTA partners is just as likely to be settled at the consultations stage as a dispute between non-PTA partners.

Table E.3 compares the number of times a particular WTO agreement has been the subject of a dispute between PTA partners with the number of times it has been invoked in all disputes. There are significant differences with respect to some of the agreements, though it may be difficult to draw conclusions in many cases given the small number of disputes involving certain agreements. The most frequently cited agreements in disputes between PTA partners are the GATT 1994, the Anti-dumping Agreement, the

Subsidies and Countervailing Measures (SCM) Agreement, the Agreement on Safeguards, and the Agreement on Agriculture. Interestingly, subsidy and safeguards disputes make up a larger share of disputes between PTA partners (intra-PTA) than of overall disputes, while intra-PTA disputes involving the GATT 1994 represent a lower share than overall.

Porges (2010) offers some possible explanations for the continued use of WTO dispute settlement by members that are partners in a PTA: the WTO's "familiar institutions" and "unblockable" dispute settlement procedures; the possibility to suspend MFN tariffs and other WTO obligations (particularly where the PTA's margin of preference is low); the broader pool of neutral panellists; the broader issue scope of the WTO; the possibility of forming alliances; access to assistance

Table E.3: WTO Agreements cited in the requests for consultations, 1995-2010

WTO Agreement	No. of references to the Agreements <sup>1</sup>		In requests where a pair of members has a PTA in force		
	Frequency	Share of references (per cent)	Frequency	Share of references in disputes between PTA partners (per cent)	Share of overall references (per cent)
GATT 1994 ( <i>adjusted</i> ) <sup>2</sup>	227	31.0	31	23.7	13.7
Subsidies and Countervailing Measures	86	11.7	16	12.2	18.6
Anti-dumping	84	11.5	27	20.6	32.1
Agriculture	66	9.0	12	9.2	18.2
TBT	41	5.6	7	5.3	17.1
Safeguards	38	5.2	15	11.5	39.5
SPS	37	5.0	6	4.6	16.2
Import Licensing	34	4.6	4	3.1	11.8
TRIPS	29	4.0	1	0.8	3.4
TRIMs	27	3.7	1	0.8	3.7
GATS	22	3.0	3	2.3	13.6
ATC	16	2.2	1	0.8	6.3
Customs Valuation	15	2.0	5	3.8	33.3
Rules of Origin	7	1.0	2	1.5	28.6
Gov. Procurement	4	0.5	0	0.0	0.0
<b>TOTAL</b>	<b>733</b>	<b>100</b>	<b>131</b>	<b>100</b>	<b>17.9</b>

<sup>1</sup> References to the DSU and the Marrakesh Agreement Establishing the WTO were not taken into account.

<sup>2</sup> See Box E.3 for a description of the adjustment methodology used.

Source: WTO Secretariat.

from the Advisory Centre on WTO Law; the multilateral surveillance process; the institutionalized framework for taking countermeasures; and the fact that the cost of WTO dispute settlement is included in a member's annual assessment, while in most PTAs, the parties pay the panellists, or pay for the cost of the tribunal.

(e) Caveats: mechanisms generating negative systemic effects

Some of the deep provisions contained in new-era PTAs can contain discriminatory aspects, creating a tension with the multilateral trading system. The most prominent examples are the area of contingency measures (anti-dumping and safeguards).

(i) *Discriminatory aspects in anti-dumping rules in PTAs*

Recent research suggests that the risk of trade diversion may extend beyond tariffs. Prusa and Teh

(2010) uncover what they call a protection analogue to the trade creation-trade diversion impact of PTAs in the area of anti-dumping. Anti-dumping provisions in PTAs result in members being spared from anti-dumping actions ("protection reduction") while non-PTA members face even greater anti-dumping scrutiny ("protection diversion").

The idea that PTAs may have this distortionary effect is not new. In a series of papers, Bhagwati (1992: 1993) and Bhagwati and Panagariya (1996) conjecture that due to its "elastic" and selective nature, anti-dumping can increase the risk of protection diversion from PTAs. According to their explanation, contingency measures are driven by import volume. Who is targeted in the anti-dumping petition is entirely up to the discretion of the domestic industry.

If anti-dumping provisions make PTA members more difficult to sanction, the domestic industry will simply target other sources. As a result, we might see an increase in anti-dumping protection directed towards

## Box E.3: Methodology

**A Data sources**

The tables and graphs in this section are based on a specialized dataset that was developed based on databases maintained by the Legal Affairs division and the Regional Trade Agreements unit of the WTO. The dataset includes a total of 419 requests for consultations submitted under the WT/DS document series as of 31 December 2010.

**B "Pairs" of members (i.e. complainant-defendant)**

Seven requests for consultations involved more than one complainant (i.e. DS16, DS27, DS35, DS58, DS158, DS217 and DS234), which meant it was not possible to establish whether a PTA was in force between the parties without creating a bias in the figures. For this reason, the 419 requests for consultations as of 31 December 2010 were re-expressed as 443 pairs of complainants-defendants. Figures relating to the prevalence of a PTA at the time of filing the request for consultations were derived on this basis.

**C Adjusting the references to the GATT 1994**

Santana and Jackson (2011) noted that, because complainants tend to cite a large number of agreements and provisions in their requests for consultations under the DSU, frequency counts of provisions cited tend to overestimate the importance of the GATT 1994. This is mainly because references to certain GATT Articles tend to be subsidiary in nature when made together with other "specialized" agreements or even Articles in the GATT. For example, the complainant in a typical anti-dumping case will normally claim that the defendant is in breach of provisions in the Agreement on Anti-dumping, Article VI of the GATT, and that the anti-dumping duty imposed is in violation of the tariff binding (Article II:1(b) of the GATT) and the MFN clause (Article I of the GATT).

In spite of the four Articles cited, the GATT normally plays a secondary role in these disputes. Similarly, a request for consultations citing both Articles II and XIX of the GATT is almost certainly a case about safeguards and not about tariff bindings. To minimize the incidence of those secondary references, and following the principle of *lex specialis*, Santana and Jackson proposed a methodology that does not take into account references to certain Articles of the GATT 1994 when cited together with other provisions. The adjustments are as follows:

1. Article I was excluded when a reference was made in the same dispute to the Agreements on Anti-dumping, Safeguards, SCM (related to countervailing duties - CVD), sanitary or phytosanitary measures (SPS), or technical barriers to trade (TBT), or when a reference was made to Article VI of the GATT (i.e. CVD or anti-dumping related).
2. Article II was excluded when a reference was made in the same dispute to the Agreements on Anti-dumping, Customs Valuations, Safeguards or SCM (CVD related), or retaliation under Article 22 of the DSU. It was also excluded when a reference was made to GATT Articles VI (i.e. CVD or anti-dumping related) or XIX (safeguards).
3. Article III was excluded when a reference was made in the same dispute to either the SPS or the TBT Agreements.
4. Article VI was excluded when a reference was made in the same dispute to Anti-dumping or SCM (CVD related) Agreements.
5. Article XI was excluded when a reference was made in the same dispute to the Safeguards, SPS, TBT Agreements, as well as GATT Articles XII and XIX.
6. Article XVI was excluded when a reference was made in the same dispute to the SCM Agreement (related to the provision of subsidies), or to Articles 3, 6-11 of the Agreement on Agriculture.
7. Article XIX was excluded when a reference was made in the same dispute to the Safeguards Agreement

On the basis of an adjusted dataset, an agreement is considered "cited" if one or more of its provisions are cited in a specific request for consultations.



non-PTA members when in fact the injury to domestic industry mostly stems from imports from other PTA members.<sup>38</sup> The work by Prusa and Teh (2010) provides the first empirical support for this conjecture.<sup>39</sup> Their findings are especially relevant given the prominence of anti-dumping in the trade policy arena. Anti-dumping has long been the contingency measure of choice and its prominence has increased over the past two decades. The number of countries using anti-dumping has increased five-fold and the annual number of anti-dumping initiations has more than doubled (Prusa, 2005).

Figure E.2 shows a discernible difference in the pattern of anti-dumping activity of countries before and after entering into a PTA. Measuring time relative to the year the PTA was enacted, year zero is the year the PTA was established, year  $t - 1$  is the year before while year  $t + 1$  is the year after, etc. Notice that during the years prior to the establishment of the PTA enactment, intra-PTA anti-dumping activity is growing. The number of anti-dumping initiations drop sharply in the year of establishment ( $t = 0$ ) and remain much lower in subsequent years as compared to the years prior to enactment. On average, during the ten years prior to establishment there were 29.5 anti-dumping cases per year and during the ten years following establishment there were just 23.6 cases per year.

There is another way to show how PTA membership changes the pattern of anti-dumping activity. Table E.4 depicts anti-dumping filings when countries are distinguished between (i) those who are members of a PTA and (ii) those who are not, and the time period is distinguished between pre- and post-PTA establishment. As seen, countries file about 58 per cent of anti-dumping cases against non-PTA countries

	Target country	
	Non-PTA country	PTA country
Pre-PTA	506	370
	58%	42%
Post-PTA	3,554	375
	90%	10%

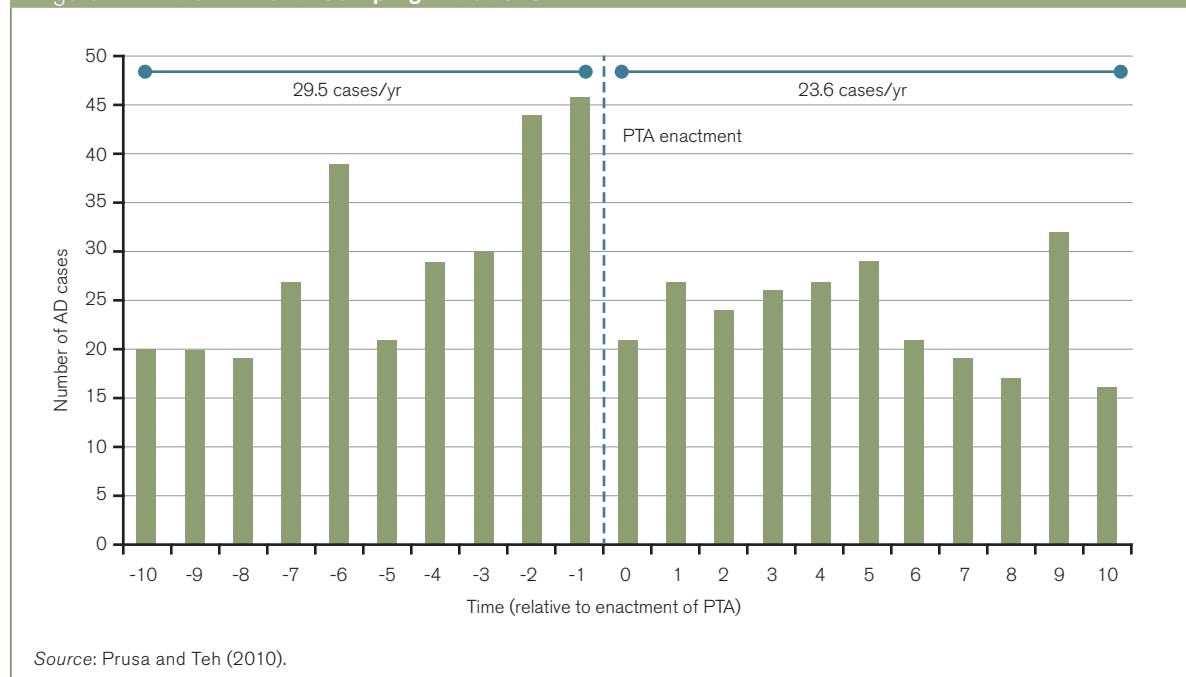
Source: Prusa and Teh (2010).

prior to PTA enactment but a remarkable 90 per cent following enactment. Again, this strongly suggests that PTAs are changing the pattern of protection.

While illustrative, are these patterns statistically significant (unlikely to have occurred by chance)? Furthermore, there may be other provisions in PTAs that can explain the pattern in the anti-dumping data. PTAs often liberalize investment, thus increasing the level of FDI flows between PTA partners. The fall in anti-dumping activity between PTA members might thus arise because imports are sourced from multinational affiliates. Another concern is that the results may be entirely driven by the big users (European Union and the United States) or targets (China) of anti-dumping.

Prusa and Teh's econometric analysis (a method known as difference-in-difference regression) establishes that the patterns do not arise simply from chance.<sup>40</sup> In addition, they find that PTAs cause as much as a 60 per cent reduction in anti-dumping disputes between PTA members. This result is not solely driven by those PTAs that have abolished anti-dumping (for whom intra-PTA anti-dumping activity is

Figure E.2: Intra-PTA anti-dumping initiations



Source: Prusa and Teh (2010).

essentially eliminated). When they only look at those PTAs that have adopted PTA-specific anti-dumping rules, they find a 33-55 per cent reduction in intra-PTA anti-dumping activity. They find no significant change in anti-dumping activity for PTAs without PTA-specific anti-dumping rules.

Their econometric estimates also suggest that PTAs cause a 10-30 per cent increase in the number of anti-dumping filings against non-PTA members. Taking the protection reduction and diversion results together, they find that the reduction in intra-PTA activity is more than offset by the increase in activity against the far larger set of non-PTA members. Overall, they conclude that PTAs increase the number of anti-dumping filings by perhaps as much as 10 per cent.

Their results appear to be extremely stable. Even when they excluded the EU, NAFTA and China individually from their analysis, the results were essentially unaffected. To take account of the possible effects of other PTA provisions, they included FDI flows and a measure of the investment liberalization in each PTA based on work done by Dee et al. (2006) and Dee (2008). While investment provisions in PTAs reduce the incidence of anti-dumping disputes, they continued to find that anti-dumping rules remain a significant independent explanation for the reduction in intra-PTA anti-dumping cases.

### (ii) *Discriminatory aspects in safeguard rules in PTAs*

There are typically two types of safeguard actions which are covered in PTAs: “bilateral” and “global” safeguard actions.<sup>41</sup> Bilateral safeguard actions are meant to apply only to the trade of other PTA members. They provide a temporary escape for members when, as a result of undertaking the commitments under the agreement, increased imports from PTA partners result in serious injury to the domestic industry. Global safeguard actions, on the other hand, are triggered under GATT Article XIX (Emergency Action on Imports of Particular Products) and the Agreement on Safeguards. Multilateral rules require that any safeguard measures be applied on a non-discriminatory basis. Typically, the PTA provisions on global safeguard actions specify the conditions under which PTA partners could be excluded from multilateral safeguard actions invoked by a member.

While most of these PTAs state that their safeguard provisions are in accordance with or do not affect their members' rights and obligations under the multilateral agreements, many go on to exclude the imports of PTA partners from global safeguard actions.<sup>42</sup>

The conditions under which imports from PTA members can be excluded from a global safeguard action are if those imports do not account for a substantial share of total imports and if they do not

contribute to serious injury to the domestic industry or the threat thereof.<sup>43</sup>

The Agreement on Safeguards requires that safeguard measures be applied to all imports irrespective of source (non-discrimination). Thus, the exclusion of PTA partners from a safeguard action poses a potential conflict between regional and multilateral rules. This conflict has been addressed in a number of WTO dispute cases (*Argentina–Footwear*, *United States–Wheat Gluten*, *United States–Line Pipe* and *United States–Steel*). In these cases, the investigating authority had included imports from all sources in making the determination that imports were entering in such increased quantities so as to cause serious injury to the domestic industry. However, instead of applying safeguard measures to all imports irrespective of their source, the country invoking the safeguard action excluded its PTA partners.<sup>44</sup> In all four cases, the Appellate Body has ruled against the WTO member which included its PTA partners in the safeguard investigation but excluded them in the application of the safeguard measure.

The key concept that underlines all these cases has been called “parallelism”.<sup>45</sup> In brief, parallelism prohibits any differences in the application of safeguards measures.<sup>46</sup> In the case of PTAs, parallelism means that when a WTO member has conducted a safeguard investigation considering imports from all sources, it cannot, subsequently, without any further analysis, exclude imports from PTA partners from the application of the resulting safeguard measure. In order to be able to exclude imports from PTA partners, the investigating authority must establish explicitly that imports from non-PTA sources alone caused serious injury or threat of serious injury to the domestic industry. The investigating authority, in its causality analysis, should further ensure that the effects of the excluded (PTA) imports are not attributed to the imports included in the safeguard measure.

While the elaboration of the principle of parallelism by the Appellate Body in these four cases has clarified one issue, WTO jurisprudence has not provided a definitive ruling to what extent GATT Article XXIV could be relied on by a WTO member to exclude PTA partners from the application of a safeguard measure.<sup>47</sup> The provisions excluding PTA partners from global safeguard actions raises concerns about increased discrimination against non-members and trade diversion. Although WTO dispute settlement panels have ruled against excluding PTA partners from safeguard measures if imports from those PTA partners had been included in the investigation, they appeared to have done so on quite narrow grounds – on the lack of parallelism in the application of safeguard measures. So far the Appellate Body has not ruled on whether such exclusions will be justifiable under GATT Article XXIV. Conceivably, under a

different set of circumstances, exclusion of PTA partners from safeguard measures could pass muster.

### (iii) *Other mechanisms*

The non-discriminatory nature of deep provisions might in principle create adverse systemic effects, namely political-economy and third-country resistances to further multilateral liberalization. If preferential liberalization is non-discriminatory in nature, it might be opposed by political-economy forces, because higher market shares (and profits) in the other member's market might be more than offset by the loss of domestic profits vis-à-vis firms from partners and non-members.<sup>48</sup>

Secondly, the non-discriminatory nature of deep provisions may undermine the willingness of developing countries to engage in multilateral negotiations with developed countries with the objective of exchanging deep regulatory commitments with market access for goods (Chauffour and Maur, 2011). This is because preferential tariffs are bound to be eroded over time, whereas regulatory commitments are both permanent and MFN; thereby they cannot be used as bargaining chips over time and vis-à-vis different countries.

Thirdly, it has been argued that lock-in effects of regulatory harmonization within a given PTA may have negative systemic effects (World Trade Organization (WTO), 2007). Competing PTAs with incompatible regulatory structures and standards may lock-in members. This can constitute a threat to the multilateral trading system for two reasons. First, it undermines the principles of transparency and predictability of regulatory regimes. Secondly, it may hinder further multilateral liberalization. A recent study (Piermartini and Budetta, 2009) has found evidence of distinct "families" of PTAs with differentiated rules on technical barriers to trade. The study shows that a number of regional arrangements that have the European Union as the hub include provisions to harmonize the standards of the spoke partner country to EU standards. To the extent that the adjustment to European standards requires making investments, these provisions may lock-in a country to the regional arrangement, thus making movement towards multilateral liberalization costly.

Finally, it has been argued above that third-party MFN clauses have the potential to reduce the discriminatory nature of preferential agreements. However, a variety of PTAs do not contain third-party MFN clauses (e.g. China – ASEAN). In this case, the provisions of the agreement effectively discriminate vis-à-vis third countries, and there is the risk of discriminatory treatment between different parties of different PTAs signed by the same country (Houde et al., 2007). In their services and investment chapters, other PTAs include sectoral exceptions to the automatic extension

of the third-party MFN treatment. Excluded sectors do not therefore automatically benefit from the better treatment of future agreements. However, as reported by Houde et al., very few sectors are concerned.

Moreover, as argued by Adlung and Morrison (2010), a number of agreements exclude some of the potentially most distortive types of intervention from third-party MFN obligations (e.g. all subsidies are excluded under the Australia-United States Free Trade Agreement – AUSFTA). The Economic Partnership Agreements (EPAs) that the EU concluded with African, Caribbean and Pacific (ACP) countries contain MFN clauses requiring that, if an ACP country concludes a subsequent PTA with a major trading economy other than the EU, such as the United States or Brazil, the EU should automatically receive the benefits conceded in such PTA. As argued by Pauwelyn (2009), inclusion of this clause in recent EPAs is controversial. It could in fact have a chilling effect on third countries qualifying as "major trading economies" that were previously interested in concluding a PTA with ACP countries.

## 3. Regionalism and the WTO: historical perspective

The MFN principle is at the core of the multilateral trading system. Nevertheless, from its very beginnings, the multilateral trading system has allowed some space for member countries to grant each other more preferential treatment under free trade areas or customs unions. As one commentator has put it, "(t)he real thrust of the GATT had been to control and contain discrimination rather than eliminate it" (Hudec, 1990). The rules applicable to free trade areas and customs unions under Article XXIV of the GATT have been incorporated into the WTO with little change and the many interpretative questions that arise under that provision remain intensely debated today.<sup>49</sup> Although there are still many observers who would like to see the rules clarified and strengthened, recent efforts have focused on improving transparency.

### (a) The origins of the GATT

Preferential trading arrangements were one of the main issues of concern of some of the countries that participated in the negotiations for the establishment of an International Trade Organization (ITO), which eventually became the basis for the GATT. In particular, some countries saw the ITO negotiations as an opportunity to dismantle certain existing preferential trade arrangements, such as the preferences between territories belonging to the British Commonwealth, while the British seemed willing to dismantle these preferences only if they obtained meaningful access to other markets, particularly the United States (Hudec, 1990). Indeed, several commentators note that this was an important objective for the United States, which made a proposal to

allow preferences only between territories that formed part of a customs union and later accepted interim arrangements that would lead to a customs union. A group of developing countries that included Syria and several Latin American countries sought to widen the exception to include free trade areas.

The language adopted at the Havana Conference of 1947-48, which was later incorporated into the GATT, allowed for free trade areas and customs unions, as well as interim arrangements leading to their formation. Several explanations have been put forward by commentators to explain the eventual acceptance of preferences under free trade areas, especially by the United States, which initially had opposed them.

In a recent historical study, Chase (2006) summarizes the reasons that were traditionally given for the acceptance of free trade areas within the framework of the GATT: the need to compromise to reach agreements (Viner, 1950); discouraging a consolidation of the Commonwealth preferences (Odell and Eichengreen, 1998); encouraging European integration (Bhagwati, 1991; Odell and Eichengreen, 1998); or pressure from certain developing countries (Haight, 1972; Mathis, 2002; World Trade Organization (WTO), 1995). Chase (2006) disagrees with these traditional views and, based on his archival research, suggests that the United States and Canada were secretly negotiating a bilateral free trade agreement and the United States changed its position on free trade areas to accommodate this eventuality. According to Chase (2006), the United States did not have to make a new proposal because it saw an opportunity in the proposal allowing free trade areas submitted by Lebanon and Syria.

Article XXIV of the GATT recognizes “the desirability of increasing freedom of trade by the development, through voluntary agreements, of closer integration”, yet cautions “that the purpose of a customs union or of a free-trade area should be to facilitate trade between the constituent territories and not to raise barriers to the trade of other contracting parties with such territories.” Article XXIV:5 establishes that the provisions of the GATT “shall not prevent, as between the territories of contracting parties, the formation of a customs union or of a free-trade area or the adoption of an interim agreement necessary for the formation of a customs union or of a free-trade area”.

For purposes of Article XXIV, a customs union is understood as “the substitution of a single customs territory for two or more customs territories, so that (i) duties and other restrictive regulations of commerce (except, where necessary, those permitted under Articles XI, XII, XIII, XIV, XV and XX) are eliminated with respect to substantially all the trade between the constituent territories of the union or at least with respect to substantially all the trade in products originating in such territories, and, (ii) ... substantially the same duties

and other regulations of commerce are applied by each of the members of the union to the trade of territories not included in the union”. A free-trade area is “a group of two or more customs territories in which the duties and other restrictive regulations of commerce (except, where necessary, those permitted under Articles XI, XII, XIII, XIV, XV and XX) are eliminated on substantially all the trade between the constituent territories in products originating in such territories”.

Article XXIV sets out additional conditions that must be met by customs unions and free trade areas. Generally speaking, in both cases, the duties and other regulations applied upon formation may not be higher or more restrictive than previously. In the case of customs unions, the duties or regulations may not be “on the whole” higher than the “general incidence” of the duties and regulations of commerce previously applicable in the constituent territories. Interim agreements for the formation of a customs union or free trade area must include “a plan and schedule” for the formation of the customs union or free trade area “within a reasonable length of time”. Certain notification requirements also apply under Article XXIV. Furthermore, Article XXIV includes provisions on frontier traffic (Article XXIV:3) and on observance of GATT obligations by regional and local governments and authorities (Article XXIV:12). Specific exceptions for preferences between certain neighbouring countries (for example, Lebanon and Syria; Belgium-Luxembourg-Netherlands) were included in Article I of the GATT.

## (b) Developments during the GATT years

The creation of the European Economic Community (EEC) and its association agreements were the principal focus of the discussions around Article XXIV during the early years of the GATT. Commentators describe intense debates among the GATT contracting parties on the consistency of the EEC with the requirements of Article XXIV. The compatibility of the Treaty of Rome with the requirements of Article XXIV was not resolved by the contracting parties. As Ladreit de Lacharrière (1987) notes, in 1958, the contracting parties considered it “more fruitful if attention could be directed to specific and practical problems, leaving aside for the time being ... debates about the compatibility of the Rome Treaty” with the GATT.<sup>50</sup>

Eventually the GATT contracting parties opted for resolving some of the tariff issues surrounding the formation of the EEC as part of the Dillon Round (Hoda, 2001). The EEC association agreements with other countries were also the subject of intense debates. Here the concern was about the lack of a clear commitment to full liberalization or membership. EFTA's notification also gave rise to discussions, particularly because of its exclusion of agriculture and fisheries (Hudec, 1990). Another agreement that was notified at the time was ALALC, which included several

Latin American countries, and which raised concerns as to the ambitiousness of the liberalization programme and its objective of promoting infant industries<sup>51</sup> (Hudec, 1990).

At the time, there was no standing body of the GATT that was responsible for reviewing agreements notified under Article XXIV. Instead, these agreements were reviewed by individual working parties. GATT contracting parties did not adopt definitive reports with respect to these agreements. Most commentators agree that, despite the many questions raised by some contracting parties with respect to the PTAs that were notified, what essentially developed was a policy of tolerance towards these agreements. Jackson (1969) observes that generally speaking the practice of the GATT was of "a high degree of tolerance for a wide diversity of regional arrangements". Nevertheless, he recognizes that "legal discussions about criteria in Article XXIV and consultations may have enabled the interests of parties that were not members to regional arrangements to influence those regional arrangements in a way that softened their detrimental impact on the trade of non-members".

Another important development during the GATT was the adoption of the Decision on Differential and More Favourable Treatment, Reciprocity and Fuller Participation of Developing Countries, commonly known as the "Enabling Clause". In addition to providing a basis for unilateral tariff preferences for developing countries, the Enabling Clause provides an exemption from the MFN obligation in Article I of the GATT for "(r)egional or global arrangements entered into amongst less-developed contracting parties for the mutual reduction or elimination of tariffs and, in accordance with criteria or conditions which may be prescribed by the CONTRACTING PARTIES, for the mutual reduction or elimination of non-tariff measures, on products imported from one another".

A total of 124 agreements were notified to the GATT between 1948 and 1994. Of these, however, only 38 remained in force in 1995 when the WTO was established. As explained in a WTO Secretariat Report, this reflects "in most cases the evolution over time of the agreements themselves, as they were superseded by more modern ones between the same signatories (most often going deeper in integration), or by their consolidation into wider groupings" (Crawford and Fiorentino, 2005).

Discriminatory treatment under PTAs became a topic of increasing concern over the years. In 1983, the Director-General of the GATT created an independent group of seven eminent persons to study and report on the problems facing the international trading system. The group issued its report in March 1985. Commonly referred to as the "Leutwiler Report", one of its conclusions is that "(t)he rules permitting customs unions and free-trade areas have been distorted and

abused" and that "(t)o prevent further erosion of the multilateral trading system, they need to be clarified and tightened".

The Report indicated that, while the European Community and EFTA met the conditions in Article XXIV, "many agreements presented under the rules, including some agreements between the European Community and its associates, fall short of the requirements". It further cautioned that "(t)he exceptions and ambiguities which have thus been permitted have seriously weakened the trade rules, and make it very difficult to resolve disputes in which Article XXIV is relevant". Accordingly, the Report proposes that "GATT rules on customs unions and free trade-areas should be examined, redefined so as to avoid ambiguity, and more strictly applied, so that this legal cover is available only to countries that genuinely use it to establish full free trade among themselves" (Leutwiler, 1985).

### (c) PTAs in the Uruguay Round

During the Uruguay Round, a group of countries that included Australia, India, Japan, New Zealand and the Republic of Korea favoured strengthening the disciplines of Article XXIV. Japan, in particular, proposed among others, improving the consultations before and after agreements were reached; establishing a firm time limit on "interim agreements", to ensure that members moved to genuinely open trade; clearly defining "general incidence" of duties or other regulations; and limiting the credit that a new customs union could claim if the general incidence of duties or regulations was actually lower than before. India, for its part, proposed reviewing the requirement that duties and other restrictive regulations be eliminated on "substantially all trade" between the PTA partners (Croome, 1995).

In a second set of proposals, Japan sought to improve the procedures for examination of preferential trade agreements, suggesting the establishment of special procedures, separate from GATT dispute settlement, to assess and discuss compensation for damages caused by preferential agreements to the trade of non-members. Some of those who opposed this proposal suggested that surveillance of preferential trade agreements could be undertaken under the newly-created Trade Policy Review Mechanism (Croome, 1995).

Another issue discussed during the Uruguay Round in connection with preferential trade agreements was the obligation in Article XXIV:12 relating to federal states. This point was initially raised by India, but was later taken up by the European Community, which presented a proposal to tighten Article XXIV:12 by affirming the full responsibility of GATT members for measures taken by their regional or local governments or authorities (Croome, 1995).

Ultimately, the discussion coalesced around the idea of negotiating an Understanding on Interpretation of Article XXIV, which would focus on the calculation of the level of duties before and after a customs union is formed, reassert the obligation to compensate, set out requirements for interim arrangements, limit the “reasonable period of transition” to ten years unless otherwise authorized, and acknowledge that matters arising under Article XXIV could be submitted to dispute settlement.

Despite initial opposition from the European Community (which wanted fuller credit in compensation negotiations for tariff reductions made by group members and was dissatisfied with the text on Article XXIV:12), India (which considered the text disproportionately weak), and Yugoslavia (which objected to the text on Article XXIV:12), the Understanding on Interpretation of Article XXIV was adopted and became part of the Uruguay Round agreements (Croome, 1995).

An additional development of significance during the Uruguay Round was the inclusion in the GATS of a provision on preferential agreements relating to trade in services.<sup>52</sup>

#### (d) Developments in the WTO

##### (i) *Committee on Regional Trade Agreements*

The WTO Committee on Regional Trade Agreements (CRTA) was established by the General Council in 1996 (WT/L/127). It was initially foreseen that the CRTA would carry out the examinations of the regional trade agreements notified to the WTO, thus taking over the functions of the individual working parties of the GATT. Despite the establishment of the CRTA in 1996, the examination of RTAs resulted in stalemate. Between 1996 and 2001 not a single examination report had been adopted by the CRTA, in part due to continuing disagreements over the inherent ambiguities in GATT Article XXIV, the lack of information submitted by RTA parties, and the fact that the determination of consistency was to be made by all WTO members, including those whose RTAs were under examination.

In December 2006, WTO members adopted on a provisional basis a new transparency mechanism for regional trade agreements (WT/L/671).<sup>53</sup> The new mechanism calls on members to provide an “early announcement” of their involvement in negotiations for a regional trade agreement, requires members to promptly notify a newly concluded regional trade agreement, and sets out a schedule for its consideration by WTO members.<sup>54</sup> The mechanism provides that consideration of notified regional trade agreements should conclude within a year from the date of notification. For this purpose, parties to a

regional trade agreement are required to submit certain data to the WTO Secretariat, such as tariff concessions, MFN duties, rules of origin and import statistics.

Based on this data, the text of the agreement, and information from other sources, the WTO Secretariat prepares a factual presentation that is intended to assist members in their consideration of the notified regional trade agreement. WTO members are currently reviewing the transparency mechanism with a view to making it permanent. The transparency mechanism places emphasis on the “consideration” of RTAs rather than on their “examination”, which may be viewed by some as a tacit acknowledgement by members that their interests would be better served by focusing efforts on improving transparency.

WTO members are also engaged in negotiations as part of the Doha Round aimed at “clarifying and improving disciplines and procedures under the existing WTO provisions applying to regional trade agreements.” Negotiations are to “take into account the developmental aspects of regional trade agreements” and have been taking place in the Negotiating Group on Rules.<sup>55</sup>

The CRTA reported that, as of 1 November 2010, 479 regional trade agreements, counting goods and services notifications separately, had been notified to the GATT/WTO, 288 of which were in force at the time.<sup>56</sup> These figures correspond to 375 “physical” agreements, of which 197 were in force (117 goods, 1 services and 79 goods and services). Of the 288 notifications, 174 were notified under GATT Article XXIV, 31 under the Enabling Clause, and 83 under GATS Article V. A total of 92 regional trade agreements had been considered under the Transparency Mechanism since its adoption in December 2006.<sup>57</sup>

##### (ii) *Dispute settlement*

Despite the concerns expressed by many observers regarding the compatibility of many notified regional trade agreements with Article XXIV of the GATT, issues relating to regional trade agreements have not figured prominently in WTO dispute settlement. The most important issue that came up was the question of whether the consistency of a regional trade agreement with Article XXIV could be examined in WTO dispute settlement. In *Turkey–Textiles*, the Appellate Body held that panels have the authority to examine whether a regional trade agreement meets the requirements of Article XXIV. The burden of establishing that the regional agreement meets the requirements of Article XXIV falls on the respondent WTO member to the extent that it invokes the regional agreement as a defence to justify a discriminatory measure.

The availability of WTO dispute settlement to challenge regional trade agreements has given rise to mixed

reactions from commentators. Roessler (2000) has argued that the examination of the consistency of regional trade agreements was a matter that should have been reserved exclusively to the WTO's political organs and specifically to the CRTA. By contrast, Davey (2011) has suggested that WTO dispute settlement could be used to further clarify the disciplines of Article XXIV. WTO members so far have been reluctant to use the WTO dispute settlement system to enforce the obligations of Article XXIV of the GATT and Article V of the GATS.

Issues concerning the relationship between the WTO dispute settlement system and the dispute settlement systems of PTAs have been discussed in connection with a handful of WTO disputes. These disputes were addressed in subsection E.2. In this subsection, we address the small number of disputes in which Article XXIV has been explicitly raised.

As noted above, the case that has dealt most directly with the requirements of Article XXIV is *Turkey – Textiles*. In this case, the Appellate Body examined the requirements applicable to customs unions under subparagraph 5 of Article XXIV and explained that a party invoking this provision to justify an otherwise WTO-inconsistent measure must establish that the following two conditions have been fulfilled. First, it “must demonstrate that the measure at issue is introduced upon the formation of a customs union that fully meets the requirements of sub-paragraphs 8(a) and 5(a) of Article XXIV”. Secondly, it must show that “the formation of that customs union would be prevented if it were not allowed to introduce the measure at issue” (Appellate Body Report, *Turkey – Textiles*, para. 58).

Article XXIV has also been raised in the context of several safeguard cases, where the issue has been whether a WTO member could exclude one of its partners in a preferential trade agreement from the application of a safeguard measure in departure from Article 2.2 of the Agreement on Safeguards. These cases were discussed in subsection E.2.

A measure taken pursuant to a PTA became relevant in a dispute in which Brazil invoked the General Exceptions in Article XX of the GATT to justify an import ban on retreaded and used tyres on public health grounds. As a result of a decision by a MERCOSUR tribunal, however, the import ban was not applied to imports of remoulded tyres from MERCOSUR members.

The panel found that “(t)he exception of remoulded tyres originating in MERCOSUR therefore does not seem to be motivated by capricious or unpredictable reasons” and that “(t)o the extent that the existence of some discrimination in favour of other members of a customs union is an inherent part of its operation, the possibility that such discrimination might arise between members of MERCOSUR and other WTO

Members as a result of the implementation of the MERCOSUR Agreement is not, in our view, *a priori* unreasonable”.

The panel nevertheless noted that “the fact that we give due consideration to the existence of Brazil's commitments under MERCOSUR in our assessment does not imply that the exemption must necessarily be justified. Rather, we must now examine the manner in which the import ban is applied, taking into account the existence of an exemption for MERCOSUR members, in order to determine whether the discrimination arising from the MERCOSUR exemption is arbitrary or unjustifiable”. Because the panel found that the “volumes of imports of retreaded tyres under the exemption appear not to have been significant”, it concluded that “the measure's ability to fulfil its objective does not appear to have been significantly undermined by the occurrence of imports from other sources, even in the presence of an exemption for MERCOSUR imports”.

Therefore, the panel concluded that “the operation of the MERCOSUR exemption has not resulted in the measure being applied in a manner that would constitute arbitrary or unjustifiable discrimination”. The panel also relied on its analysis of the volume of imports to conclude that the MERCOSUR exemption did not result in the import ban being a disguised restriction on international trade (Panel Report, *Brazil–Retreaded Tyres*, paras. 7.272-7.289 and 7.354-7.355).

The Appellate Body disagreed with the panel's finding, explaining that the ruling of the MERCOSUR arbitral tribunal was not an acceptable rationale for the discrimination, because it bore no relationship to the protection of public health, the legitimate objective pursued by the import ban under Article XX(b), and “even [went] against this objective, to however small a degree”. The Appellate Body held “that the MERCOSUR exemption has resulted in the Import Ban being applied in a manner that constitutes arbitrary or unjustifiable discrimination”.

Moreover, the Appellate Body disagreed with the panel's consideration of the volumes of imports. According to the Appellate Body, the analysis of “whether discrimination is 'unjustifiable' will usually involve an analysis that relates primarily to the cause or the rationale of the discrimination”, and does not depend on “the quantitative impact of this discrimination on the achievement of the objective of the measure at issue”. For the same reason, the Appellate Body reversed the panel's finding that the import ban was not applied in a manner that constituted a disguised restriction on international trade (Appellate Body Report, *Brazil–Retreaded Tyres*, paras. 228-229).

A point emphasized by the Appellate Body was that “before the arbitral tribunal established under MERCOSUR, Brazil could have sought to justify the challenged Import Ban on the grounds of human,

animal, and plant health under Article 50(d) of the Treaty of Montevideo”, yet Brazil decided not to do so. The Appellate Body observed that “Article 50(d) of the Treaty of Montevideo, as well as the fact that Brazil might have raised this defence in the MERCOSUR arbitral proceedings, show, in our view, that the discrimination associated with the MERCOSUR exemption does not necessarily result from a conflict between provisions under MERCOSUR and the GATT 1994” (Appellate Body Report, *Brazil–Retreaded Tyres*, para. 234).

#### 4. The relationship between PTAs and the WTO

##### (a) Coherence in international trade governance

The quest for coherence between regionalism and multilateralism is nothing new. In the early days of the multilateral trading system, economic thinking focused on the welfare effects of PTAs. As explained in Section C, the main finding was that these effects were ambiguous for members and generally negative for third parties. As PTAs were mostly about tariff reductions, multilateral market opening which, even if it does not mean completely open trade, reduces discrimination, was seen as superior to preferential opening.<sup>58</sup> In this context, ensuring coherence was understood as accepting that PTAs and the multilateral system could complement each other while imposing disciplines aimed at minimizing the negative effects that PTAs could have.

As mentioned above, in the 1990s, the expansion of regionalism brought the coherence issue back to the forefront. Many analysts re-examined the relationship between the two approaches, this time focusing on the systemic effects of regional integration. They showed that PTAs could either be stepping stones or stumbling blocks on the road to multilateral market opening. This literature, however, did not provide much guidance on how to improve coherence.

Whether they view the multilateral trading system and PTAs as complementing each other or think that the multilateral system is simply superior to the regional approach, observers broadly agree that “the case for finding ways of strengthening the ability of the WTO to influence and discipline PTAs, or at least to blunt their more exclusive and distorting features, remains strong” (Low, 2008).<sup>59</sup> Subsection 3 has shown how since its inception the multilateral system has accommodated preferential trade agreements. GATT/WTO members have largely taken a non-confrontational and non-litigious approach. Approaches to improving coherence have focused on the weaknesses of multilateral disciplines and how they could be fixed. This subsection summarizes the debate and briefly discusses the main proposals. It appears that feasibility is the main issue and political economy is the key.

Recent developments in PTA activity may well change the perspective on coherence. As documented in Section B, PTA activity accelerated noticeably from 1990 onwards. The number of PTAs had more than doubled by 1995 and more than quadrupled by 2010, resulting in close to 300 active PTAs today. As previously discussed, new PTAs – or at least some of them – are qualitatively different from older ones. While part of recent PTA activity has consisted of the consolidation and rationalization of bilateral arrangements, there has also been a trend towards bilateral deals across the world. Since 1995, PTA activity has increasingly crossed regional boundaries. The coverage of PTAs in terms of both policy areas and products has also widened and deepened over time.

This has led some observers to think that regionalism has entered a “new era” where the old analytical framework is no longer valid and where ensuring coherence no longer means merely imposing multilateral disciplines on discrimination. Baldwin (2010), for instance, sees recent PTAs as providing the framework to underpin the “production unbundling” that characterizes a growing share of world trade. In his view, twenty-first century regionalism is more about reducing frictional trade barriers and the cost of doing business and removing domestic entry barriers than about tariff preferences. Given that preferential agreements on such behind-the-border measures do not typically induce trade diversion, their systemic implications cannot be analysed using the traditional stumbling block/stepping stone framework (see Section C).

The political economy of more recent PTAs is also about a lot more than preferential tariffs. First, according to Baldwin (2010), only a few countries can play a leading role in such agreements. PTAs motivated by production sharing, in particular between developed and developing countries, may be seen as an exchange of factories for the relaxation of behind-the-border barriers and assurances to offshoring firms that their investments and intellectual property will be safe. Few countries, in Baldwin's view, have the sort of factories that can be exchanged for deep reform of behind-the-border measures.

Secondly, negotiating behind-the-border reform in the WTO may not help to directly foster inward investment. Thirdly, the nature of behind-the-border policies makes it difficult to multilateralize PTAs. For example, the principle of subsidiarity (see below) may apply in that some areas may best be disciplined at the regional or bilateral level. These considerations lead Baldwin (2010) to the conclusion that “it is, thus, possible and even likely that the new disciplines form an independent system of governance that does not intersect much, or at all, with Marrakesh rules”. If this is the case, the coherence challenge posed by recent trends in regional agreements may be quite different from that arising from discriminatory tariff reductions. It may be that new international trade rules are being negotiated and



decided outside the WTO in a setting where differences in power are greater and in the absence of the basic principles of non-discrimination and reciprocity.<sup>60</sup>

Whether and how this new challenge needs to be addressed is an open question. Further research will be necessary to understand better the systemic effects of deep integration. One issue that may require further investigation is the effects of power asymmetries and options for mitigating them. Also, as already mentioned, the principle of subsidiarity could be used to assess whether measures agreed at the bilateral or regional level need to be submitted to multilateral disciplines.<sup>61 62</sup> This principle states that “action to achieve agreed policy objectives should be taken at the lowest level of government capable of effectively addressing the problem at hand” (Sauvé and Beviglia-Zampetti, 2000). Because countries have different tastes, cultures, endowments, or institutions, their social choices differ. At the same time, efficiency criteria suggest that regulatory regimes should apply to the largest possible communities.

Given this trade-off, the subsidiarity principle states that the determination of regulatory regimes should be as decentralized as possible unless action in one jurisdiction has an impact in others (spillovers) – resulting in cross-border external effects (externalities), or the creation of economies of scale or public goods, in which case they too should be consulted. In other words, “unless there are significant spillovers, there is no efficiency case for imposing one set of standards across different regulatory domains” (Rollo and Winters, 2000).

A basic rationale for international cooperation on regulation is that the cost of complying with different standards may be high. Economies of scale (across countries) and scope (across issues) are likely to exist in rule-making. However, conflicts of interest can arise between countries with permissive regulations and countries with strict regulations that make multilateral coordination hard and perhaps in some instances undesirable. If these factors are sufficiently prevalent, mutual recognition and harmonization of product norms and testing may work better bilaterally and plurilaterally (between relatively similar countries) than multilaterally. While there may be concerns regarding possible negative third-party effects of common or mutually recognized standards and shared conformity assessment in PTAs, empirical evidence suggests that the EU's single market programme increased access at least as much for third-party firms (Mayer and Zignago, 2005).<sup>63</sup>

Finally, the fact that PTAs where preferential tariffs are still important have not disappeared means that both the new and the old coherence challenges need to be tackled at the same time. The evidence presented in Section D suggests that only a (relatively small) number of the new PTAs have little or nothing to do with preferential tariffs, and that tariff preferences still play a role in many new agreements. The next sub-

section provides a short summary of the debate on existing multilateral disciplines. This overview is followed by a discussion of some of the main options for improving coherence.

## (b) Multilateral disciplines on PTAs

As explained in subsection 3, the multilateral system has generated three core provisions to deal with regionalism. The first provision is GATT Article XXIV, which allows departures from MFN for customs unions and FTAs. The Uruguay Round Understanding on the Interpretation of Article XXIV of the GATT seeks to clarify the criteria and procedures for assessing new or enlarged agreements and to improve transparency. The second provision is the “Enabling Clause”, which relaxes (some of) the GATT provisions on PTAs for developing countries in the name of “special and differential treatment” for this group of countries. The third provision is Article V of the GATS, which sets out the rules for PTAs in the services field. As discussed above, WTO members more recently also adopted on a provisional basis a new transparency mechanism for regional trade agreements.

Over the years, a number of concerns regarding the effectiveness of the multilateral oversight of regional agreements have emerged (Davey, 2011; Low, 2008). First, it has been argued that a number of Article XXIV provisions defy uncontested legal interpretation and, more generally, are deficient.<sup>64</sup> The debate has focused on the interpretation of:

- Paragraphs 5(a) and 5(b) of GATT Article XXIV, which state that “the duties and other regulations of commerce” imposed on third parties should not “on the whole be higher or more restrictive than the general incidence” of the pre-PTA duties and regulations;<sup>65</sup>
- Paragraphs 8(a) and 8(b) of GATT Article XXIV, which state that duties and other restrictive regulations of commerce should be eliminated with respect to “substantially all the trade” between the constituent territories, and Paragraph 1(a) of GATS Article V, which states that an RTA should have “substantial sectoral coverage”;
- Paragraph 3 of the Understanding on the Interpretation of Article XXIV of the GATT, which states that the “reasonable length of time” within which the implementation of an RTA should take place should exceed ten years only in exceptional cases.

Secondly, several gaps in the GATT/WTO legal and institutional framework have been identified. The absence of disciplines regarding rules of origin for free trade agreements, in particular, has become an issue with the multiplication of such agreements and the resulting expansion of a spaghetti/noodle bowl. Similarly, there is no indication regarding how agricultural tariff quotas should be treated in

preferential agreements, whether members of such agreements are allowed to exclude their PTA partners from the application of contingency measures applied to the trade of third parties, or whether PTA parties may or may not apply safeguards on their trade with each other. Another question that has been raised is whether the special and differential treatment provisions for developing country PTAs should be extended beyond those in the Enabling Clause.<sup>66</sup>

Thirdly, while the law of the GATT/WTO may have influenced PTA negotiations, in practice, it has never been used to impose discipline on discriminatory reciprocal trade agreements (Davey, 2011; Low, 2008). Governments have almost never agreed through established procedural arrangements whether any given PTA is in conformity with the multilateral rules. Procedural requirements such as notifications have been partially observed at best and dispute settlement findings have not helped address existing weaknesses in the disciplines.

In the eyes of some observers, it is revealing that the Transparency Mechanism for Regional Trade Agreements is the only result of the Doha Round negotiations that has been allowed to go forward independently of the full results of the Round.<sup>67</sup> This suggests both that WTO members are aware of the need to understand better what regional trade agreements are about and that they continue to privilege a cautionary approach (Low, 2008). Others go even further and consider that the Transparency Mechanism advantageously substitutes the “old” review process (Mavroidis, 2010). With trade diversion reduced as a result of multilateral tariff reductions, along with empirical evidence suggesting that PTAs can be welfare improving, and with PTAs covering a number of issues not covered by the WTO, existing rules are considered to be of limited relevance. Mavroidis (2010) argues that the Transparency Mechanism should become the *de jure* new forum to discuss PTAs within the multilateral trading system.<sup>68</sup>

### (c) Possible ways to improve coherence

This report has discussed the idea that there may be a case for maintaining separate regimes for regional and multilateral cooperation. This would be the case where particular types of cooperation are more appropriately managed at the regional rather than the multilateral level. By the same token, there are issues that cannot be addressed adequately at the regional level. In between these two polar realities, the coherence question arises. Essentially, the challenge is to identify where there are gains from ensuring greater coherence among PTAs and between PTAs and the multilateral trading system.

A number of different approaches have been proposed for improving coherence between PTAs and the multilateral trading system (Davey, 2011; Low, 2008;

Sutherland Report, 2004; The Warwick Commission, 2007; World Trade Organization (WTO), 2003). This subsection reviews these proposals and groups them under four headings: i) accelerating multilateral trade opening; ii) fixing the deficiencies in the WTO legal framework; iii) adopting a softer approach as a complement to the existing legal framework; and iv) multilateralizing regionalism. These approaches are not necessarily mutually exclusive. They all aim at reinforcing compatibility and coherence, which essentially means making sure that PTAs contribute to trade cooperation and opening in a fundamentally non-discriminatory manner. They differ mainly in terms of what they see as a politically feasible strategy to reach this objective.

Lowering MFN tariffs would reduce discrimination and thereby blunt the adverse effects of PTAs. The Sutherland Report, for instance, recommended that all developed country tariffs should be bound at zero in WTO members' schedules of commitments at some agreed upon time in the future. While a reduction to zero of all developed country tariffs on industrial products may not seem impossible to achieve in a not too distant future, the Doha Round negotiations suggest that this may not happen without a measure of reciprocity from emerging economies. As for the elimination of all tariffs on agricultural products, this does not seem to be politically feasible in the current context. Also, binding all tariffs at zero may take care of tariff-induced trade diversion but it would not eliminate all potentially adverse effects of deeper integration measures.

As for the idea of filling gaps in the WTO legal framework, the Doha Round includes a mandate to negotiate with a view to “clarifying and improving disciplines and procedures under the existing WTO provisions applying to regional trade agreements”. The negotiations have been pursued along two tracks. On the one hand, members addressed procedural issues relating to the transparency of PTAs. On the other hand, they tried to identify issues for negotiation, including “substantive” issues, such as systemic and legal issues.<sup>69</sup> As already mentioned, negotiations on the procedural issues resulted in the adoption on a provisional basis of a new transparency mechanism for regional trade agreements (WT/L/671). The negotiations on the “substantive” issues have so far generated proposals by various members mainly aimed at clarifying the provisions of GATT Article XXIV. While these proposals contribute usefully to the debate, they do not seem to have converged towards any form of consensus on possible reforms to the rules.<sup>70</sup>

This should not come as a complete surprise as previous discussions have not led to much progress on substantive issues.<sup>71</sup> One possible explanation for the lack of progress is that members who have entered PTAs in the past may be reluctant to sign off on clarifications in the rules that might suggest that the

PTAs they belong to did not comply with Article XXIV (Davey, 2011). Considering that efforts to clarify concepts such as “substantially all trade”, “other restrictive regulations of commerce”, etc. have had limited success so far, it seems unlikely that the second option referred to above – that of clarifying and strengthening existing rules – would be viable.

Moreover, WTO members have been reluctant to use the WTO dispute settlement system in order to clarify existing rules and it does not seem likely that they will change this posture in the near future. This does not mean that revised and improved rules will not one day be part of any significant progress towards more coherence, only that this does not seem to be a promising starting point. In that context, economic analysis could help strengthen the existing provisions. It shows, for example, that the condition in GATT Article XXIV that the protection applicable to non-members should not increase with the creation or extension of a PTA will not necessarily protect the latter from a welfare loss.<sup>72</sup>

The third option noted above would be to adopt a “soft law” approach to complement the “hard law” and the dispute settlement mechanism. There is no agreement in the literature regarding the definition of the concept of “soft law”, although legal scholars often seem to define hard law as binding and soft law as non-binding (Shaffer and Pollack, 2010). One example of soft law would be the Code of Good Practice for the Preparation, Adoption and Application of Standards annexed to the WTO Agreement on Technical Barriers to Trade. Following the Code is optional for WTO members and WTO dispute settlement is unavailable as a remedy under the Code. Another example would be APEC’s Best Practices for Free Trade Agreements and Regional Trading Agreements.<sup>73</sup> The rationale for using a soft law approach would be to allow WTO members to better understand their respective priorities and interests, with a view eventually to unblocking progress towards legal interpretations of particular provisions that would ensure coherence.

The soft law approach is not without risk. As pointed out by Shaffer and Pollack (2010), soft law and hard law could become antagonistic to one another if the underlying conditions for cooperation are absent. Low (2008) argues that a shared perception of objectives and the nature of the transition to hard law would increase the chances that soft law could help rebuild hard law. In view of these considerations, he proposes a three-stage approach. The first stage would involve increased transparency and information sharing under the new Transparency Mechanism. This reinforced exchange of views would pave the way for the progressive development of soft law in the form of a code of good practices in the second stage. Finally, in a third and last stage, when governments become comfortable with the soft law, negotiations aimed at improving the hard law provisions could be undertaken.

The fourth and last proposal is to multilateralize regionalism (Baldwin, 2006; Baldwin and Thornton, 2008). Baldwin (2009) defines a process of multilateralization as the extension of existing preferential arrangements in a non-discriminatory manner to additional parties, or a fusion of distinct PTAs. The idea is that, as a result of global production sharing, political economy forces that were behind the proliferation of PTAs and the creation of the so-called spaghetti bowl have weakened and are being progressively replaced by new forces favourable to the multilateralization of preferences. This translates into a number of multilateralization initiatives both at the regional and at the multilateral level.

Examples of initiatives taken at the regional level to reduce the tangle of PTAs include APEC’s Best Practices for PTAs or the Pan European Cumulation System, which reduced the distortions of international economic production within the zone through the harmonization of rules of origin and diagonal cumulation. An interesting example of multilateralization at the multilateral level is the Information Technology Agreement, which established a mechanism for the elimination of MFN tariffs on information technology products and thus made rules of origin and rules of cumulation non-operative.

Recent research has highlighted the potential cost of overlapping PTAs and complicated rules of origin to today’s world of geographically fragmented production chains (Baldwin et al., 2009). There may be a role for the WTO to reduce these transaction costs by serving as a forum for the coordination/standardization/harmonization of preferential rules of origin.<sup>74</sup> Another way that greater coherence can be established has already been discussed and consists of identifying “best practices” in PTAs.<sup>75</sup> As noted in Section D, the extent to which deep integration measures in PTAs have the potential to generate the same sort of costly spaghetti/noodle bowl as tariff preferences is still being debated. Baldwin et al. (2009) explore six different areas, discussing for each of them whether PTAs have created a spaghetti bowl and how PTA provisions have been or could be multilateralized.

A final thought with respect to moves towards the multilateralization of PTAs concerns decision-making procedures. Several authors (Lawrence, 2006; VanGrasstek and Sauv e, 2006; Cottier, 2009; Elsig, 2009; Low, 2011) have considered the possibility of developing a multilateral approach to a modified consensus rule, often referred to as critical mass decision-making. The approach proposed by Low (2011) is very similar to the so-called “code” approach that emerged in the Tokyo Round agreements on non-tariff measures, but which was subsequently eliminated by the “Single Undertaking” (whereby nothing is agreed until everything is agreed) that accompanied the creation of the WTO in 1995. A revival of the critical mass approach occurred with the

post-Uruguay Round agreements on basic telecommunications and financial services, as well as the Information Technology Agreement.

The adoption of a critical mass approach would make it possible to multilateralize trade rules without implicating the entire WTO membership – a proposition that may look attractive where there is a case for more broadly shared regulatory approaches to trade but not necessarily on a global basis. A critical mass may be said to exist when a sufficiently large subset of the entire membership agrees to cooperate under the auspices of the WTO. An important characteristic of the approach is that agreements do not involve any discrimination vis-à-vis non-signatory countries.

Appropriately chosen institutional and procedural safeguards could protect the system against the risk of fragmentation and dilution of the multilateral basis for trade cooperation. Regarding the definition of critical mass, for example, a simple but effective approach could be to let the critical mass define itself. Critical mass would be reached when those prepared

to go ahead with an agreement consider that support and commitment for the agreement in the membership is sufficient. Those left outside would then be considered too small to undermine the agreement and there would not be any reason for refusing to apply the MFN rule in respect of all the benefits to all non-signatories.

Another important question is whether and when consensus decision-making would need to be applied to critical mass initiatives. In the absence of multilateral participation through a consensus-based process, a risk exists that a sub-set of the membership could shape rules from which they benefitted, but at the expense of members that were not part of the critical mass. The suggestion here is that critical mass agreements would need to be approved by consensus before they enter into force. Not only would the risk of damaging the interests of non-members of the critical mass be guarded against, but critical mass agreements would also remain within the ambit of the multilateral system.

## Endnotes

- 1 "Systemic effects" are defined for the purpose of this report as the static and dynamic effects of PTAs on the multilateral trading system. An example of static effect is the possibility of conflicting rules, for instance on trade remedies. An example of a dynamic effect is the impact of a PTA on the probability of engaging in further multilateral negotiations.
- 2 There is some theoretical and empirical work studying the inverse question of whether multilateralism drives the proliferation of PTAs. Ethier (1998) and Freund (2000) build theoretical models where PTA formation is an endogenous response to the multilateral trading system. Using data on multilateral tariff cuts and duty-free access concessions granted by the United States at the tariff-line level, Fugazza and Robert-Nicoud (2010) find empirical evidence in support of the claim that past MFN opening sows the seeds of future preferential opening.
- 3 There are practical problems with this argument. First, assuming the availability of international lump-sum transfers may not be realistic, and in their absence, it may very well be that, at some point, some bloc members will veto further enlargements. Secondly, nothing forces PTA members to set their external tariffs as assumed by Kemp and Wan and they may indeed have reasons to set them differently (see Section C.1).
- 4 "Preference erosion" refers to declines in the preference margin that some exporters enjoy in foreign markets as a result of preferential trade treatment. It can occur when export partners eliminate preferences, expand the number of preference beneficiaries, or lower their MFN tariff without lowering preferential tariffs proportionately (Alexandraki and Lankes, 2004).
- 5 Excluded countries suffer from the PTA because the border price faced by their exporters falls. From the perspective of member countries, the gains of moving to global free trade are better access to third-country markets and more liberalization in their import markets. However, these gains are small for low initial tariffs, giving no incentive to PTA member countries to move to multilateral tariff reductions.
- 6 However, Amiti and Romalis (2007) argue that for many developing countries, actual preferential access is less generous than it appears because of low product coverage or complex rules of origin. Therefore, lowering tariffs at the multilateral level (Doha Round), especially on agricultural goods, is likely to lead to a net increase in market access for many developing countries.
- 7 This is the so-called "juggernaut" logic (Baldwin and Robert-Nicoud, 2008).
- 8 Note that the effect could be reversed if the PTA resulted in a higher level of protection for the home import competing sector. In this case, as argued below, the PTA would inhibit multilateralism.
- 9 Enhanced protection is obtained when producers from the low-(external) tariff member can export all their output to the high-tariff member without affecting prices there. In that case, producers in the high-tariff country are not hurt while producers from the low-tariff country enjoy higher protection rents (Freund and Ornelas, 2010).
- 10 As discussed in Section C, Ornelas (2005b), (2005a) qualifies the argument in models where the external tariff is endogenous. The possibility that trade-diverting PTAs are formed is more limited, but cannot be ruled out.
- 11 Schiff and Winters (1998) argue, however, that PTAs based on such factors are likely to be transitory, since optimum trade preferences tend to decline over time. In their model, the PTA's external trade policy becomes increasingly open over time.
- 12 Notice that this result is independent of the existence of political economy motivations in excluded countries. If, however, the governments of non-member countries put a disproportionately high value on the profits of producers, they are even more likely to oppose global trade opening.
- 13 Since it is not possible to observe the degree of multilateral liberalization to which a country that is a member of a PTA would have committed to in its absence, these empirical studies have to rely on differences in liberalization patterns over time, across countries or across sectors, making it harder to identify the causal effect of PTAs.
- 14 Unilateral tariff reductions have accounted for two-thirds of the 21 percentage point cuts in average weighted tariffs of all developing countries between 1983 and 2003, according to the World Bank (2005). Tariff reductions associated with the multilateral commitments in the Uruguay Round accounted for about 25 per cent, and the proliferation of regional agreements amounted to about 10 per cent of the reduction.
- 15 Both studies find that Uruguay Round liberalization was smaller in products where preferences were granted.
- 16 This interpretation is strongly criticized by a number of scholars (Baldwin, 2009). According to Baldwin (2009), it is Canada and Mexico's change of mind that triggered the rise of regionalism in North America.
- 17 This and the following paragraph draw on World Trade Organization (WTO) (2007).
- 18 As explained in more detail below, the PECS arrangements came into being because industrial trade was almost duty-free in Europe, but trade flows were beset by complex and intertwining origin and cumulation rules. Trade in information technology products was virtually duty free, but the impediments to efficiency arising from multiple preferential arrangements built pressure on governments to simplify arrangements – hence the ITA.
- 19 The point is more general than service liberalization. It applies, for instance, to policies that reduce or eliminate technical barriers to trade (TBTs) across the board, by way of regulatory harmonization or mutual recognition. Empirical evidence suggests that the EU's single market programme (a large part of which is based on non-discriminatory regulation) increased access at least as much for third-party firms as for EU members (Mayer and Zignago, 2005).
- 20 First-mover advantage defines cases in which the supplier that first gets into the market can benefit from a long-lasting advantage, even if other suppliers are not subsequently prohibited from entering. See Mattoo and Fink (2004) and Manger (2008).
- 21 GATS Article V:6 mandates the establishment of liberal RoOs for PTAs involving developed countries. The Article establishes that "A service supplier of any other Member that is a juridical person constituted under the laws of a party [...] shall be entitled to treatment granted under such agreement, provided that it engages in substantive business operations in the territory of the parties to such agreement". GATS Article V:3(b) provides that PTAs involving only developing countries may "limit trade preferences to service suppliers owned or controlled by persons of the parties". Yet most PTAs among developing countries have not taken advantage of this option. Among the

- reasons why countries have agreed to include liberal RoOs in the GATS and not to use the special and differential treatment provision specified above, Fink and Jansen (2009) mention: i) the fact that established non-party service suppliers are seen as part of the domestic economy; ii) in the presence of network economies, it is more efficient for services providers to simultaneously serve several markets, which is made easier by flexible rules of origin; iii) participation in global production sharing creates an incentive to abandon idiosyncratic service standards as a way of boosting the competitiveness of own exporters and improving the attractiveness of nations to FDI.
- 22 For instance, the Closer Economic Partnership Arrangements (CEPA) between China and Hong Kong, China and Macao, China, respectively, follow the wording of GATS Article V:6 very closely. However, Emch (2006) argues that the necessity to accumulatively comply with six requirements (nature and scope of business; years of operations; payment of taxes; business premises; employment of staff; exclusion of intra-group services) to qualify for the "substantial business operations" requirement may *de facto* grant access only to a few service suppliers, on a selective basis.
  - 23 It should be noted that GATS Article V:6 only recognizes the interests of juridical, but not of natural persons of third countries who supply services under mode 4 in the territory of one of the PTA members. For instance, a Japanese national with a degree from a French university and a licence to practice in France who wants to work in Germany would not be entitled to the treatment granted to EU nationals.
  - 24 According to UNCTAD (2009), 2,676 BITs were in place at the end of 2008. Eighty-two BITs were signed in 2009, and six during the first five months of 2010 (United Nations Conference on Trade and Development, 2010).
  - 25 In the context of investment, MFN requires that all investors from PTA-member countries are accorded the best treatment accorded to any other foreign investor. NT requires that investors from PTA-member countries are treated as well as domestic investors.
  - 26 NAFTA-based agreements accord the better of MFN and NT. See Kotschwar (2009) and the discussion of investment provisions in Section D.
  - 27 The bilateral agreements that flourished in Europe from the mid-nineteenth century until World War I included such unconditional non-discrimination clauses. The end result was *de facto* multilateral non-discriminatory liberalization (Lampe, 2009).
  - 28 There are, however, a number of caveats that limit the role of such MFN clauses as automatic multilateralizers of preferential treatment. These caveats are discussed in Section E.2(e) below.
  - 29 See Baldwin et al. (2009) for details.
  - 30 The trade effects of PECS are discussed in Box C.4 of Section C. For a discussion of the effects of the "multilateralization" of rules of origin on the multilateral trading system, see Box E.2.
  - 31 A radical solution would be the elimination of MFN tariffs on industrial goods, which would render rules of origin unnecessary. This is obviously politically unpalatable.
  - 32 Article 23.2 of the DSU "prohibits certain unilateral action by a WTO member". More specifically, under Article 23.2, a WTO member "cannot unilaterally: (i) determine that a violation has occurred, benefits have been nullified or impaired, or that the attainment of any objective of the covered agreements has been impeded; (ii) determine the duration of the reasonable period of time for implementation; or (iii) decide to suspend concessions and determine the level thereof". (Appellate Body Report, *US / Canada – Continued Suspension*, para. 371).
  - 33 See the GATT ruling in *United States – Margins of preference*, BISD II/11.
  - 34 For a detailed discussion of jurisdiction of international adjudicative bodies and of these doctrines, see Shany (2005).
  - 35 This can happen, for example, where the complainant in one forum is a government, while the complainant in the other forum is a private party.
  - 36 For a contrary view, see Kuijper (2010).
  - 37 It should be clarified that the existence of conflicting decisions was not the basis for the reversal of the WTO panel by the Appellate Body.
  - 38 Notice that the welfare effects of this increased discrimination are, however, unclear, because there is potentially both trade creation within the PTA and trade diversion away from cheaper sources of imports from non-members.
  - 39 Teh et al. (2009) and Prusa and Teh (2010) map the anti-dumping provisions of about 80 PTAs, covering almost 50 per cent of worldwide exports. Because anti-dumping use is governed by the WTO Anti-dumping Agreement, they expect that if PTA rules have any impact, they will serve to make AD duties more difficult to impose on PTA members. This can take a number of forms. Some PTAs increase the threshold required to apply anti-dumping duties, or in the event that a duty is applied, either reduces it below the dumping margin or shortens the applicable duration. Other PTAs give a role to regional bodies to conduct investigations and/or review the final determinations of national authorities.
  - 40 To explain the method, imagine observing anti-dumping activity against two groups of countries (PTA members and non-PTA members) for two time periods (pre- and post-PTA establishment). The PTA countries are "treated" to some additional anti-dumping rules that possibly affect activity in the post-PTA period but not in the pre-PTA period. The non-PTA countries are not exposed to the treatment during either period. Thus, any observed difference in anti-dumping activity between the two groups of countries can be causally attributed to the treatment – the anti-dumping rules.
  - 41 The discussion in this subsection closely follows Prusa and Teh (2010).
  - 42 PTAs which exclude PTA partners from global actions include Australia-Thailand, Australia-US, Canada-Chile, Canada-Israel, EU-Chile, Group of Three, Mexico-Chile, Mexico-Israel, Mexico-Nicaragua, Mexico-Northern Triangle, Mexico-Uruguay, NAFTA, US-CAFTA-DR, US-Jordan and US-Singapore.
  - 43 Most of the PTAs describe very precisely what "substantial share" of total imports and "contribute importantly to serious injury" mean. In some PTAs, "not substantial share of total imports" means if the partner is not among the top five suppliers during the most recent three-year period. The phrase "not contribute importantly to serious injury or threat thereof" means that the growth rate of imports from the PTA partner is appreciably lower than the growth rate of total imports from all sources.
  - 44 In *Argentina–Footwear*, Argentina included MERCOSUR imports in the analysis of factors contributing to injury to its domestic industry. But it excluded MERCOSUR countries from

- the application of the safeguard measure. In *United States–Wheat Gluten*, the United States excluded Canada from the application of its safeguard action although imports of wheat gluten from Canada were included in the investigation phase. In the *United States–Line Pipe* case, the United States excluded imports from its NAFTA partners from the safeguard measure while including them in the analysis of factors contributing to injury. And in *United States–Steel*, the United States included all sources of imports in its analysis of increasing imports, serious injury and the causal nexus. However, it excluded its NAFTA partners, Israel and Jordan from the application of its safeguard action.
- 45 While the word parallelism is not found in the text of the Agreement on Safeguards, the Appellate Body considered that the requirement of parallelism is found in the language used in the first and second paragraphs of Article 2 of the Agreement on Safeguards. See Appellate Body Report, *US–Steel*, para. 439.
  - 46 See Pauwelyn (2004) for a critique of the Appellate Body's use of this principle.
  - 47 One dispute (between the United States and the Republic of Korea) in which this issue was given some consideration was the *United States–Line Pipe* case. There the United States argued that GATT Article XXIV gave it the right to exclude its NAFTA partners from the scope of the safeguard measure. The panel accepted the US argument that the exclusion of its PTA partners from safeguard actions forms part of the required elimination of "restrictive regulations of commerce" on "substantially all the trade" among the free trade area members, which is a condition required by GATT Article XXIV. The panel decision was subsequently appealed by the Republic of Korea. On appeal, the Appellate Body declared the ruling by the panel on Article XXIV as moot and having no legal effect. The question whether Article XXIV of the GATT 1994 permits imports originating from a PTA partner to be exempted from a safeguard measure becomes relevant only in two circumstances. The first was when the imports from PTA members were not included in the safeguard investigation. The second was when imports from PTA members were included in the safeguard investigation it nevertheless was established explicitly that imports from sources outside the free-trade area, alone, satisfied the conditions for the application of a safeguard measure. Since neither of these applied to the circumstances surrounding the *United States–Line Pipe* case, the issue was not relevant to the case. The Appellate Body was careful to point out though that, in taking this decision, it was not ruling on the question whether Article XXIV of the GATT 1994 permits exempting imports originating in a member of a free-trade area from a safeguard measure. This decision thus leaves the question of an appeal to GATT Article XXIV still very much open.
  - 48 However, Baldwin et al. (2009) argue that production unbundling is likely to soften political opposition to non-discriminatory deep provisions. See Section E.2 (e).
  - 49 Two minor amendments were made to Article XXIV of the GATT in 1955-1957. The term "constituent territories" was replaced with "parties", and the term "included" was replaced with "provided for" (Jackson, 1969).
  - 50 Certain measures that were linked to the formation of the European Economic Community or its expansion were challenged in GATT dispute settlement. (See, for example, US Action Under Article XXIII (Chicken War) and EEC Citrus Preferences (and Association Agreements)). At the same time, as Hudec (1990) notes, the formation of the European Economic Community meant that disputes between EEC members were no longer brought to WTO dispute settlement. He further observed that for some time the EEC was reluctant to initiate disputes against other contracting parties fearing that it would invite challenges to EEC measures.
  - 51 Hudec (1971) suggests that Article XXIV may not have been "drafted with the developing countries in mind". He explains that while the GATT recognizes the right to raise trade barriers for the purposes of industrial development - that is, to promote infant industries - the requirements of Article XXIV may limit this possibility, as they call for elimination of internal barriers and a *status quo ante* ceiling on external barriers.
  - 52 For a history of this provision, see *Systemic Issues related to 'Substantially all the Trade'*, Background Note by the Secretariat (Revision), WT/REG/W/21/Rev.1, 5 February 1998. By contrast, a provision on preferential trade agreements was not included in the TRIPS Agreement.
  - 53 On 14 December 2010, the General Council adopted a Decision on a Transparency Mechanism for Preferential Trade Arrangements (WT/L/806), which was drafted as a result of the mandate given by the General Council to the Committee on Trade and Development in 2006. This mechanism covers: preferential trade agreements falling under paragraph 2 of the Enabling Clause, with the exception of regional trade agreements under paragraph 2(c); preferential trade agreements taking the form of preferential treatment accorded by any member to products of least-developed countries; and any other non-reciprocal preferential treatment authorized under the WTO Agreement. Paragraph 2(c) of the Enabling Clause refers to "Regional or global arrangements entered into amongst less-developed contracting parties for the mutual reduction or elimination of tariffs and, in accordance with criteria or conditions which may be prescribed by the CONTRACTING PARTIES, for the mutual reduction or elimination of non-tariff measures, on products imported from one another".
  - 54 Agreements notified under GATT Article XXIV and GATS Article V are considered by the CRTA. Agreements notified under the Enabling Clause are considered in the Committee on Trade and Development (CTD).
  - 55 At the request of the Negotiating Group on Rules, the WTO Secretariat has prepared a compendium of issues related to PTAs that have been generated by work within the CRTA and discussions in other WTO bodies up to 2002 (see *Compendium of Issues related to Regional Trade Agreements*, Background Note by the Secretariat, TN/RL/W/8/Rev.1, 1 August 2002).
  - 56 These figures correspond to notifications of new regional trade agreements, as well as accessions to existing ones.
  - 57 Eighty-eight regional trade agreements were considered in the CRTA and four in the Committee on Trade and Development.
  - 58 Multilateralism is also considered superior to regionalism because large countries can behave in a more hegemonic way when they negotiate bilaterally with smaller countries.
  - 59 See also Davey (2011).
  - 60 A similar point is made by Brown and Stern (2011).
  - 61 The traditional theory of trade agreements focuses its attention on terms-of-trade effects. In terms-of-trade theory, the motivation for entering into trade agreements depends on whether a country can influence the price of its imports through its trade policy. If two large countries enter into a trade agreement to escape a prisoners' dilemma, this agreement should be multilateral rather than preferential. This is because if they do not extend the benefit of their

- bilateral agreement to any third country through some form of MFN treatment, one or the other of the two large countries could indulge in “bilateral opportunism” by making an agreement with a third party which excluded the other large country partner (World Trade Organization (WTO), 2007).
- 62 Section C presents the Oates decentralization theorem, which provides the economic rationale for the subsidiarity principle.
- 63 See the discussion of TBT commitments in PTAs in Baldwin et al. (2009).
- 64 See Davey (2011), the overview of the debate in the WTO's *World Trade Report* (2007) and Marceau and Reiman (2001).
- 65 Both the definition of the “other regulations of commerce” and the question of how the requirement that RTAs should not result in higher barriers against third parties were intensely debated.
- 66 Procedural issues relating to the administration of the PTA provisions of the Enabling Clause have been addressed through the Transparency Mechanism for Regional Trade Agreements.
- 67 Note that in December 2010 the WTO General Council adopted a Transparency Mechanism for Preferential Trade Agreements (WTO document WT/L/806), which extends the Transparency Mechanism for RTAs to non-reciprocal preferences.
- 68 Evenett (2009) emphasizes that the WTO General Council Decision establishing the provisional Transparency Mechanism (WT/L/671) mentions “consideration” rather than “examining” or an “evaluation” of RTAs, which, in his view, suggests that the collective WTO membership does not want this new mechanism to have “teeth”.
- 69 Note that some issues, such as for instance those pertaining to the internal coherence of WTO provisions that apply to PTAs, have both a procedural and a substantive or legal dimension.
- 70 See Davey (2011). While there does not appear to have been much consideration of these issues in recent years, there is now a new proposal on the table and discussions have restarted. It remains to be seen whether they will be substantive.
- 71 See the summary of discussions prepared by the WTO Secretariat (TN/RL/W/8/Rev.1).
- 72 For a more detailed economic discussion of the proposals, see *World Trade Report 2007* (World Trade Organization (WTO), 2007).
- 73 See Marceau (2007).
- 74 On the multilateralization of rules of origin, see also Box E.2.
- 75 A “best practice” has alternatively been defined as a rule that allows convergence to some multilateral benchmark. See Plummer (2006) for a possible approach.