The Regionalism Debate:
An Overview

Arvind Panagariya

1. INTRODUCTION

This paper offers an overview of the debate on the merits of regional arrangements as an instrument of trade liberalisation. While the advocates of these arrangements view them as moving the member countries and the world towards freer trade, multilateralists argue that they detract from true liberalisation and fragment the global trading system. Because I count myself among multilateralists, the paper has a strong bias in favour of the latter view.¹

In this introductory section, I spell out some key definitions, describe the WTO provisions permitting regional arrangements, and recount the main historical events which led to the current wave of regionalism. This is followed by a review of the economic effects of PTAs on union members (Section 2) and the global trading system (Section 3).² Two terms, ‘Open Regionalism’ and ‘Deep Integration’ have received much attention in recent years from the advocates of PTAs. Given the importance these terms have acquired in policy discussions, it is appropriate to subject them to a close examination. Thus, open regionalism is dissected in Section 4 and deep integration in Section 5. I conclude the paper in Section 6 with suggestions for reforming the WTO rules on PTAs.

ARVIND PANAGARIYA is Professor of Economics at the University of Maryland, College Park, USA. He is grateful to Jagdish Bhagwati for collaborations and conversations over the years on the subject matter of this paper. Comments from Douglas Irwin on an earlier draft are gratefully acknowledged.

¹ A more neutral overview can be found in Winters (1996).
² The two sets of effects are not entirely separable so that the division should not be viewed as water tight. An alternative classification, employed in Bhagwati (1993), Bhagwati and Panagariya (1996a and 1996b) and Bhagwati, Greenaway and Panagariya (1998), divides the implications of PTAs into static welfare effects and the ‘dynamic’ time-path issue. The key feature distinguishing the classification in the present paper is that it places all systemic effects, including those relating to the time-path question, into a single category.
a. Definitions

Three key concepts, all relating to arrangements confined to trade in goods, appear frequently in the academic as well as policy literature: Preferential Trade Area, Free Trade Area and Customs Union. A Preferential Trade Area or PTA is a union between two or more countries in which goods produced within the union are subject to lower trade barriers than the goods produced outside the union. A Free Trade Area or FTA is a PTA in which member countries do not impose any trade barriers on goods produced within the union but do so on those produced outside the union. A Customs Union or CU is an FTA in which member countries apply a common external tariff (CET) on a good imported from outside countries. The CET can, of course, differ across goods but not across union partners.

In policy documents and debates, the acronym FTA is often used to refer to a Free Trade Agreement or Free Trade Arrangement rather than Free Trade Area. This usage of the term gives the misleading impression that an FTA is equivalent to non-discriminatory free trade. The term PTA, whether used to stand for Preferential Trade Area, Preferential Trade Agreement or Preferential Trade Arrangement, avoids this confusion by making explicit the discriminatory nature of the arrangement. For this reason, at the urging of Bhagwati (1995), economists have increasingly adopted the term PTA to refer to Free Trade Areas. The term has the additional advantage of being wider in that it can be used to describe FTAs, CUs and arrangements involving partial trade preferences.

In practice, PTAs rarely do away with all trade barriers among member countries. For instance, in the North American Free Trade Agreement (NAFTA), anti-dumping measures can be used by member countries against one another. Similarly, in the European Union (EU), competition policy can be invoked to restrict the flow of imports from partner countries. In addition, PTAs may exclude entirely certain goods or sectors from liberalisation. A prime example of such an exclusion is agriculture in the EU.

PTAs are sometimes accompanied by agreements in areas other than trade in goods. For example, the EU has gone some way towards introducing harmonisation of product standards, competition policies and tax laws and is poised for a complete monetary union among at least a subset of member countries. NAFTA has provisions relating to investment liberalisation, intellectual property rights, dispute settlement and, through side agreements, environmental and labour standards.

The term 'Regional Arrangement' (RA) is sometimes used either as substitutes for PTA or to describe an arrangement which is not a PTA. I will use this term

---

3 Two other terms, 'Regional Trade Arrangement' and 'Regional Integration Arrangements' also appear in the literature sometimes. But these are poor substitutes for either PTA or RA and I will not use them in this paper. Qualifier 'regional' in terms such as RA, RTA and RIA should be interpreted to mean that the membership falls short of all countries on the globe without any implication that it is concentrated in a specific geographical region of the world.
only for arrangements that cannot be justifiably called PTAs. For instance, the Asia Pacific Economic Cooperation (APEC) forum is not a PTA and is, as such, better described as a regional arrangement. As mentioned in the previous paragraph, arrangements such as the EU and NAFTA go beyond a PTA but I will continue to call them PTAs rather than risk creating confusion by introducing a new term distinguishing them from pure PTAs.

b. WTO Provisions for PTAs

The Most Favoured Nation (MFN) clause in Article I of the General Agreement on Tariffs and Trade (GATT) forbids Member countries from pursuing discriminatory trade policies against one another. Indeed, the language in the 1947 Agreement, subsumed within the 1994 Marrakesh Agreement establishing the World Trade Organisation (WTO) is quite unequivocal:

With respect to customs duties and charges of any kind ... any advantage, favour, privilege or immunity granted by any contracting party to any product originating in or destined for any other country shall be accorded immediately and unconditionally to the like product originating in or destined for the territories of all other contracting parties (GATT, 1994, p. 486).

PTAs have, nevertheless, been accommodated into GATT through Article XXIV which allows Members to form such arrangements provided they eliminate, rather than just lower, within-union trade barriers on 'substantially all trade.' The arrangements must also not raise trade barriers on goods produced outside the union. These provisions together legalise PTAs in which member countries do not raise their tariffs on outside countries and CUIs in which the CET is chosen so as to leave the average external tariff unchanged. The provisions rule out partial PTAs, however. Article V in the General Agreement on Trade in Services (GATS) introduces provisions for PTAs in services which parallel those in Article XXIV of GATT for goods.

In the original GATT, signed in 1947, Article XXIV offered the only provisions for new PTAs. Subsequently, however, with the addition of Part IV in 1965 and the Enabling Clause in 1979, GATT permitted partial PTAs under two circumstances: developed countries were allowed to grant one-way partial tariff preferences to developing countries (through Article I exemption) and two or more developing countries were given the right to exchange two-way partial trade preferences. Under the former provision, developing countries have benefited from tariff preferences granted by developed countries under the rubric of the Generalised System of Preferences (GSP). Under the latter

4 Trade barriers such as anti-dumping which are permitted by other provisions of GATT can be maintained within the union.

5 The existing arrangements involving partial preferences, principally those between former colonies and their colonial powers, were granted exemption from the MFN provision within Article I.
provision, developing countries have exchanged partial tariff preferences within arrangements such as the ASEAN Preferential Trading Area (APTA) and South Asian Preferential Trading Area (SAPTA).° Not being full-fledged FTAs or CUs, these arrangements would not have been permitted under Article XXIV.

It is important to emphasise that the presence of even a single developed country in a PTA imposes an important constraint on the nature of the arrangement: to be GATT consistent, it must satisfy the requirements of Article XXIV which are far more demanding than those of the Enabling Clause. This has made, for example, the exchange of partial tariff preferences among members of the Asia Pacific Economic Cooperation (APEC) forum a difficult task. On the other hand, the bulk of PTAs consisting exclusively of developing countries, including even the prominent ones such as the Southern Cone Common Market (MERCOSUR) and ASEAN Free Trade Area (AFTA), have been concluded under the Enabling Clause provisions and can stop well short of a genuine FTA or CU if the members so desire. Admittedly, in the past, developed-country PTAs have also violated the spirit of Article XXIV provisions; on balance, however, the Enabling Clause offers a much greater scope for the exchange arbitrary preferences than Article XXIV.

c. Historical Evolution

Prior to the early 1980s, PTAs were limited to arrangements within Western Europe, among developing countries, and trade preferences by developed to developing countries. Because the developing-country arrangements, undertaken principally in Latin America and Africa, were largely ineffective and trade preferences by developed to developing countries limited, effective PTAs were confined to the two arrangements in Western Europe: the European Community (EC) and the European Free Trade Area (EFTA). The limited role for PTAs meant that the architects of the global trading system did not have to fear that regional arrangements might undermine the multilateral process of trade liberalisation. Indeed, with the EC negotiating its common external tariffs as a single unit, the United States, which was strongly committed to the multilateral process, was successful in leading the world into as many as seven rounds of multilateral trade negotiations (MTNs).

Throughout this earlier period, while the European Community widened and deepened its integration, the United States remained singularly committed to the multilateral approach. Having witnessed the pernicious effects of discriminatory trade and payments regimes during the Great Depression, at the end of the Second World War she had emerged as the champion of a non-discriminatory global trade regime, grounded firmly in the MFN principle. Speaking for the US policy

° ASEAN stands for the Association of South East Asian Nations.
makers, Howard Ellis (1945) denounced bilateral arrangements in the strongest terms:

There are good reasons for believing that no device portends more restrictions of international trade in the postwar setting than bilateral arrangements.

Proposals were made during the 1960s for a North Atlantic Free Trade Area but received no attention from the United States. All this changed, however, when, at the GATT Ministerial in November 1982, the United States began efforts to start the eighth round of multilateral trade negotiations. Unable to persuade the EC to go along, she felt obliged to abandon her long-standing opposition to regional arrangements. Recognising that PTAs were the only means left for keeping the process of trade liberalisation afloat, the United States went on to conclude an FTA with Israel in 1985 and Canada in 1989. Though the Uruguay Round was launched in the meantime, because the EC remained a reluctant player at the negotiating table, the United States moved ahead with yet another PTA, this time jointly with Mexico and Canada. Side by side, the European Community continued its expansion, adding Greece in 1981, Portugal and Spain in 1986 and Austria, Finland and Sweden in 1995.

Today, with the Uruguay Round having been successfully concluded and the multilateral process working well, the original rationale for the US pursuit of PTAs has disappeared. But what was originally viewed as a temporary diversion to force the European Community to the negotiating table has turned into a race for securing preferentially the neighbours’ markets for one’s exports. The European Community, renamed the European Union following the Maastricht Treaty, has moved aggressively to conclude FTAs with its neighbours in Eastern and Central Europe and with the Baltic Republics while the United States has gone on to promote the idea of a Free Trade Area of the Americas. This race between two giants has, in turn, led to a renewal of efforts for PTAs by and among smaller countries in Africa, Latin America, South and Central Asia, Central and Eastern Europe and the Baltic Republics. The only region which has so far remained firmly committed to the MFN approach to liberalisation is East Asia.

Because PTAs are inherently discriminatory, their proliferation has led to fears that they may undermine the multilateral process of trade liberalisation. The issue became a key subject of discussions at several WTO working parties which had been looking at the regional arrangements notified to the WTO in recent years. In February 1996, recognising the importance of the issue, the WTO appointed a Committee on Regional Trade Agreements (CRTA) to give coherence to these discussions. A key charge of CRTA is to examine in detail whether regional arrangements are compatible with multilateralism.

Though the Committee’s work is still in progress, the Singapore Ministerial Declaration provides a preliminary impression of the Member countries’ view on the subject. The paragraph dealing with regional arrangements states:
We note that trade relations of WTO Members are being increasingly influenced by regional trade agreements, which have expanded vastly in number, scope and coverage. Such initiatives can promote further liberalization and may assist least-developed, developing and transition economies in integrating into the international trading system. In this context, we note the importance of existing regional arrangements involving developing and least-developed countries. The expansion and extent of regional trade agreements make it important to analyse whether the system of WTO rights and obligations as it relates to regional trade agreements needs to be further clarified. We reaffirm the primacy of the multilateral trading system, which includes a framework for the development of regional trade agreements, and we renew our commitment to ensure that regional trade agreements are complementary to it and consistent with its rules. In this regard, we welcome the establishment and endorse the work of the new Committee on Regional Trade Agreements. We shall continue to work through progressive liberalization in the WTO as we are committed in the WTO Agreement and Decisions adopted at Marrakesh, and in so doing facilitate mutually supportive processes of global and regional trade liberalization.

Thus, the view taken by the Conference and, indeed, by the participating Ministers in their individual statements, is that regional arrangements are compatible with multilateralism. As we will see, this view is not shared by free-trade purists who see PTAs as a source of fragmentation of the global trading system rather than a unifying force.

2. EFFECTS ON UNION MEMBERS

The effects on union members can be studied under four headings: (i) Vinerian analysis which focuses on the welfare effects of a PTA, holding the tariffs on outside countries unchanged, (ii) the implications of differences in transport costs across potential union members, (iii) the implications of the rules of origin, and (iv) non-traditional gains including guaranteed market access, shelter from contingent protection, locking-in the reforms and dispute settlement. We take each of these effects in turn.


Advocates of regional arrangements sometimes claim that the conventional static welfare analysis is too narrow a criterion to judge overall desirability of the arrangements. But this is not a defensible position since PTAs are the principal, often only, component of regional arrangements with serious economic effects of which the static welfare effect is a key and best understood component. It is no accident that economists evaluating regional arrangements, be it various EC extensions or NAFTA, have almost always focused on their static welfare effects.

In static analysis, Viner's (1950) seminal concepts of trade creation and trade diversion remain central today. Viner noted that since PTAs liberalise trade
preferentially, on the one hand, they 'create' new trade between union members while, on the other, they 'divert' trade from low-cost outside suppliers to high-cost within-union suppliers. The former effect arises from a union partner undermining another union member's less efficient industry and is beneficial. The latter effect arises from a union member displacing a more efficient outside supplier by taking advantage of the tariff preference it enjoys in a partner country. This effect is harmful. Unions which are primarily trade creating are beneficial and those that are primarily trade diverting are harmful to member countries taken together and to the world as a whole. As we will see shortly, however, an individual member of the union can reap large benefits even from a primarily trade diverting union through a shift in intra-union terms of trade in its favour.

It is easy to see that when trade is multilateral, that is, countries import from and export to union members as well as outside countries, trade diversion is inevitable. Moreover, if potential union members are small in relation to the outside world as is likely, little trade creation will be forthcoming. This rather surprising point has not been fully appreciated but can be made easily with the help of a simple partial-equilibrium example.

Suppose Indonesia and Singapore form an FTA under which Indonesia removes its tariff on imports of video-cassette recorders (VCRs) from Singapore but retains it on outside countries (which include such competitive suppliers as China, Republic of Korea and Taiwan). The outcome of such an FTA is shown in Figure 1, where Indonesia is distinguished by subscript I, Singapore by S and the rest of the world by W. Curve $M_WM_I$ represents Indonesia's import demand for VCRs (assumed to be homogeneous), $E_SE_S$ represents Singapore's supply of exports of VCRs, and $P_WE_W$ gives the supply of VCRs from the world market. It is assumed that Indonesia and Singapore are small in relation to the world and take the world VCR price as given.

Initially, Indonesia levies a non-discriminatory tariff on imports equalling $P_WP_W$ per VCR so that export supply curves of Singapore and the rest of the world, as viewed by Indonesian consumers and producers, are given by $E_SE_I'$ and $P_WE'_W$, respectively. The price of VCRs in Indonesia is $P_W$. The country imports $OQ_I$ from Indonesia and $Q_1Q_3$ from the rest of the world, collecting $GSNH$ in tariff revenue. The consumers' surplus is given by triangle $KSG$.

An FTA with Singapore results in the elimination of the tariff on Singapore but its retention on the rest of the world. This change shifts the supply curve from Singapore down to $E_SE_S$. As long as any VCRs continue to come from the rest of the world, the price in Indonesia remains unchanged at $P'_W$. Imports from Singapore rise to $OQ_2$ and those from the rest of the world fall to $Q_2Q_3$.

Several conclusions can now be drawn. First, as noted above, the FTA turns out to be wholly trade diverting. With no change in the external tariff, as long as VCR imports continue to come from the rest of the world, the domestic price of VCRs in Indonesia cannot change and domestic producers are unaffected by increased
competition from Singapore. The margin on which they must compete is defined by the rest of the world before as well as after the formation of the FTA. There is no trade creation. On the other hand, taking advantage of the preference, Singaporean suppliers of VCRs are able to displace some of the more efficient outside suppliers, leading to trade diversion of $Q_1Q_2$ and an associated deadweight loss of triangle $FLU$.

Second, the FTA generates a redistributive effect which hurts Indonesia who extends the preference and benefits Singapore who receives it. Indonesia loses rectangle $GFLH$ which it previously collected in tariff revenue but no longer does. The rectangle goes to exporting firms in Singapore who boost their profit by trapezium $GFUH$. Triangle $FLU$ is deadweight loss due to trade diversion and pays for the higher production costs of output $Q_1Q_2$ by Singaporean firms relative to outside firms. Due to the large redistribution effect, which is a rectangle, the loss to Indonesia in this example is much larger than the deadweight loss from trade diversion which is a triangle. And the larger the quantity of trade with Singapore, the larger the redistribution and the greater the loss. Of course, if Indonesia was to receive trade preferences from Singapore on its exports to the latter, its exports will receive a rectangle and Singapore will lose it, thus neutralising the redistribution effects. The problem, however, is that if Singapore’s tariffs are near zero, as is true, it can offer Indonesia no trade
preferences. Under such circumstances, the FTA amounts to one-way preferences in which the high-tariff country loses and low-tariff country gains.7

Finally, if Indonesia chooses to remove the tariff on an MFN basis as it has actually done during the last several years, no loss occurs. In this case, the tariff is removed on the rest of the world as well and the price of VCRs in Indonesia falls to the world level. The lost tariff revenue, which was transferred to Singaporean exporters under the FTA, is now transferred to Indonesian consumers. In addition, the country makes a net gain of triangle SNR from improved efficiency. Thus, for small countries, unilateral liberalisation on an MFN basis remains a vastly superior option to a PTA.

It has been argued by many, most prominently Summers (1991), that while considering PTAs, we should not worry about trade diversion. To quote him:

I find it surprising that this issue is taken so seriously – in most other situations, economists laugh off second best considerations and focus on direct impacts.

The analysis just presented shows that trade diversion is not something that can be laughed off. Specially when unions are between a high-tariff country such as Mexico and a low-tariff country such as the United States, losses from the PTA to the former can be considerable. In Panagariya (1997a), I have estimated that the redistributive effect shown in Figure 1 may be costing Mexico as much as $3.25 billion per year.

Summers rests his conclusion on the claim that when union members are ‘Natural Trading Partners,’ that is, they already trade a lot with each other and are geographically proximate, the risk of trade diversion is minimal. In his own words,8

Are trading blocs likely to divert large amounts of trade? In answering this question, the issue of natural trading blocs is crucial because to the extent that blocs are created between countries that already trade disproportionately, the risk of large amounts of trade diversion is reduced.

Bhagwati (1995), Panagariya (1996) and Bhagwati and Panagariya (1996a) have offered detailed critiques of this view. Bhagwati and Panagariya (1996a), in particular, show systematically that the natural trading partners hypothesis has no analytic basis. To give some flavour of their critique, note that this criterion is neither symmetric nor transitive. According to Summers’ definition, the United

---

7 Using a simulation model, a recent paper by Spilimbergo and Stein (1996) shows that the 'rectangle effect' continues to dominate in models of product differentiation of the Krugman (1980) variety. They find that a high-tariff country is better off forming an FTA with a high-tariff country than a low-tariff country.

8 The natural trading partners idea was originally proposed by Wonnacott and Lutz (1989) and was also discussed by Jacquemin and Sapir (1991). Krugman (1991a) endorses the idea in the following excerpt. 'To reemphasize why this matters: if a disproportionate share of world trade would take place within trading blocs even in the absence of any preferential trading arrangement, then the gains from trade creation within blocs are likely to outweigh any possible losses from external trade diversion.'

© Blackwell Publishers Ltd 1999
States may be the natural trading partner of Mexico but the reverse is not true. Similarly, the United States is a natural trading partner of both Mexico and Canada but these two are not natural trading partners of each other. More directly, it should be obvious from Figure 1 that trade diversion is a marginal concept and, therefore, has nothing to do with the initial level of trade. While the scope for trade diversion may depend on the extent of intra-union trade, the actual trade diversion depends entirely on the response of partner country's exports to the tariff preference at the margin. Yet another point, made forcefully by Panagariya (1995 and 1996) and illustrated in Figure 1, is that if a country forms a PTA with another country with substantially lower tariffs than its own, its losses are larger the more it imports from the partner!

Because weaker, uncompetitive industries are usually the ones that succeed in lobbying against foreign competition, PTAs get voted in precisely when trade diversion is the dominant force. This is a point made formally in the recent political-economy-theoretic analyses of Grossman and Helpman (1995) and Krishna (1996). The careful empirical work by Kowalczyk and Davis (1996) on the tariff phase out in NAFTA shows that the sectors which were allowed the longest phase out periods in the United States were the ones in which import-competing lobbies were the strongest.

More direct empirical evidence supporting trade diversion is also beginning to accumulate. A recent, widely-publicised, World Bank study by Yeats (1996) provides systematic evidence of wholesale trade diversion in MERCOSUR. Similarly, Wei and Frankel (1996) find that various extensions of the European Community were accompanied by a considerable trade diversion. Referring to the EC expansion during 1980s, Wei and Frankel note:

Overall, the evidence suggests massive trade diversion resulted from the membership expansion.

According to the key finding on which this conclusion is based:

Imports from non-member countries in 1990 were 30% lower than in 1980, after controlling for economic growth. Thus, the possibility of trade diversion cannot be taken lightly.

b. Transport Costs and PTAs

Following an initial suggestion by Krugman (1991a), Frankel (1996) and his associates (Frankel, Stein and Wei, 1995; and Frankel and Wei, 1997) have

---

9 Richard Eckaus of M.I.T. and Robert Scott of the University of Maryland have found that even in the case of NAFTA, there has been a substantial trade diversion in the textile and clothing industry. 10 It is important to note that establishing trade creation and trade diversion empirically is a difficult task. As such these studies are not without flaws. In the case of Yeats (1996), despite pitfalls of specific measures used, the overall evidence is compelling. Wei and Frankel (1996) employ the gravity equation which, despite many modifications by different authors, is likely to remain misspecified and, thus, open to criticism (e.g., see Polak, 1996).
strongly pushed the idea that the presence of transport costs make PTAs among proximate countries an attractive option. This is a surprising development since none of the original pioneers including Viner (1950), Meade (1955) and Lipsey (1958) had mentioned transport costs or proximity as a factor in determining the desirability of PTAs. Building on the earlier critique in Bhagwati (1993) and Bhagwati and Panagariya (1996a), I have recently subjected this view to a thorough examination (Panagariya, 1997b). My unequivocal conclusion is that transport costs are not different than any other costs and as such deserve no special attention in considering PTAs. Here I discuss some key points.

There is nothing in economic theory which says that countries located far apart must gain less from trade than those located close to each other. This should be clear enough from the analysis associated with Figure 1 which does not depend on the source of cost differences between Singapore and the rest of the world. As long as the delivery price of Singapore is given by the height of the curve $E_S$, regardless of whether transport costs contribute a large, small or no part of it, the analysis remains unchanged.

Alternatively, consider an example, given in Panagariya (1997b), in which trade blocs between proximate countries confer no gains while those between distant countries do. Suppose the world consists of two continents, two countries per continent, and two goods produced at constant but different labour costs à la Ricardo. Suppose further that the countries located on the same continent are identical in all respects but differ across continents. Despite positive transport costs across continents but none within a continent, there are no gains from forming continental blocs whereas, with sufficiently large comparative cost differences across continents, gains are available to blocs between countries across continents.

Bhagwati and Panagariya (1996a) show that in general even a limited proposition which makes a PTA between proximate partners *ceteris paribus* superior to that between distant partners is false. These authors provide an example in which a country facing a proximate and a distant partner who are otherwise identical is better off giving a tariff preference to the latter.

c. Rules of Origin in FTAs

In FTAs as distinct from CUs, member countries retain their own outside tariffs. Ignoring internal transport costs, this feature opens the possibility that a product destined to a high-tariff member country will be first imported into the lowest-tariff member country and then re-exported to the former. Or, more subtly, if inputs imported from outside countries constitute a large part of the value added of a product, producers in the member country with lowest tariffs on inputs can undercut producers in other member countries. To guard against these possibilities, FTAs are usually supplemented by rules of origin which can take
various forms. For example, the agreement may specify that a product will qualify for duty-free movement within the union only if a pre-specified proportion of its value added has originated within the union. Alternatively, it may require that a product undergo a substantial transformation in a country before being permitted to cross the border of a partner country free of duty.

Analytic literature on the rules of origin in the context of PTAs is still in its infancy. But from a policy perspective, some simple points can be made. First, at least in the absence of traded intermediate inputs, rules of origin have an unambiguously harmful effect. If there are no rules of origin, each product will be imported through the member country that has the lowest external tariff. This will minimise the trade diversion effect of internal preference in higher-tariff member countries by lowering the effective external tariff down to the level of the member with the lowest tariff on the product. Effectively, the FTA will be turned into a CU with the lowest tariff among union members serving as the common external tariff. Rules of origin, by contrast, allow more trade diversion in a member country with a higher tariff by requiring that goods destined to a high-tariff country be imported through the border of that country.

Second, if traded inputs are used in production, rules of origin can paradoxically reduce trade diversion. Thus, recall that tariff preferences can give rise to harmful trade diversion by giving within-union producers an extra advantage over outside suppliers. To the extent that rules of origin may counteract this advantage, they may be beneficial. For example, taking advantage of a tariff preference in a partner country, the producer of a final good in a member country may be able to outcompete his outside competitor provided he can import the input from outside sources. But if the rules of origin require that he purchase the input from more expensive internal sources to qualify for the tariff preference on the final good and such a purchase raises the cost of production above the tariff inclusive price of outside suppliers, harmful trade diversion in the final good will be avoided.

Third, for the same reason that rules of origin may counteract trade diversion, they may also counteract trade creation. The ability of a partner country to undermine an inefficient domestic industry is reduced by rules of origin which require it to purchase inputs from less efficient internal sources.

Fourth, in intermediate goods production itself, the rules of origin are likely to be harmful. To take advantage of the tariff preference in the final good, union members will shift their purchases towards intra-union sources of inputs even though these sources may be more expensive. It is this harmful role of rules of origin which is emphasised by Krueger (1993) in her critique of FTAs.

---

11 See Krueger (1993) and Krishna and Krueger (1995) for a preliminary analysis of some policy issues. Theoretical literature on the subject is so far sparse. Thus, see Falvey and Reed (1997 and 1998) and the references therein.
Fifth, the rules of origin can multiply distortions as overlapping FTAs begin to form. For example, if Chile, who already has an FTA with MERCOSUR, joins NAFTA, a Chilean firm will have to buy components in Brazil if it wants to take advantage of the preferential tariff in MERCOSUR and in the United States if it wants to exploit the preference in NAFTA. This notwithstanding the fact that the most efficient supplier of the components may be in Asia.

Finally, in a political-economy context, rules of origin can turn into yet another instrument available to lobbies for adding to trade diversion or frustrating trade creation. That this may indeed have happened in NAFTA is illustrated by the fact that the stringent rules of origin were sought by the United States even though US tariffs were generally lower than those in Mexico. In the absence of rules of origin, the possibility of final goods imports coming into the United States through Mexico was minimal. Nor would it have made sense for a producer in the United States to import inputs through Mexico. The intent of the rules of origin had to be protectionist: a stringent rule of origin would undermine Mexico’s ability to outcompete an inefficient US firm producing final goods and also make internal market in inputs more profitable. In one area, textiles and clothing, where US tariffs were high and the scope for trade creation substantial, the triple transformation rule of origin was adopted to maintain a high level of protection for US producers.

d. Non-traditional Gains: Guaranteed Market Access, Shelter from Contingent Protection, Locking-in the Reforms and Dispute Settlement

The debate on regionalism initiated by NAFTA brought several new elements into the analysis of FTAs. Many economists and policy analysts expressed the view, though without supporting evidence or analysis, that the gains to a small developing country from a PTA with a large developed economy go well beyond the traditional static welfare effects. In addition it has been argued that since the dispute settlement mechanism in a PTA such as NAFTA or the EU gives direct access to private parties, it works more effectively than the dispute settlement mechanism of the WTO which is available exclusively to member-country governments.

Three non-traditional benefits to a developing country from entering into a PTA agreement with a large developed country were identified during the NAFTA debate. First, such arrangements guarantee access to a large market. The agreement ensures the small developing country that if the rich partner turns protectionist in the future, its access to the latter’s market will be preserved. The cost in terms of opening one’s own market preferentially to the large country can, thus, be seen as the cost of insurance against possible loss of access to the partner’s market.

Second, the PTA can shield the developing-country member from administered protection in the rich country. For instance, the country may escape anti-
dumping and safeguards actions by the rich partner to which other trading partners can be subject.

Finally, a regional arrangement with a large, rich trading partner can be an effective instrument of imparting credibility to reforms. An international treaty with a large and rich country can ‘lock’ the reforms, making it difficult for more protection-minded future governments to reverse the actions of their predecessors. An important element of this ‘lock,’ emphasised recently, is the more effective dispute settlement process of PTAs which in the case of NAFTA is available to private parties including labour unions, business groups and activists. The WTO dispute settlement process, by contrast, is available to the governments of member countries only.

Panagariya (1995 and 1996) offers a systematic critique of these arguments. Since the benefits were identified originally as accruing to Mexico from NAFTA, let us examine them in that context. Consider first the market access argument. Under normal circumstances, access to the US market is guaranteed by WTO agreements. In order for market access to be curtailed in the absence of NAFTA, the United States will have to go back on its WTO obligations including the possibility of a complete withdrawal from that institution. Under such circumstances, it is not clear that the United States will adhere to its NAFTA obligations. The bottom line is: Is the US commitment to NAFTA more credible than to WTO? If not, the insurance argument is quite weak.

Next, take the argument that NAFTA shelters Mexico from administered protection. At least in principle, this is a misleading claim since the NAFTA agreement contains nothing that says that member-country firms are to be treated differently from firms in outside countries for purposes of administered protection. Indeed, the special agreements on sugar and orange juice explicitly allow for the play of administered protection in the event of import surges from Mexico in these sectors. Furthermore, side agreements on labour standards and environment give the United States new powers to subject Mexico to dispute settlement procedures which can lead to fines of up to $20 million.

Turning to ‘lock-in’ effects, it is unlikely that NAFTA can ‘lock in’ all reforms. For instance, it surely does not guarantee macroeconomic stability as was demonstrated by the peso crisis. It also does not advance the cause of privatisation in any meaningful way. The main area where the lock-in argument may apply is trade policy. But here there are two problems. First, the lock applies only to trade preferences granted to NAFTA members. If these preferences were harmful to Mexico in the first place, the lock is not a benefit but a cost. Second, Mexico could have as easily locked in its trade reforms on a multilateral basis by binding tariffs with the WTO at the applied rates. Instead, it chose to bind tariffs at levels much higher than applied rates. Paradoxically, NAFTA may have contributed to the water in the bindings. Recognising that it will not be possible to raise tariffs on the bulk of imports coming from the United States, Mexican
THE REGIONALISM DEBATE: AN OVERVIEW

authorities may have decided to leave themselves considerable room in the choice of external tariffs in case pressures from domestic industry necessitate a rolling back of trade liberalisation. As it turned out, this flexibility was used after the peso crisis with tariffs on 503 items rising from less than 20 per cent to 35 per cent.

Finally, consider the benefits of the alternative dispute settlement process in the EU and NAFTA. The argument here is that it is more effective since it is available directly to private parties whereas the WTO process is available to member-country governments only. But as Levy and Srinivasan (1996) show, this is not a compelling argument. They show that the access of private parties can lead governments not to sign agreements that are otherwise beneficial.

3. IMPLICATIONS FOR THE GLOBAL TRADING SYSTEM

There are several questions one may ask about the relationship between PTAs and multilateral liberalisation. Can a PTA expand continuously to yield worldwide free trade? Can it make an otherwise feasible multilateral liberalisation infeasible? Does it offer an incentive to increase or reduce trade barriers against outside countries? What kind of transition trade regimes are we likely to have with the formation of criss-crossing PTAs? Let us take these and related questions in turn.

a. Can PTA Expansion Lead to Global Free Trade?

Assuming the PTA and multilateral processes to be independent, can the expansion of a PTA continue until we achieve worldwide free trade? This is what Bhagwati and Panagariya (1996a) labelled Question I relating to the time path issue, introduced originally by Bhagwati (1993). As noted by them, this question is not about the existence of a PTA path along which world welfare increases monotonically as analysed by Kemp and Wan (1976) and Panagariya and Krishna (1997).\textsuperscript{12} Nor is it about the relationship between the number of trade blocs and welfare as analysed by Krugman (1991b). Instead, it concerns the monotoncity of incentive to seek entry on the part of outsiders and to offer entry on the part of insiders until global free trade is achieved.

Baldwin (1995) deals with this issue focusing on the incentives of outsiders to seek entry. Using a variant of what have come to be known as models of

\textsuperscript{12} Kemp and Wan (1976) demonstrated that two or more countries can always form a customs union which makes member countries better off without making outside countries worse off. An analogous result for PTAs was not available until the recent contribution of Panagariya and Krishna (1997).
economic geography, he identifies a ‘domino’ effect according to which idiosyncratic events such as the European Single Market initiative create economic incentives for outside countries to seek entry into an existing PTA. Unless there are sufficiently strong non-economic factors that counter these incentives, as the PTA expands, eventually all countries want to enter the PTA. Then, as long as entry into the PTA is free, as indeed assumed by Baldwin, this process can lead to global free trade.\textsuperscript{13}

There are two key limitations of Baldwin’s otherwise elegant analysis. First, working in the tradition of economic-geography models, he formalises trade barriers as transport costs. As such, accession to the PTA becomes equivalent to a reduction in transport costs. This means that the tariff revenue aspect of trade barriers is completely absent in his analysis. It is not clear whether his result will remain valid once transport costs are replaced by tariffs with tariff-revenue effect of the entry into the PTA taken into account. Second, even if we ignore this problem, Baldwin (1995) assumes that ‘insiders’ have no incentive to block the entry. It may be hypothesised that after the PTA reaches a certain size, insiders will have an incentive to block further entry.

This is indeed the message of a recent elegant paper of Andriamananjara (1998) who explicitly models the incentives facing outsiders to seek entry and willingness of outsiders to give entry. He uses a Cournot oligopoly model of identical countries in which the outside tariff is assumed to be fixed and decisions to seek and offer entry are driven by profits. He shows that in this model as the PTA expands, profits of insiders first rise, reach a maximum and then decline. Moreover, the maximum-profit point is reached before the PTA comes to encompass all countries. Profits of outsiders, on the other hand, decline monotonically as the PTA expands. Thus, while outsiders have an increasing incentive to seek entry, insiders stop short of taking all of them into the club. The PTA fails to expand into a global bloc.

Using the symmetric Krugman (1991b) model, modified to allow for comparative advantage, Bond and Syropoulos (1996) ask the following related question. Suppose countries and blocs behave as Nash players and choose tariffs to maximise welfare. The world is initially divided into two or more identical blocs. If one of these blocs now begins to expand by drawing one country from each of the remaining blocs at a time, with Nash-optimum tariffs chosen at all times by all blocs, will the expanding bloc eventually turn into a global bloc or stop short of it. Using simulations, they show that as this bloc expands, the welfare of its members peaks before it absorbs the other blocs in their entirety. The process stops short of yielding global free trade.

\textsuperscript{13} A similar result is obtained by Yi (1996) in a somewhat different model than that of Baldwin.
b. Do PTAs Make Multilateral Liberalisation Less Likely?

Assuming the PTA process interacts with the multilateral process, does it serve as a building block or a stumbling block to global free trade? This is the alternative time-path Question II identified in Bhagwati and Panagariya (1996a). There are both formal models and informal arguments focusing on this issue. Let us take them in order.

(i) Formal models

Using the median voter model, Levy (1997) asks the question whether presenting the option to form an FTA can make a previously infeasible multilateral liberalisation feasible and whether it can render a previously feasible multilateral liberalisation infeasible. The question is addressed in a multi-country world in which voters in two countries are given the option to vote on an FTA and multilateral liberalisation in that order. The FTA is modelled as complete free trade between union members with a prohibitive tariff on non-members. Levy shows that in the Heckscher-Ohlin set up, a previously feasible multilateral liberalisation cannot be rendered infeasible by the FTA option. Essentially, when multilateral liberalisation is feasible, the FTA offers a lower utility than global free trade to the median voter in at least one country so that the FTA is necessarily rejected by at least one country. Levy also shows that the FTA cannot make a previously infeasible multilateral liberalisation feasible. This is simply because the FTA is voted favourably only if it raises the median voter’s utility over the autarky equilibrium. But that change only raises his reservation utility.

If the problem is considered in a Krugman (1980) type of model with product differentiation, Levy shows that the FTA can turn into a stumbling block to global free trade. With the gains from trade deriving from differences in factor endowments as well as increased variety, it now becomes possible for median voters in both countries forming the FTA to achieve a higher level of utility under this alternative than available under the multilateral option.

Krishna (1998) offers another example of FTAs turning into stumbling blocks to global liberalisation. He uses a three-country, partial-equilibrium, oligopoly model in which trade policy is chosen to maximise national firms’ profits. He shows that more trade diverting the FTA between two countries in this set up, the greater the backing it receives and more it reduces the incentive to eventually liberalise with the third country. With sufficiently large trade diversion, an initially feasible multilateral liberalisation can be rendered infeasible by the FTA option.

(ii) Informal arguments

In addition to these formal analyses, several informal arguments have been made to support the building blocks or stumbling blocks hypothesis.
First, it has been suggested by Summers (1991) and others that multilateral negotiations will move more rapidly if the number of negotiators is reduced to approximately three via bloc formation. This argument gained some popularity at the time the Uruguay Round negotiations were stalled but has lost force since the successful completion of the round. The argument is that due to a large number of members involved at the WTO and the associated free rider problem, negotiations there are slow and difficult. If the world is first divided into a handful of blocs, multilateral negotiations will become easier.

There are at least two problems with this argument, however. First, theoretically speaking, if blocs take the form of FTAs, they have no effect on the number of participants; for FTA members retain their own external tariffs and must negotiate these tariffs individually. So far, the only PTA which participates as a single unit in multilateral negotiations is the EU. Second, as a practical matter, it is not clear that the existence of the European Community, now the EU, has been an unmixed blessing for multilateral negotiations. Preoccupied with its internal problems and agenda, it was the EC which first delayed the launching and then the conclusion of the Uruguay Round. It is possible that if EC members were participating individually in negotiations, they would see greater merit in multilateral liberalisation and be more eager to negotiate.

According to the second informal argument, PTAs may serve as a threat to force unwilling parties to negotiate in earnest at the multilateral level. Bergsten (1994) argues that the upgrading of the November 1993 APEC ministers meeting in Seattle by President Clinton to high-profile Leaders’ Meeting signalled to the European Community that if they dragged their feet at the Uruguay Round, the United States would go ahead with an FTA with Asian countries. This threat led the European Community to conclude negotiations. Bhagwati (1996a) disagrees with this interpretation of events. He argues that the Uruguay Round was completed essentially because the United States wisely decided to close the deal, taking the offer on the table rather than seeking more concessions. Given how distant and remote was, and still is, the possibility of an APEC FTA, it is difficult to imagine that the European Community would have taken the threat seriously, suggesting that Bhagwati’s interpretation is correct.

Third, on the stumbling blocks side, it is argued that due to their high visibility, PTAs can energise and unify protectionist lobbies, turning them into effective obstacles against multilateral liberalisation. This is specially true since the PTAs likely to be negotiated by developed countries are with developing countries. Such PTAs are associated in public mind (in developed countries) with large inflows of labour-intensive goods and reduced wages for the unskilled. Multilateral negotiations, by contrast, involve both developed and developing countries and draw less attention of protectionist lobbies. Thus, the NAFTA debate in the United States was more fierce than the debate for any multilateral trade negotiation including the Uruguay Round. More recently, President
Clinton's efforts to obtain the fast track authority, aimed principally at bringing Chile into NAFTA, met with failure. It is possible that if the President had sought this authority to negotiate a multilateral round, instead, he would have faced less opposition.

Finally, there is the related issue of attention diversion and scarce negotiating resource. If the President of the United States and his Trade Representative are preoccupied with cutting deals in Latin America, they will have less time and motivation for multilateral negotiations. Furthermore, even if they had the time and motivation, it may become difficult to persuade the US Congress to go along with a multilateral round with negotiating capital having been used up for regional deals.

c. Do PTAs Lead to a Rise in Trade Barriers Against Non-members?

Closely related to the issue whether PTAs make multilateral liberalisation less likely is the question whether they lead to a rise in trade barriers against non-members. Like the former question, this one is also a political-economy theoretic question and the answer depends on the political process at work in a given situation. The theoretical literature, summarised below, offers arguments on both sides. Although the analysis is generally couched in terms of tariffs, it should be interpreted to apply to trade barriers in general which include WTO-sanctioned instruments such as anti-dumping and safeguard measures.

(i) An FTA leading to a reduction in tariffs on inputs

Suppose that in the pre-PTA equilibrium, relative to its potential partner, a union member faces lower tariffs on final goods and higher tariffs on intermediate inputs in some sectors. The PTA then places the country's final goods producers at a disadvantage vis-à-vis its union partner in these sectors. If the final goods producers are politically powerful, they will succeed in getting the external tariff on the input reduced to the partner country's level. A similar process will operate in the partner country in sectors with lower tariff on final goods and higher tariffs on intermediate inputs. Thus, the PTA will lead to further liberalisation. It has been suggested that this factor has played some role in the MFN liberalisation of certain inputs in Canada.

This seemingly neat argument favouring PTAs has three important limitations. First, the liberalisation envisaged by it is in all likelihood a welfare-reducing proposition. As is well known, a reduction in the tariff on an input increases effective protection to the final good and can turn out to be a move away from

---

14 In the case of developing countries, applied tariffs are generally below GATT bindings so that barriers can be raised by raising tariffs. In developed countries where actual tariffs are constrained by GATT bindings, this objective can be achieved through safeguard measures including anti-dumping actions.

© Blackwell Publishers Ltd 1999
rather than towards free trade. In trying to explain why so many free traders of his time supported PTAs, Viner (1950) had warned against precisely this type of liberalisation. To quote him,

The major explanation (of why so many free traders support PTAs) seems to lie in an unreflecting association on their part of any removal or reduction of trade barriers with movement in the direction of free trade. Businessmen, however, and governments which had to try simultaneously to satisfy both special interests seeking increased protection and voters hostile to protection, have long known ways of making increased protection look like movement in a free-trade direction. ... Let us suppose that there are import duties both on wool and on woolen cloth, but that no wool is produced at home despite the duty. Removing the duty on wool while leaving the duty unchanged on woolen cloth results in increased protection for the cloth industry while having no significance for wool-raising (Viner, 1950, p. 48).

Second, the argument assumes that the option of raising the external tariff on the final product is not available. In virtually all developing countries, external tariffs are bound at levels above their applied levels. As a result, lobbying by producers of final goods can very well result in increased tariff on the final good rather than a reduction in the tariff on the input.

Finally, the argument relies on an essentially flawed political-economy model. It assumes that the producers of intermediate inputs do not lobby against the reduction in tariff and, thus, relies on an ad hoc asymmetry between the availability of protection to final goods and intermediate inputs. The asymmetry could be justified if intermediate inputs were not produced at home. But in that case, there will be no reason for a tariff on them in the first place.

\( (ii) \) Reducing the external tariff to reverse trade diversion

Richardson (1993) has provided an interesting case in which, following the formation of PTA, a reduction in the external tariff can result from the government’s desire to maximise a political support function. This case, illustrated in Figure 2, occurs when a union member faces a constant supply price of a product from its partner as well as the rest of the world and the partner’s price is higher. It is well known that the FTA between Singapore and Indonesia in Figure 2 is trade diverting. Letting \( P_w P_w' \) be pre-FTA per-unit tariff, all imports come from the rest of the world initially and the price in Indonesia is \( P_w' \). The FTA brings the internal price down to \( P_S \) with all imports diverted to the union partner, Singapore. In Indonesia, pre-FTA tariff revenue, represented by areas \( a + b \), disappears and the expansion of imports due to the price decline increases joint gains of consumers and producers by area \( a + c \). Thus, the net gain to Indonesia from the FTA is \( b - c \) which is likely to be negative though not necessarily.

Richardson’s argument is that once the FTA is in place, by reducing the tariff on the outside world by \( P_S P_w' + \varepsilon \) where \( \varepsilon \) is an arbitrarily small number, Indonesia can switch imports back to the rest of the world and, thus, reverse the initial trade diversion. The switch allows the country to recover the lost area \( b \) in
tariff revenue. Richardson shows that such a tariff reduction necessarily obtains in an incentive-theoretic model in which the government maximises a political support function which gives a positive weight to consumers' surplus and tariff revenue, no matter how small in relation to the producers' surplus.

There are three key limitations of this argument which considerably limit its practical relevance. First, if we take Richardson's political support function seriously, it is not clear why the FTA will be approved in the first place. The domestic price declines due to FTA from $P_w^t$ to $P_S$ so that producers will most surely resist it. Moreover, area $b - c$ is likely to be negative implying that the sum of tariff revenue, consumers' surplus and producers' surplus declines. The FTA will be resisted. Alternatively, a political support function compatible with FTA will also be compatible with an MFN tariff reduction by $P_w^t P_S$ which is precisely what Richardson obtains through the FTA plus external tariff reduction of $P_w^t + \varepsilon$.

Second, for a moment, let us make the unrealistic assumption that despite their equivalence, the FTA and MFN liberalisation up to $P_w^t P_S$ are subject to different political processes so that the former is feasible but not the latter. Even then, Richardson's argument has a very unrealistic setting in that it requires a model in which, at any time, the good in question is imported from a single source. Note that in Figure 2, all imports come from the outside country under an MFN tariff and from the partner under an FTA. If we were to assume, as in Figure 1, that the partner's supply is upward sloped permitting imports from both sources, Richardson's argument no longer holds.
Finally, it is not clear why Singapore will enter into a PTA with Indonesia in the setting shown in Figure 2. Singapore has nothing to gain from the FTA so that it will not be willing to pay the political costs of the FTA. If Singapore did have to gain as in Figure 1, it will necessarily resist outside liberalisation by Indonesia after the FTA is formed.

Empirically, so far, no evidence has been provided by advocates of PTAs supporting the point that Richardson type of considerations have led to liberalisation of tariffs on outside countries. On the contrary, as discussed later in this section, much of the evidence suggests increases in tariffs following the formation of PTAs.

(iii) Lobbying and increased barriers against outside countries

Bhagwati (1993) argues that in a political-economy setting in which producers play the central role in determining trade policies, liberalisation through FTA is likely to be replaced by increased protection against outside countries. To the extent that the country’s applied external tariff is below its GATT binding, the country can accomplish this increase in protection by increasing tariffs. And if the actual tariff corresponds to the bound tariff, the increase in protection can be accomplished through increased anti-dumping actions. This argument is demonstrated with the help of a diagrammatic technique in Bhagwati and Panagariya (1996a).

Panagariya and Findlay (1996) offer a formal general-equilibrium model in which tariffs in importable sectors are determined by firms’ lobbying activity which uses labour. There are two importables of which one comes from the potential partner and the other from the outside country. The FTA leads to an institutional change which renders lobbying for protection against imports of the partner’s good ineffective. This releases labour, lowers the wage and makes lobbying in the other import-competiting sector more profitable. The tariff on that good, imported from the outside country, rises. This rise in the tariff can make an otherwise welfare-improving FTA a welfare-reducing proposition and makes the world trading system less liberal.

Cadot, Melo and Olarreaga (1996) also consider the issue of the external tariff in a model in which trade policy is determined by the Grossman-Helpman (1995) political support function. Their analysis turns out to be more complicated and the answer ambiguous. If rules of origin are present, the FTA may result in one or both partners raising the outside tariff though such an outcome is not necessary.

(iv) A tariff-revenue objective and the outside tariff

If a country is dependent on tariffs for revenue purposes as is true of the countries in Africa, South Asia and even Central and Eastern Europe, an FTA which requires the removal of a tariff on the partner country may force it to raise the external tariff. The more the country imports from the FTA partner (i.e., more
natural is the FTA partner according to Summers), the larger the loss of revenue, the greater the increase in the external tariff required to maintain fiscal balance and the greater the trade diversion. In the same vein, after the formation of an FTA, if a country is faced with a fiscal crisis and has to resort to tariffs, since the bulk of imports coming from the union partner can no longer be in the tax base, the necessary increase in the tariff rate is likely to be much larger than in the absence of the FTA.

(v) Empirical evidence

Theoretical arguments aside, what is the empirical evidence on the impact of PTAs on outside tariffs? There is certainly a considerable evidence of increases in outside tariffs following the implementation of PTAs. After Israel concluded FTAs with both the United States and EU, tariffs on outside sources of imports – principally located in Asia – went up (Halevi and Kleiman, 1994). In the aftermath of the peso crisis in Mexico, tariffs on outside countries on 503 items went up from 20 per cent or less to 35 per cent (Bhagwati and Panagariya, 1996a). The same phenomenon has also been observed recently in the Central African Customs and Economic Union (UDEAC) which introduced an across-the-board increase in tariffs on non-members to support trade preferences among member countries. In the wake of the recent fiscal crisis in Brazil, MERCOSUR had to raise its common external tariff by three percentage points. In the EC, it has been found that internal liberalisation there was accompanied by more vigorous anti-dumping against outside countries (Hindley and Masserlin, 1993). Finally, now that Eastern and Central European countries must start implementing tariff preferences under the Association Agreements with the EU, they too are having to raise tariffs on outside imports to make up for the revenue shortfall (Masserlin, 1997).

Many advocates of PTAs argue that tariff increases in Mexico in the aftermath of the recent peso crisis were minuscule in comparison to the increases in the wake of the macroeconomic crisis there in the mid-1980s. They, thus, conclude that the effect of NAFTA was to actually contain tariff increases on outside countries which could have been much more widespread in its absence. There are three problems with this argument. First, since the early 1980s, there has been a complete reversal in the conventional wisdom on how countries should respond to balance-of-payments crises. In the past, the uniform advice, including that given by the International Monetary Fund, to countries facing balance-of-payments crises was to raise trade barriers. Today, the advice is to take the opportunity to carry out trade reforms that are difficult in times of stability. As a result, even in India, a bastion of protectionism until 1990, the balance-of-payments crisis of 1991 led to an unprecedented and unthinkable wave of trade liberalisation which has not been reversed so far and, with some pressure from the WTO, is likely to continue. India, of course, has no PTA with a powerful entity
such as the United States or EU. Second, Mexico was given a massive $40 billion debt-relief package to deal with the recent peso crisis which was not available at the time of the previous crisis. If Mexico had shown a lack of commitment on its reforms, the fate of the package would have been in grave doubt. Finally, though the reversal of trade reforms in Mexico was deeper in the mid-1980s than that following the crisis of 1994, trade liberalisation resumed soon and achieved sufficient success to impress the United States into calling for NAFTA negotiations. By contrast, since NAFTA and the crisis of 1994, Mexico has shown no significant progress towards cutting its outside tariffs.

d. The Spaghetti-Bowl Phenomenon

As discussed in Section 2c, in FTAs as opposed to CUs, member countries fear that imports from outside countries destined to a high-tariff member may enter through a low-tariff member. Or more subtly, entrepreneurs in the low-tariff country may import a product in almost finished form, add a small value to it and export it to the high tariff country free of duty. To avoid this trade deflection, FTA agreements usually include the rules of origin according to which products receive the duty-free status only if a pre-specified proportion of value added in the product originates within the union.

Even in the absence of any increase in trade barriers against non-members, the proliferation of criss-crossing FTAs leads to a replacement of the non-discriminatory MFN tariff by a spaghetti bowl whereby tariffs vary according to the ostensible origin of the product. This process is indeed at work as Figure 3, adapted from Snape (1996) and taken directly from Bhagwati, Greenaway and Panagariya (1998), illustrates. Complicated as it looks, the figure actually understates the complexity of the trade regime which has come to prevail in Europe as a result of FTAs among different countries there and in neighbouring countries in North Africa. For each FTA has its own rules of origin which vary across products and transition phase. As a result, for a given product, there are several different tariff rates depending on what origin is assigned to it. A similar picture can be drawn for Latin America.

e. WTO-Ilegal Policies in PTAs

Yet another way in which PTAs can undermine the global trading system by introducing into such arrangements measures which are otherwise WTO inconsistent. One such example is the trade-balancing requirement within MERCOSUR. The WTO has just outlawed this Trade-Related Investment Measure or TRIM. Yet it has been introduced by the members of MERCOSUR on firms operating within the union. Thus, an Argentine company operating in Brazil must export as much Brazilian goods to Argentina as it imports from the
FIGURE 3
The European Spaghetti Bowl

EFTA\(^a\)

Switzerland

Liechtenstein

Iceland
Norway

EEA\(^b\)

Israel

Cyprus
Malta

EEA\(^b\)

EU\(^c\)

Turkey

7 Mediterranean Agreements\(^d\)

70 ACP (Lomé)

Notes: Does not include countries of the former Soviet Union other than the Baltic countries

\(^a\) European Union comprises Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and the United Kingdom.

\(^b\) European Economic Area.

\(^c\) European Free Trade.

\(^d\) Algeria, Egypt, Jordan, Lebanon, Morocco, Syria and Tunisia.

EU Single Market.

\[\text{Customs Union.} \]

\[\text{Free Trade Area.} \]

\[\text{EU Association Agreements.} \]

\[\text{Non-reciprocal agreements.} \]

Source: Adapted from Snape (1966).
latter. Yet another example of this kind is appearance of a voluntary export restraint or VER on tomatoes in NAFTA. Recall that the Agreement on Safeguards in the Uruguay Round Agreement had outlawed VERs. At the moment, we do not have evidence of such WTO-inconsistent measures being widespread, nor are we clear on its legality within the WTO, but it is a problem which deserves attention.

4. OPEN REGIONALISM

In the policy debate, advocates of the current wave of PTAs also defend it on the ground that it represents ‘open regionalism’ in contrast to closed, import-substituting regionalism of the 1950s and 1960s. While references to open regionalism are repeatedly made, few attempts have been made to define the term systematically. In a recent paper, even Bergsten (1997a), the most staunch advocate of open regionalism, acknowledges:

Yet neither APEC nor any other official body has defined ‘open regionalism.’ … There is thus considerable confusion about the implications, and even the relevance, of the basic idea.

As for critics, they find the term inherently contradictory: arrangements that are open cannot be regionally confined and those that are regionally confined cannot be open. Srinivasan (1995), the most vocal critic of the idea, goes so far as to call open regionalism an oxymoron.

An early attempt at defining open regionalism was made by the Eminent Persons’ Group (EPG), appointed by APEC and headed by Fred Bergsten, in its second report (APEC, 1994). The Council of Economic Advisors to the President of the United States, from which I quote below, subsequently defined the idea in very similar terms:

Open regionalism refers to plurilateral agreements that are non-exclusive and open to new members to join. It requires first that plurilateral initiatives be fully consistent with Article XXIV of the GATT, which prohibits an increase in average external barriers. Beyond that, it requires that plurilateral agreements not constrain members from pursuing additional liberalization either with non-members on a reciprocal basis or unilaterally. Because member countries are able to choose their external tariffs unilaterally, open agreements are less likely to develop into competing bargaining blocs. Finally, open regionalism implies that plurilateral agreements both allow and encourage non-members to join (CEA, 1995, p. 220).

This definition offers three criteria for open regionalism: (i) open membership with positive encouragement to non-members to join, (ii) consistency with GATT Article XXIV, and (iii) freedom to member countries to liberalise further unilaterally or with non-members on a reciprocal basis. Though the report is not entirely explicit about it, the requirement seems to be that all three criteria be applied rather than just any one of them. Let us examine the criteria carefully, taking them in the reverse order.
The requirement that members be free to pursue unilateral or bilateral liberalisation rules out customs unions as being compatible with open regionalism even though they are perfectly compatible with GATT Article XXIV. Recall that in a customs union, individual members are not free to lower their tariffs; the common external tariff cannot be lowered unless all members are willing to do it. Nor are members permitted to conclude PTAs with outside countries on their own. Entry of new members into the customs union or a bilateral FTA has to be a joint decision. The criterion rules out the EU as pursuing open regionalism while it accepts NAFTA to be doing so despite the fact that ex post the former has signed more new PTAs than the latter. It must be abandoned as either necessary or sufficient for defining a PTA as open.

The second criterion, compatibility with Article XXIV, can serve as a necessary but not sufficient condition for an arrangement to be open. If two countries start with prohibitive tariffs and then form an FTA, keeping their prohibitive tariffs on outside countries, they will satisfy the requirements of Article XXIV. Yet, such arrangements can hardly be characterised as open regionalism. In more practical terms, this criterion says nothing about why the regionalism of the 1950s and 1960s was closed while that being pursued today is open. The criterion can be regarded as necessary but far from sufficient.

The remaining criterion, open membership, is perhaps the most important one and is, indeed, what gives the term 'open regionalism' substance. It opens the possibility that if outsiders find it attractive to seek membership, a PTA can eventually encompass the entire world and thus lead to multilateral free trade. Despite this possibility, the open-membership criterion has three important limitations which give critics reason to be sceptical of open regionalism. First, discrimination against non-members at any point in time remains in place by definition as long as the regionalism is of Article XXIV variety. Therefore, 'open' club is still likely to harm non-members. Second, openness is not as innocuous as it sounds; as Bhagwati (1995 and 1997) notes, the admission price can include several unpleasant 'side payments' that are essentially unrelated to trade. These include acceptance of a stronger intellectual-property-rights regime, investment rules, and higher labour and environmental standards. Finally, open membership does not necessarily translate into speedy membership. It has taken the EU more than 40 years to grow from 6 members to 15. The Canada-US Free Trade Agreement was concluded almost a decade ago and, taking into account NAFTA, its membership has grown to only three so far. Attempts by even a tiny country such as Chile have faced serious resistance. And it will be a long time, if ever, before countries in South Asia and Africa are admitted to either the EU or NAFTA. Until that happens, the world trading system can become fragmented with complex rules of origin and tariff phase-outs contributing to the spaghetti-bowl phenomenon.

It is this fear of fragmentation which prompted Mr. Renato Ruggeiro, the Director General of the WTO, to go a step beyond the CEA (1995) in defining
open regionalism. In WTO (1996), he contrasted two interpretations of open regionalism. The first interpretation stops at consistency with Article XXIV of GATT 1994 and the understanding on its interpretation incorporated in the Uruguay Round agreements on Trade in Goods. In the second interpretation:

... the gradual elimination of internal barriers to trade within a regional grouping will be implemented at more or less the same rate and on the same timetable as the lowering of barriers towards non-members. This would mean that regional liberalization would in practice as well as in law be generally consistent with the m.f.n. principle (WTO, 1996, p. 11).

Interestingly, this was the option considered in detail and preferred by the author in the context of regionalism in East Asia (Panagariya, 1993 and 1994). The Director General expressed a clear preference for this option when he concluded:

The choice between these alternatives is a critical one; they point to very different outcomes. In the first case, the point at which we would arrive in no more than 20 to 25 years would be a division of the trading world into two or three intercontinental preferential areas, each with its own rules and with free trade inside the area, but with external barriers still existing among the blocs (ibid., p. 11).

In my judgement, if regionalism is to be truly open, it cannot escape being non-discriminatory. As we have seen, Article-XXIV consistency is a necessary but far from sufficient criterion for ensuring the openness of the world trading system. Open membership can, in principle, lead to global free trade but will lead to fragmentation during transition which will be long and, in practice, may even dead-end at two or three large blocs. Adherence to the MFN principle, by contrast, will eliminate the need for any rules of origin and prevent the growth of the spaghetti-bowl phenomenon.

Some critics find the idea of adherence to the MFN principle to be incompatible with regionalism, however. Thus, Srinivasan (1997) retorts:

if regional liberalization is to be extended on the same time table 'in practice and in law' to non-member countries on an MFN basis, it would be multilateral and not regional. If that is the case, why would any group initiate it on a regional basis in the first place?

There is one other definition of open regionalism which deserves consideration. Some observers have noted that a distinguishing feature of current regional arrangements is that they are taking place in an environment in which trade barriers have been brought down to low levels through eight rounds of multilateral trade negotiations. Therefore, open regionalism can be defined as FTAs and customs unions with low trade barriers on outside countries. While this definition distinguishes recent regionalism from earlier ones, it does not overcome the fundamental contradiction between openness and discrimination: outsiders may face low barriers but they are nonetheless subject to discrimination relative to insiders.

But suppose for a moment we accept the view that it is possible to discriminate, albeit with low barriers, and still be open. But what is the level of
discrimination at which closed regionalism turns into open regionalism? This is not idle discourse. When Mexico maintains a large number of external tariffs in the 15–20 per cent range, keeps open the option of raising them up to 35 per cent, and also has the instrument of anti-dumping at its disposal, can we say that it is pursuing open regionalism under NAFTA? Or should we adopt external trade barriers in the United States as the critical level which distinguishes open from closed regionalism? But this latter definition is also problematic: even after the Uruguay Round liberalisation, more than half of textiles and clothing products will remain subject to 15–35 per cent tariffs in the United States and anti-dumping will remain a major trade barrier. To put the matter differently, despite lower (though by no means low) external trade barriers today, the motivating force behind regional arrangements is no different than in the 1950s and 1960s. Now, as then, discrimination is the name of the game as member countries continue to be driven by a desire to secure a preferential access to the partner country’s market. This is clearly the impression conveyed, for example, in the recent report by the United States Trade Representative, ‘Operation and Effect of NAFTA.’

Since the term open regionalism has been used most frequently to describe the developments in Asia, our discussion will be incomplete without brief comments on regionalism in that region. Asian regionalism has been truly ‘open’ in that liberalisation undertaken by members of Asian regional groups has been largely non-discriminatory. APEC has created no trade preferences and the members of ASEAN, despite having signed the ASEAN Free Trade Area (AFTA) agreement which is a PTA in principle, have undertaken much of their liberalization on a non-discriminatory basis. Thus, if one was looking for an example of Director General Ruggeiro’s preferred form of regionalism, Asia provides it. Regional arrangements in North America, Latin America or Europe do not meet the standards of openness laid out in the Director General’s definition.

5. ‘DEEP’ INTEGRATION

Lawrence (1997) has advocated regional arrangements by arguing that, unlike multilateral trade liberalisation, they promote ‘deeper’ integration. Rather than being confined to ‘shallow’ integration in terms of liberalisation of trade among members, they involve ‘deep’ integration through coordination, if not complete

---

15 As discussed in Panagariya (1997a), the United States certainly sees APEC as an instrument for seeking increased access for the US goods and has refused to give tariff or other trade concessions under the APEC auspices except on a discriminatory basis.

© Blackwell Publishers Ltd 1999
harmonisation, of other policies including competition policies, product standards, regulatory regimes, investment codes, environmental policies, labour standards and so on. He argues that such deep integration may confer gains on member countries by lowering the costs of production and improving efficiency in general.

The phrase ‘deep integration’ conveys the impression that whatever it stands for must be a good thing – both deep and integration are ‘good’ words. But there are reasons to be sceptical of such a conclusion. First, a substantial part of the deep integration agenda is itself undesirable from the viewpoint of efficiency and welfare. The careful analysis in Bhagwati (1996b) and Wilson (1996) shows that there are good reasons for diversity in domestic polices across nations and that harmonisation is, in general, not a welfare-enhancing proposition. Optimal pollution and labour standards depend on income levels. Developed countries may be more concerned about air pollution while developing countries may be worried about water pollution. Likewise, optimal standards with respect to minimum wage, worker-safety, and child labour depend on income levels.

Second, in regional arrangements between countries with uneven bargaining power, smaller, developing countries fear that deep integration can become an instrument for extracting concessions of all kinds not just in trade but in other ‘non-trade’ matters by their larger, more powerful counterparts. The agenda for deep integration is likely to be determined by rich, developed countries. And it is the smaller, developing countries who will have to adjust their standards to those of developed countries, regardless of whether these are appropriate to their conditions. As it happens, the non-trade agenda for future regional arrangements has been dominated in the United States by demands for higher labour and environmental standards in potential partners. Whatever their social merit, deep integration along these dimensions is unlikely to be a welfare-enhancing move for countries integrating with the United States.

According to the available evidence, when deep integration is attempted among entities with uneven bargaining power, penetration may end up being one way. Lawrence (1997) offers NAFTA as an example of deep integration. Yet, all said and done, it offers increased scope for the penetration of Mexican

---

16 This sentiment was expressed in slightly different form by Whalley (1993) when he wrote, 'The current danger for Mexico in NAFTA is that the price will be paid in the form of adverse exclusionary arrangements which are against Mexican national interest, particularly autos, with potential impediments to third-country inward capital flows and arrangements which effectively reserve the Mexican import market for US suppliers.'

17 Bhagwati (1994) makes the further argument that a hegemonic power is likely to gain a greater payoff by bargaining sequentially with a group of non-hegemonic powers than simultaneously. He cites provisions with respect to intellectual property protection and environmental and labour standards as extra benefits secured by the United States through the uneven bargain. Mexico was unable to obtain similar benefits in return.
markets by the United States but not the other way around. Fred Bergsten who championed the cause of NAFTA, has himself gone on record stating that under NAFTA the US made no concessions to Mexico while she got every concession she sought.\textsuperscript{18} The rules of origin including triple transformation rules in textiles and clothing, side agreements on orange juice and sugar, and subsequent restrictions on Mexican tomatoes bear testimony to the fact that NAFTA generally resisted deep penetration of the US markets by Mexico.

Finally, even if we are able to identify dimensions along which deep integration is desirable, it does not follow that a PTA is a necessary complement to it. In principle, much of deep integration agenda can be pursued independently of a PTA. To justify PTA, one must identify extra gains resulting from a simultaneous pursuit of PTA and other deep integration agenda. Short of that, the two policies must be justified on their own merit. This the proponents of deep integration have not done.

6. CONCLUSIONS: MINIMISING THE ADVERSE EFFECTS OF PTAS

Rather than try to summarise, I conclude the paper with suggestions that have been made to minimise the adverse effects of PTAs. Because the enthusiasm for PTAs remains strong, this is a useful exercise. Many proposals have been made of which I list the more important ones here.\textsuperscript{19}

First and foremost, we can consider placing a moratorium on the expansion of PTAs beyond those already in an advanced stage of negotiation, particularly the big FTAs such as the Transatlantic Free Trade Area and APEC. At the same time, we should move speedily to bring the unfinished agenda of multilateral trade liberalisation to its logical conclusion. Such liberalisation will not only promote the cause of multilateral liberalisation directly, it will also do so indirectly by neutralising the impact of trade preferences within PTAs. In the limit, if all external barriers drop to zero, PTAs will be effectively eliminated.

Second, for future FTAs, we should modify the GATT Article XXIV so as to require each FTA member to bind its tariffs to the actual levels prevailing at the time of negotiations for the FTA. This will ensure that tariff protection against non-members is not raised in the future. For example, if this provision had been present, Israel and Mexico would not have been able to raise their external tariffs following FTAs with the United States. In the case of customs unions, as Bhagwati (1991) suggested, it may be required that the common external tariff

\textsuperscript{18} To quote him (Bergsten, 1997b, p. 26), 'NAFTA amounted to a 4\% expansion of the American economy, to include a country that accepted virtually every demand placed upon it in the negotiations and which made virtually all the concessions.'

\textsuperscript{19} Serra et al. (1997) and Panagariya and Srinivasan (1998) provide more details in this context.
be set at the minimum of the pre-union import tariffs of the member countries.\footnote{In an early contribution to the debate on how GATT rules should be modified to minimize trade diversion, McMillan (1993) had suggested that member countries of PTAs be required to ensure through appropriate policies that their trade with outside countries does not decline. The logic behind this suggestion is that current rules that FTA members not be permitted to raise their external tariffs and a customs union adopt a common external tariff no higher than the average of the member countries’ tariffs leaves considerable room for a reduction in the union’s trade with outside countries. A difficulty with McMillan’s proposal, however, is that in a world of growing incomes, trade diversion is consistent with growing trade with outside countries. Moreover, it substitutes an outcome variable for an instrument variable which is harder to predict ex ante.}

Third, since anti-dumping and safeguard measures can be substituted for higher tariffs on outside countries by enforcing these measures more aggressively against the latter, we also need to introduce changes in Article XXIV which will prevent such a shift in policy.

Fourth, to minimise trade diversion, we should require that within a specified time period, say, five to ten years, following the grant of a tariff preference to an FTA member, the external tariff should be reduced to the level of the preferential tariff as well. This means that countries forming a full FTA will have to go to free trade with the rest of the world as well within the specified time limit.\footnote{See Srinivasan (1996) for this suggestion. The suggestion is a modification of the proposal by Panagariya (1994) that in pursuing regional trade liberalisation, the countries in East Asia should proceed in a non-discriminatory manner. He suggested that the countries could negotiate reductions in trade barriers regionally and implement them on a non-discriminatory basis.}

Finally, to contain the spaghetti-bowl phenomenon, we could require that there be no rule of origin on a product in a member country with the lowest tariff in the union on that product. Such a requirement would have eliminated the need for the rules of origin in the United States in a large majority of products. We could also require that whenever the difference between the highest and lowest external tariff rates in the union is less than a prespecified limit, say, five per cent, no rules of origin are permitted.\footnote{See Wonnacott (1996) for proposals along these lines.} With multilateral liberalisation moving ahead, this requirement could help kill the rules of origin on many products in many FTAs faster than otherwise.

An advantage of the proposals on binding the external tariff and then bringing it down to the levels of the preferential tariff is that it will help arrest the proliferation of trade-diverting PTAs. For in that case only the countries truly interested in eventual multilateral free trade will muster the courage to enter into FTAs.
REFERENCES


Panagariya, A. (1993), 'Should East Asia Go Regional? No, No and Maybe,' WPS 1209 (World Bank, October).


THE REGIONALISM DEBATE: AN OVERVIEW

Srinivasan, T.N. (1995), 'APEC and Open Regionalism' (Yale University, processed).