Trading Preferentially: Theory and Policy

Jagdish Bhagwati; David Greenaway; Arvind Panagariya


Stable URL:
http://links.jstor.org/sici?sici=0013-0133%28199807%29108%3A449%3C1128%3ATPTAP%3E2.0.CO%3B2-D

*The Economic Journal* is currently published by Royal Economic Society.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/res.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
TRADING PREFERENTIALLY: THEORY AND POLICY

Jagdish Bhagwati, David Greenaway and Arvind Panagariya

The best kind of economic theory has almost always reflected policy concerns, while informing policy in turn. This is particularly so when it comes to the theory of international trade, going back to Adam Smith’s discovery of the demerits of mercantilism and his invention of economic science, both in The Wealth of Nations. The theory of preferential trading is no exception.

The original burst of creative theorising about Preferential Trade Agreements (PTAs), associated with what Bhagwati (1991) has called the First Regionalism, is well known to have come from Jacob Viner’s (1950) work on what he called the ‘customs union issue’ and was a result of his having been commissioned by the Carnegie Endowment to examine the appropriate design of the world trading system with the end of the Second World War. In turn, the impending formation of the Common Market, leading to the Treaty of Rome in 1957, played a role in the further development of the theory at the hands of James Meade (1955), Richard Lipsey (1958) and others.

The recent burst of theorising about PTAs is also a reflection of the new policy questions raised by the fact that the United States abandoned in the early 1980s its policy of avoiding PTAs, even though sanctioned by Article 24 of the GATT, and concentrating exclusively on multilateral trade negotiations

* Thanks are due to Kym Anderson, Kyle Bagwell, Robert Baldwin, Richard Baldwin, Claude Barfield, Christopher Bliss, Magnus Blomström, Max Corden, Clive Crook, Don Davis, Alan Deardorff, Vivek Dehejia, Elias Dinopoulos, Henry Flam, Robert Feenstra, Caroline Freund, Ross Garnaut, Earl Grinols, Gene Grossman, Elhanan Helpman, Gary Hubbauer, Douglas Irwin, Pravin Krishna, Paul Krugman, Anne Krueger, Patrick Lane, Phil Levy, Rachel McCulloch, Chris Milner, John Nash, Andre Sapir, Gary Saxonhouse, Alasdair Smith, Robert Staiger, Robert Stern, Sethapat Suthiwartnarueput, T. N. Srinivasan, Wendy Takacs, David Tarr, John Whalley, Alan Winters, Martin Wolf, Kay-yiu Wong, Ron and Paul Wonnacott, and Ian Wootton for conversations that have contributed to the development of our thinking over the last decade on the issues covered in this paper. Some of these economists have their own take on the rapidly expanding theory of PTAs; e.g. Winters (1996), and Baldwin and Venables (1996). A set of key contributions to the theory of PTAs, starting from Viner (1950), organised by major analytical approaches to the subject, is given in Bhagwati et al. (1998). Ajay Panagariya drew the figures skillfully for us.

1 The late 1950s and early 1960s, which constitute the period of the First Regionalism, were witness to attempts at forming customs unions (CUs) and free trade areas (FTAs) in several developing countries, inspired partly by the Treaty of Rome but reflecting a different rationale which is discussed below in Section 1.4. It should be stated that the general usage of the word Regionalism interchangeably with PTAs or FTAs is not desirable since PTAs are occasionally formed with members who do not meet any reasonable definition of a ‘region’: e.g. United States and Israel. For a discussion of this relationship, or rather lack of it, see Bhagwati and Panagariya (1996a, pp. 4 and 31–5) and the spoof, ‘The Watering of Trade’ by Bhagwati (1996).

2 Other non-Vinerian developments of the ‘static’ approach to the analysis of PTAs, pioneered by Viner, are discussed in the text below. These too have been interactive with policy questions, as we argue in the text.
(MTN) which lead to reductions of trade barriers characterised by MFN. The subsequent proliferation of PTAs has made this Second Regionalism a period of 'success' whereas the First Regionalism was generally marked by aborted efforts.

Whereas the Vinerian analysis of PTAs in the First Regionalism reflected 'static' questions concerning the welfare effects of unions with defined membership, the Second Regionalism has been preoccupied with what Bhagwati (1993) has described as the 'dynamic' time-path question: i.e., in broad terms (to be refined below), whether PTAs can provide an impetus to, or whether they will detract from, the worldwide nondiscriminatory freeing of trade. In other words, in the phrasing and conceptualisation of Bhagwati (1991), will PTAs be 'stumbling blocks' or 'building blocks' in the freeing of trade multilaterally?

The very success of the Second Regionalism, and the continuing proliferation of PTAs which we confront today even as their number is already in three digits (Bhagwati and Panagariya, 1996a, Chapter 1, Appendix), has led to a substantial shift away from the complacency among international economists that attended this proliferation. Not merely have the old concerns about trade diversion revived, with several studies now challenging the assertions that trade diversion is not an important issue in practice, but the systemic effects of having numerous PTAs present in the world trading system have also attracted analytical and policy attention.

In the following review of the present state of the theory of PTAs, we consider in Section 1 the 'static' theories, in Viner's tradition and outside of it; in Section 2 we address the 'systemic' issues which have been raised in the static framework; in Section 3, the 'dynamic' time-path theories are reviewed; and in Section 4, we consider briefly the state of the current policy debate.

1. 'Static' Theories: Viner and Others

There are by now four alternative theoretical approaches to the 'static' implications of PTAs, with recent theoretical developments of interest in most of them.

1.1. Viner and Extensions: Trade Diversion and the 'Natural Trading Partners' Issue

Two issues in particular can be highlighted in the modern extensions of Viner's analysis of the 'static' welfare implications of PTAs.

(i) Is Trade Diversion a Red Herring? Viner distinguished between trade creation and trade diversion. The notion that any move towards free trade, even if preferential, would necessarily be welfare-improving, appeared intuitive until Viner's seminal work. Today, these pre-Vinerian attitudes have reappeared, as in Lawrence Summers' (1991, page 299) remarkable statement: 'I find it surprising that this issue [of trade diversion] is taken so seriously—in most other situations, economists laugh off second best considerations and

© Royal Economic Society 1998
focus on direct impacts. But the pro-PTA implication is not sustainable for at least three reasons.

First, from an analytical viewpoint, Panagariya (1995; 1996) has shown that, if the effect of a PTA on a member country’s welfare is to be examined, the transfer of revenue among the members following the abolition of tariffs within the PTA will generally lead to tariff revenue redistribution among the members. This ‘Panagariya rectangle’, measuring the loss of tariff revenue on inter-member trade, can outweigh any net gain on the Harberger-Johnson ‘triangles’ that trade creation and trade diversion imply. This point is of relevance when a high-tariff country like Mexico joins a low-tariff country like the United States in a PTA.4

Second, increasing numbers of empirical studies are now beginning to show that trade diversion is not necessarily a negligible phenomenon in current PTAs. Thus, Yeats’ (1996) study of MERCOSUR (among Uruguay, Argentina, Paraguay and Brazil) turned up significant evidence of trade diversion; the diversion of textile trade to Mexico from the Caribbean, thanks to NAFTA which excludes the latter and includes the former, has been a source of discord; and Wei and Frankel (1996) have argued that the EU has led to trade diversion. We are sure that more evidence of trade diversion will mount as economists, no longer under political pressure to swim with the tide in favour of PTAs, return to their first economic principles and begin to distinguish à la Viner between preferential and nonpreferential trade liberalisation.

In fact, now that we have interesting analytical political-economy-theoretic papers by Grossman and Helpman (1995) and Krishna (1996) which show that trade diversion is an important motive (among others noted by us in Section 4) leading to PTAs, it is not surprising that the evidence of trade diversion occurring as a result of PTAs is not hard to find. It is not remarkable that those who indulge their preference for preferential trade arrangements are seen also to use the preferences.

Third, the complacency about PTAs not leading to trade diversion was aided by the notion that the external trade barriers were no longer very high and that, therefore, preferences could not lead to significant trade diversion. But this is not true for several reasons.

To begin with, even trade tariffs are still very high, both in developing and in developed countries. In the latter, the Uruguay Round has still left several peak tariffs in specific products whereas the tariffication of agricultural support has created truly substantial tariffs. In the former, countries in South Asia and in Latin America are also not free from high trade barriers, making PTAs particularly dangerous.

---

3 The several analytical fallacies in this statement have been noted in Bhagwati and Panagariya (1996a, pp. 5–7). The fact that empirically trade diversion is not negligible and cannot be dismissed is yet another criticism.

4 On the importance of analysing revenue effects in undertaking trade policy changes more generally, see Greenaway and Milner (1991). Greenaway (1993) considers other pertinent issues as well.

© Royal Economic Society 1998
Besides, the external trade barriers are today only a part of the protectionist story. 'Administered protection', the term we owe to Michael Finger, consisting of instruments such as anti-dumping (AD) actions, has become the favoured policy of protectionists who cleverly use the appealing notion of 'fair trade' to unfairly gain protection and advantage against successful foreign rivals. But then these instruments typically yield protection that is elastic and selective. Thus, AD duties, which bear little relationship to 'predation' in the economic sense and hence have in practice no economic justification, are often based on adjusted prices that are estimated in ways calculated to find dumping or on essentially arbitrary 'reconstructed cost' when the AD methodology compares, not foreign and domestic prices, but foreign costs and domestic prices. Thus, within broad margins, it is arguable that AD calculations and actions will seek to accommodate the needs of the protectionist petitioners in the spirit of the story where the interviewing commissar asks candidates what the sum of 2 and 2 is, and the job goes to the candidate who answers: whatever you want, sir. In addition to their being therefore elastic, AD actions are selective in the sense that they are mounted against specific countries and even specific firms within those countries. Thus, it is possible to use them to zero in against your most potent foreign rivals.

It follows then that, unless such administered protection is severely regulated, the temptation on the part of PTA members will be to protect each other with such protection at the expense of nonmembers, when internal competition among members breaks out with the PTA formation. In short, protection against nonmembers then becomes endogenous to the PTA. The consequence is that, as trade creation occurs within a PTA, the endogenous raising of protection converts it into trade diversion instead. E.g. as Mexico starts crowding out inefficient US producers, the United States accommodates imports from Mexico by reducing imports from the most efficient nonmember supplier, Taiwan, using AD actions against Taiwan. This possibility, noted in Bhagwati (1993, pp. 36–7), has been formally demonstrated in Bhagwati and Panagariya (1996a, pp. 38–41), and Panagariya and Findlay (1996).

Such a phenomenon is not an idle theoretic speculation. Instances of such endogenous raising of protection against nonmembers have been observed. An example is the raising of (unbound) tariffs on over 500 non-NAFTA tariffs by Mexico during the 1995 peso crisis, while tariffs against NAFTA members were not. EU experience has been discussed by Hindley and Messerlin in their contribution to Anderson and Blackhurst (1993). Hence, Bhagwati (1993) and Serra et al. (1996) have argued that the reform of GATT Article 24 on PTAs is not enough; enhanced discipline on administered protection,

5 An excellent review of the theoretical and policy literature on predatory dumping and the lack of its relationship to AD practice in trade law is to be found in Clarida (1996).

6 See Hindley (1997) for current EU practice; he has been a longstanding critic of the intricate calculating methodology that stacks the cards against foreign suppliers.

© Royal Economic Society 1998
especially AD actions, is equally necessary to reduce the damage from PTAs today.

(ii) The ‘Natural Trading Partners’ Issue But if the issue of trade diversion has been a matter of controversy among trade theorists and policymakers, a related issue has also attracted attention, again with claims made by prominent economists and being rebutted by serious new analytical research. The claim has been that if PTAs are formed among ‘natural trading partners’, then they can be expected to be welfare-improving for the members. The phrase ‘natural trading partners’ comes from Wonnacott and Lutz (1989, page 69) who used it in several different senses, of which two have been put into circulation by their endorsement by Summers (1991) and Krugman (1991): that the initial volume of trade among the partners is high, and that the distance between them is low. But, as Bhagwati and Panagariya (1996a) and Panagariya (1997a,b) have demonstrated, each criterion is flawed and should be rejected.

(a) The Volume-of-Trade Criterion: Of course, this criterion is neither symmetric nor transitive. The United States is Mexico’s largest trading partner, but the reverse is not true. Then, the United States is the largest trading partner of Canada and Mexico, but the latter have little trade with each other. But even waiving these objections, there are difficulties with the Summers-Krugman assertion.

The presumption that high initial volume of trade among members reduces the potential for trade diversion is presumably based on the view that if there is little trade with nonmembers, the scope for trade diversion is reduced. But what we need to know is the likelihood of trade diversion instead. The scope for trade diversion in an instance may be twice as high as in another; but the actual trade diversion that occurs may be only half as much.

The actual trade diversion will reflect the underlying fundamentals such as elasticities of substitution among products, not the average initial trade volumes! This was understood, of course, by the early theorists of PTAs; see the discussion of Meade (1955) and Lipsey (1958) in Bhagwati (1993) and Panagariya’s (1997a) analysis of the classic Meade (1955) model.

In fact, as Bhagwati and Panagariya (1996a, pp. 22–7) show, analysing a model based on Meade (1955) with each member country specialised on a different product when all products are imperfect substitutes, the steady reduction of tariff preferentially by one country on another will first improve its welfare and then progressively reduce it at some stage, so that reaching a 100% reduction, implying an FTA, may reduce welfare even below the starting level. This is shown in Fig. 1 where the superscripts $b$ and $c$ refer to countries $B$ and $C$; the reduction of $A$’s tariff on $B$ preferentially, maintaining the tariff on nonmember $C$ throughout, is plotted; and so is the welfare of home country $A$. In this diagram, as the tariff on partner country $B$ reduces steadily to the left, the volume of trade by $A$ with $B$ also grows absolutely and relatively to trade with nonmember country $C$, turning $A$ and $B$ progressively into more and more ‘natural trading partners’. But, any further such increase in ‘naturalness’ beyond the point at which welfare of $A$ begins to fall makes any further preferential tariff reduction in favour of
$B$ harmful, not beneficial: so, the more natural trading partner you get to be, the more likely it is that you will be in the zone where preferential trading will reduce welfare.\footnote{That a 100\% preferential tariff reduction à la Article 24 could be harmful whereas a lesser preference could be beneficial has been known to readers of Lipsey (1958) for over four decades. But it does not follow that we should therefore weaken Article 24 to allow all sorts of preferential reductions provided they are calculated to produce greater welfare. That question raises a number of other issues, such as the wisdom of having PTAs made admissible on the basis of such calculations. For a discussion of the economics and the political economy of Article 24 requirement that only 100\% preferential reduction of inter-member trade barriers is to be allowed, see Bhagwati (1991, 1993).}

Yet another implication of the demonstration in Figure 1 is that there need be nothing ‘natural’ about initial high volumes of trade. These may reflect earlier preferences, leading the proponents of the natural trading partners

\footnote{© Royal Economic Society 1998}
argument in effect to advocate more preferences on the basis of existing preferences! This possibility is not a theoretical curiosum: the volume of trade between the United States and Mexico has certainly increased thanks to the Offshore Assembly Provision which has differentially assisted Mexican production and exports to the United States; similarly, the trade with Canada has profited considerably from the auto agreement which established free trade in the sector between the two countries but not for producers elsewhere.8

(b) The Transport-cost Criterion: If the volume-of-trade criterion fails to work, so does the transport-cost criterion, whose appeal is intuitive to many but still rests on error. Of course, any freeing of trade can be frustrated if high transport costs—in the limit, physical inability to transport goods—prevent trade opportunities so opened up from being utilised. But this does not translate into the current prescription concerning desirable PTAs.

To see this, it is enough to see the counterexample constructed in Bhagwati and Panagariya (1996a, pp. 36–8).9 Thus, consider Fig. 2 and three countries, A, B and C. Country A has the option to join a PTA with either B or C. Countries B and C are identical except for the fact that B is less distant. Three panels are drawn. The first two panels show the export supply curves of C and B, whereas the third panel shows the combined supply curve. Assuming a fixed transport cost per unit, country C’s supply curve (c.i.f.) is higher than country B’s since B is closer. The third panel shows the trade equilibrium when both B and C pay identical tariffs in A; A’s demand curve is not shown. Note that the initial volume of imports from the proximate country B is larger than from the distant country C.

Now, by drawing the marginal cost curves derived from the export supply curves, we see immediately the paradoxical conclusion: at the initial non-discriminatory equilibrium, the marginal cost of imports from the more distant country C is lower than that from B which is the proximate country and hence is the ‘natural trading partner’ by the transport-cost criterion. It follows, of course, that the efficient choice of the PTA partner would be the distant country C.

It is easy to see why. Country A is best seen as a discriminating monopsonist. So, for any quantity of total purchases, the supplier with the higher elasticity should be paid the higher price. Here, this translates into a lower duty on the supplier with the higher elasticity. Transport costs make the distant country C’s supply curve more elastic than that of the proximate country B, making the PTA with C more attractive for country A.10

8 Among other criticisms, we must include also the ‘Panagariya rectangle’ concerning the transfer of revenue among member countries, which we noted in the text earlier. This adverse effect will be the larger, the larger is the volume of initial trade with the member country whose exports to oneself are freed from duties, turning the ‘natural trading partners’ presumption of no damage into just the opposite presumption of larger damage.

9 In the following, we omit some of the more subtle points that these authors consider, such as what would happen if the supply curves were perfectly elastic.

10 Since, in this example, the proximate country is also the one with the higher initial volume of imports, this example serves also as a demonstration of the unreliability of the volume-of-trade criterion as a guide to selecting PTA partners.

© Royal Economic Society 1998
It is fair then to say that these analyses have laid to rest the 'natural trading partners' hypotheses, whether of the volume-of-trade or of the transport-cost variety, that became fashionable when prominent economists such as Summers (1991) and Krugman (1991) embraced them, and they gained the attention of policymakers and were reported in influential magazines such as *The Economist* (1991).\(^{11}\)

1.2. *The Kemp-Wan Approach: Parting with Viner's Formulation*

The Vinerian attack on the presumption that preferential trade liberalisation is beneficial is clearly then of great importance. But there is an altogether different question that Kemp and Wan (1976) addressed and answered completely: suppose that any arbitrarily-specified subset of countries were to form a CU, and could choose their external tariff (so that the external tariff

\(^{11}\) *The Economist* ran an Economics Focus Column on the subject, as a report on the 1991 Jackson Hole Conference where the Summers and Krugman papers were presented.

© Royal Economic Society 1998
became endogenous whereas the Vinerian analysis took it as given), could they always devise a CU which left the welfare of the nonmembers where it was while it improved member country welfare?

Kemp and Wan showed that they could. The essence of their argument was beautifully concise.\(^{12}\) If the external trade vector of the arbitrarily chosen members of a CU was frozen, the nonmembers' welfare would be frozen too. If then the inter-member barriers were fully removed, the resulting competitive equilibrium would be Pareto-superior and hence, with lumpsum transfers, each member country could be made better off.\(^{13}\) The difference between the domestic price vector that emerged in this CU equilibrium, when set against the frozen external price vector, would then yield the endogenously-determined trade tariff and subsidy vector that would support the necessarily welfare-enhancing Kemp-Wan CU.\(^{14}\)

Kemp and Wan would appear to be restoring the preVinerian intuition about PTAs. In fact, however, the preVinerian intuition is, strictly speaking, not restored by Kemp and Wan, since that intuition was that all PTAs (and indeed all kinds of trade liberalisation short of free trade), were desirable because they all moved trade barriers down, i.e. that welfare of liberalising countries and world welfare would improve monotonically as trade barriers came down, no matter how they came down. Rather, Kemp and Wan show that a particular PTA which reduces trade barriers among member states, could always be crafted which would improve the welfare of member countries and of the world. In essence, therefore, the Kemp-Wan argument is an existence argument. In itself, it does not provide a clue to the structure of the endogenous tariff that would emerge in the Kemp-Wan welfare-improving CU.\(^{15}\)

1.3. The Brecher-Bhagwati Approach: Distributional Effects within a Common Market

Suppose, however, that a Common Market (variety of PTA) exists, such as the European Union 'core', where there is also free mobility of factors of production within the PTA. In that case, the analyst can ask a variety of important policy questions: e.g. what happens to the welfare of each 'national' set of factors, defining a member country, as the common external tariff changes, or

\(^{12}\) The Kemp-Wan paper was preceded by a fragment of an argument by Kemp earlier and more systematic attack on the problem by Ohyaama (1972) in an obscure Japanese journal. Baldwin and Venables (1996) claim that Meade anticipated the Kemp-Wan argument; but this claim cannot be sustained, as shown in Panagariya (1997a). Yet another way of arriving at the Kemp-Wan existence solution is provided in Krishna and Bhagwati (1996).

\(^{13}\) The internal barriers must be removed altogether, so that we can invoke without difficulty the welfare theorem that a competitive equilibrium is a Pareto-efficient equilibrium.

\(^{14}\) Although the Kemp-Wan proof relates to the formation of a CU, with a common external tariff, we suspect that the proof can be extended to the formation of an FTA with each country having its external tariff; in that case, the proof must have as a building block the freezing, not just of the aggregate external trade of all members together, but of each country's external trade vector.

\(^{15}\) On the other hand, the Kemp-Wan tariff structure has been explored by Grinols (1981), Bliss (1994) and Srinivasan (1995).
as capital accumulates in one country or technical change occurs in another, or as new members of varying sizes relative to the average size in the PTA are added? In essence, this is identical to the income distributional analysis within a single nation (wherein we conventionally assume factors to be mobile), where the focus is not on overall welfare of the nation but on the welfare of individual factors by class: e.g. by race, by gender, by ownership of assets etc., when fundamentals such as capital and technology are assumed to change and policies such as trade policy are changed.

Bhagwati and Tironi (1980), reacting to the policy concerns in South America that freeing of trade would benefit the foreign multinationals operating in South America while immiserising the host countries, had already opened up the analysis of the effect of change in trade policy on national welfare when foreign factors of production are present in the economy, so that 'national' welfare (defined over national factors of production, and hence not including the foreign-owned factors) must be distinguished from 'domestic' welfare (which does not do so). But their analysis was limited to the freeing of trade and foreign investment issue and the broad implication regarding the analysis of PTAs in shape of Common Markets was also not seen.

Brecher and Bhagwati (1981) generalised the analysis to a formal treatment of the income distributional issues when welfare analysis must be disaggregated down to classes of factors, analysing the effects of parametric and policy changes on such groups of factors, while explicitly stating the applicability of the analysis to changes in national welfare in a Common Market of the EU variety. Indeed, the effect on individual member countries of liberalisation of the common EU tariff was precisely the problem that the Brecher-Bhagwati analysis equipped one to analyse. More recently, Casella (1996) has analysed a similar problem, differentiating among different member countries by their size, in the framework of the theory of imperfect competition with scale economies.

1.4. The Cooper-Massell-Johnson-Bhagwati Approach: Forming a CU to Minimise the Cost of Import Substitution

The policy problem for developing countries in the 1960s, as they confronted the high cost of each country's import substitution strategy, was: given any level of import substitution vis-à-vis the developed countries, could the developing countries open up trade preferentially among themselves and reduce the cost of their individual import substitution? By invoking scale economies, Cooper and Massell (1965a,b), Johnson (1965) and Bhagwati (1968) argued that they could. In fact, attempts were made at forming such PTAs in East Africa and in Latin America. These attempts, part of the aborted First Regionalism, failed because these countries were wedded to planning at the time and saw trade as accommodating to a planned allocation of the import-substituting industries among the member countries, instead of letting trade decide which industry went where, thus putting the cart before the horse and killing the forward momentum.

© Royal Economic Society 1998
But, while scale economies were invoked to advance the theoretical underpinning of such an approach to forming PTAs, recently Krishna and Bhagwati (1996) have shown that the Cooper-Massell-Johnson-Bhagwati argument can be formalised simply without them. In essence, they show that the solution involves, as is intuitive once you think of it, a Kemp-Wan CU along with production tax-cum-subsidies to achieve the import-substituting industrialisation objectives of member states as indicated by the theory of optimal intervention in the presence of noneconomic objectives.

This approach is now principally of a historical-explanatory value since nearly all developing countries are convinced today of the pitfalls associated with the import substitution strategy; they have also seen that export oriented strategies have led to far more rapid industrialisation in the Far East and elsewhere. The renewed preference for PTAs among several developing countries is now premised on quite different grounds, many political rather than economic (as argued in the final section below).

2. The Systemic Issues Raised by Proliferation of PTAs: The ‘Spaghetti Bowl’ Phenomenon

The static welfare effects of PTAs would be seriously incomplete today if it were not extended to their ‘systemic’ implications because of the massive proliferation of PTAs, ridiculed in Bhagwati (1997), in the period of Second Regionalism.

These systemic implications arise because PTAs magnify the problems that arise in essence because we try to restrict or liberalise trade on the basis of which product comes from which country, or what Bhagwati (1995) has called the ‘who is whose’ problem (in Bhagwati and Krueger (1995)). Thus, for instance, as soon as the United States wishes to liberalise the imports from Israel preferentially, it must decide whether an import coming from Israel is Israeli. I.e. it must establish a ‘rule of origin’ which usually takes the form of some sort of ‘content’ rule such that a product is considered to be Israeli only if its Israeli content exceeds an arbitrarily specified share in gross value.\(^\text{16}\) The arbitrariness of this share specification is further compounded by the arbitrariness inherent in computing such content. Thus, consider the import of ingots of steel which, in conjunction with homogeneous domestically-produced ingots, go into producing scissors and forks. How is one to determine which of these two products got what share of the imported as against the domestic ingots?\(^\text{17}\) Again, if forks need to be coated with plastics, we know that even if the plastics are immediately produced at home, their gross value (as in Leontief’s analysis of direct and indirect requirements to support a vector of final demand) would generally include imported intermediates at several

\(^{16}\) An alternative rule is the ‘substantial transformation’ test.

\(^{17}\) In development economics, this problem has long been known in the context of the impractical Hirschman recommendation that backward linkages be utilised to create the maximum inducement to invest. In the Leontief input-output framework, one cannot meaningfully break down the \(a_{ij}\) coefficients into \(d_{ij}\) (domestic) and \(m_{ij}\) (imported) coefficients for competitive imports.

\(\text{© Royal Economic Society 1998}\)
stages of manufacture, which are impossible for the same reasons to identify and quantify meaningfully. Again, even if we were to estimate such imported shares meaningfully, the imports are likely in turn to include, in today’s globalised production, intermediates produced by us and used by the producers abroad. The difficulties are myriad, even endless.

All of these problems, which inherently lead to absurd arbitrariness in trying to identify the origin of products, are seriously present when an FTA inevitably requires that the origin be established for virtually all traded products. Since there are different external tariffs among members in an FTA, the problem is additionally acute since the fear of nonmember goods coming into one’s territory at a lower tariff than one’s own, simply by entering through another lower-tariff member country, is palpable. But the problem does not disappear in a CU despite the common external tariff. Thus, whether a Nissan produced in Britain is a Japanese car that must be allowed free entry into France as a British car or must pay the external tariff as a Japanese car, has been a source of discord: the problem arises since the EU would like to treat the tariff on Japanese cars differently from that on other cars. By wanting to treat Japanese cars at a preferential disadvantage, the CU (which is a PTA) also runs inherently into problems of arbitrary definitions of origin.

In reality, FTAs have created yet further problems by having many different rules of origin, varying by products (as in NAFTA, for instance) and by FTAs (when, say, the EU has FTAs with different rules with several different non-EU countries). Indeed, by making possible the manipulation of rules of origin, PTA countries are open to protectionist capture. The problems inherently posed by PTAs in regard to the rules of origin that they require are yet further compounded since FTAs are on different schedules of tariff cutting by sector and are not synchronised (having been negotiated at different points of time and with different schedules for reaching zero tariff outcomes), so that we typically find a large and chaotic set of applicable tariffs on the same good, depending on which source the good is assigned to.

The result is what Bhagwati (1995) has called the ‘spaghetti bowl’ phenomenon of numerous and crisscrossing PTAs and innumerable applicable tariff rates depending on arbitrarily-determined and often a multiplicity of sources of origin. In short, the systemic effect is to generate a world of preferences, with all its well-known consequences, which increases transaction costs and facilitates protectionism. In the guise of freeing trade, PTAs have managed to recreate the preferences-ridden world of the 1930s as surely as protectionism did at the time. Irony, indeed!

---

18 The pioneering account of the protectionist outcomes resulting from rules of origin manipulation is by the lawyer David Palmer (1993) whereas the analytics of such protectionist outcomes with rules of origin have been developed by Krishna and Krueger (1995).

19 A beautiful example of how this spaghetti bowl phenomenon arises is provided by Richard Snape (1996), whose illustration of the phenomenon for the European Union is reproduced here as Fig. 3. Such illustrations were pioneered by Ron Wonnacott who has also focused the profession’s attention on what is now known as the ‘hub-and-spoke’ problem.

© Royal Economic Society 1998
3. 'Dynamic' Time-Path Analysis: Building versus Stumbling Blocks

In contrast to the question whether the immediate (static) effect of a PTA is good, we may ask whether the (dynamic time-path) effect of the PTA is to accelerate or decelerate the continued reduction of trade barriers towards the goal of reducing them worldwide.\textsuperscript{20} I.e. we have now, in the dynamic time-path

\textsuperscript{20} The following analysis draws heavily on Bhagwati and Panagariya (1996a).

© Royal Economic Society 1998
NB: Does not include countries of the former Soviet Union other than Baltic countries.

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

(b) European Economic Area

(c) European Free Trade Area

(d) Algeria, Egypt, Jordan, Lebanon, Morocco, Syria, Tunisia

---
| EU Single Market
---
| Customs Union
---
| Free Trade Area
---
| EU Association Agreements
---
| Non-reciprocal agreements

Fig. 3 (cont.)

case, the key concepts, introduced by Bhagwati, of PTAs acting as 'stumbling blocks' or 'building blocks' towards worldwide nondiscriminatory trade liberalisation, just as Viner introduced the key concepts of trade diversion and trade creation for the static analysis.

3.1. **Defining the 'Dynamic' Time-Path Question**

The time-path question needs to be properly specified. It can be formulated analytically in two ways:

**Question 1:** Assume that the time-path of MTN (multilateral trade negotiations) and the time-path of PTAs are separable and do not influence each other, so that neither hurts nor helps the other. Will then the PTA time-path be characterised by stagnant or negligible expansion of membership; or will we have expanding membership, with this even turning eventually into worldwide membership as in the WTO, thus arriving at nondiscriminatory free trade for all? The analysis can be extended to a comparison of the two time-paths, ranking the efficacy of the two methods of reducing trade barriers to achieve the goal of worldwide free trade for all.

**Question 2:** Assume instead, as is plausible, that if both the MTN and the PTA time-paths are embraced simultaneously, they will interact. In particular, the
policy of undertaking PTAs will have a malign or a benign impact on the progress along the MTN time-path.

Question 1 can be illustrated with the aid of Fig. 4 which portrays a sample of possibilities for the time-paths. World welfare is put on the vertical and time on the horizontal axis. For the PTA time-paths drawn, an upward movement along the path implies growing membership; for the MTN time-paths, it implies nondiscriminatory lowering of trade barriers among the nearly worldwide WTO membership instead. The PTA and MTN time-paths are assumed independent of each other, not allowing for the PTA time-path to either accelerate or decelerate the course of MTN (thus ruling out Question 2-type issues). The goal can be treated as reaching $U^*$, the worldwide freeing of trade barriers on a nondiscriminatory basis, at a specified time.

Then, Question 1 above can be illustrated by reference to the PTA paths I–IV. Thus, PTAs may improve welfare immediately, in the static sense or reduce it. In either case, the time-path could then be stagnant (as with time-paths II and III), implying a fragmentation of the world economy through no further expansion of the initial PTA. Else, it can lead (as in time-paths I and IV) to multilateral free trade for all at $U^*$ through continued expansion and coagulation of the PTAs. Under process multilateralism, i.e. MTN as a multilateral process of reducing trade barriers as distinct from multilateralism as the goal desired, the time-path may fail to reach $U^*$ and instead fall short at $U_m$ because of free-rider problems.

As indicated, if the PTA and MTN time-paths are interdependent, we can address Question 2. In that case, the MTN time-path becomes a function of whether the PTA time-path is traveled simultaneously.

The questions that we have distinguished above spring, as we noted at the outset, from a shift in US policy in favour of going Article 24 when the Europeans blocked the initiation of a new MTN Round at the GATT in 1981. In Bhagwati (1991; 1993), the challenge to international trade theorists to analyse these questions was first identified and a preliminary set of arguments offered. We now review the theoretical literature that has developed subsequently. At the outset, we consider recent theoretical approaches which, however interesting in themselves, are not helpful in thinking seriously about the time-path questions at hand. The more pertinent literature will be considered next.

3.2. Exogenously-determined Time Paths: Conventional Approaches

Kemp-Wan: The approach of Kemp and Wan (1976) seems to be pertinent to our questions but is not. Evidently, the PTA time-path to $U^*$ in Fig. 4 can be made monotonic provided the expanding membership of a PTA always satisfies the Kemp-Wan rule for forming a CU. But what this argument does not say, and indeed cannot say, is that the PTA will necessarily expand and, if so, in this Kemp-Wan fashion.

(ii). Krugman: The same argument applies to the theoretical approach introduced by Krugman (1993), where again the expansion of membership is
This figure illustrates the ‘building blocks’ and ‘stumbling blocks’ concepts in the context of the question whether the regionalism (that is, PTA) dynamic time-path will show increasing or stagnant membership. The PTA may improve welfare immediately, from $U^0$ to $U_p^2$, or (because trade diversion dominates) reduce it to $U_p^1$. The time-path with PTA, in either case, could then be stagnant (paths II and III), implying a fragmentation of the world economy through no further expansion of the initial trading bloc. Or, it could lead (paths I and IV) to multilateral free trade for all at $U^*$ through continued expansion and coagulation of the PTA. Under ‘process multilateralism’, the time-path may fail to reach $U^*$ and instead fall short at $U_m$ because of free-rider problems. Or it may overcome them and reach $U^*$. This diagram assumes that the time-paths are independent: embarking on the PTA path does not affect the process-multilateralism path. This interdependence is discussed in the text.
treated as exogenously specified, as in Viner, and the world welfare consequences of the world mechanically dividing into a steadily increasing number of symmetric blocs are examined. Srinivasan (1993) has critiqued the specific conclusions as reversible when symmetry is dropped. Deardorff and Stern (1995) have also advanced a similar critique.

But the main problem is that the Krugman formulation is again in the conventional mould of taking the membership of a PTA and its expansion as exogenously specified and examining its consequences. Instead, what we need is a rigorous political-economy-theoretic incentive-structure analysis which endogenises the question of membership expansion and thus helps to address the time-path questions.

3.3. Endogenously-determined Time Paths: Recent Theoretical Analyses

The analysis of the time-path question has therefore moved into formal political-economy-theoretic modelling. We provide here a synoptic review of the few significant contributions to date, organising the literature analytically in light of the two questions distinguished above.

Question 1: The single contribution that focuses on Question 1, i.e. the incentive to add members to a PTA, is by Baldwin (1993), who concentrates, in turn, on the incentive of nonmembers to join the PTA. He constructs a model to demonstrate that this incentive will be positive: the PTA will create a ‘domino’ effect, with outsiders wanting to become insiders on an escalator. The argument is basically driven by the fact that the PTA implies a loss of cost-competitiveness by imperfectly-competitive nonmember firms whose profits in the PTA markets decline because they must face the trade barriers that member countries’ firms do not have to face. These firms then lobby for entry, tilting the political equilibrium at the margin towards entry demands in their countries. The countries closest to the margin will then enter the bloc, assuming that the members have open entry, thus enlarging the market and thereby increasing the cost of nonmembership and pulling in countries at the next margin. Given the assumptions, including continuity, this domino model can take the PTA time-path to $U^*$ in Fig. 4.

Question 2: The rest of the theoretical contributions address Question 2, i.e. whether the PTA possibility and/or time-path helps or harms the MTN time-path. Here, the two major analyses to date, addressed directly and quite aptly to this question, by Krishna (1995) and Levy (1994), reach the ‘malign-impact’ conclusion.

Krishna models the political process in the fashion of the government acting in response to implicit lobbying by firms as a ‘clearinghouse’, showing in his oligopolistic-competition model that the PTA reduces the incentive of the two member countries to liberalise tariffs reciprocally with the nonmember world and that, with sufficient trade diversion, this incentive could be so reduced as to make impossible an initially feasible multilateral trade liberalisation.

Levy, who models the political process instead as in a median-voter model works with scale economies and product variety to demonstrate that bilateral
FTAs can undermine political support for multilateral free trade. At the same time, a benign impact is impossible in this model: if a multilateral free trade proposal is not feasible under autarky, the same multilateral proposal cannot be rendered feasible under any bilateral FTA.

The Krishna and Levy models therefore throw light on the incentive-structure questions at hand when the agents are the lobbying groups and interests that are affected by different policy options.

4. Policy Today

The policy debate on PTAs today finds economists arrayed on different sides. The main divisions appear to be among those who find them a mixed bag and would therefore content themselves with an effort at reducing their downside by reforming Article 24 and other disciplines (such as on AD actions) at the WTO so as to minimise the damage they can create, those who would reject them as a pox on the trading system and would like to see either a standstill or even a rollback, and those who would go along with them on a 'GATT plus' view of trade policy. It is fair to say that those who are skeptical or hostile have gained some ground on the intellectual side.

Their success is not commensurate on the policy level. However, there is definite progress. The US desire to turn APEC, the Pacific ocean trade grouping, into an FTA has been rejected so far by its Asian members who have embraced instead the notion that any trade liberalisation under APEC auspices should be on an MFN basis. Similarly, the proposal for TAFTA, floated initially by Foreign Minister Klaus Kinkel of Germany, covering the Atlantic ocean, has been transmuted into a non-FTA initiative. The Indian Ocean initiative, started by India, is also unlikely to become an FTA.

Meanwhile, the European Council of Ministers has recently resolved, in light of the concerns expressed by the anti-PTA economists, that the 'current architecture' of the EU’s trading system will be frozen, implying a standstill on new PTAs, except when a strong case can be made, satisfying several criteria. Also, the WTO is addressing the problem of Article 24 reform through its Committee on Regional Trade Agreements. The media have also reported extensively on the ‘new thinking’ on the wisdom of PTAs and their proliferation.

The only substantial FTA project that is still making the headlines is the FTAA, the Free Trade Area of the Americas, backed by the Clinton administration which seeks to leave its own stamp on something big—the Uruguay Round and NAFTA were inherited by them. It is also an agenda that has endorsement from several South American nations, whose enthusiasm for PTAs reminds one of the earlier attachment of Latin policymakers to import substitution and to preferences for the developing nations. 21

21 The former was advanced by the Prebisch school, though it had its own intellectual origins elsewhere. Its costs have been fully analysed in numerous studies by now. The latter, embodied in the GSP schemes advanced at the UNCTAD, proved to be a fruitless distraction from MFN liberalisation, as argued by several economists who have studied it, among them Robert Baldwin, Richard Cooper and Andre Sapir.

© Royal Economic Society 1998
Meanwhile, the ‘small’ PTAs, usually without some big country providing the centre to them, continue to proliferate, chiefly (though not exclusively) among the developing countries.\textsuperscript{22} The incentives to have these are often political. Each Trade Minister, and his Prime Minister, wish to leave their names behind on some trade grouping of their own; multilateralism produces no such rewards. Then there is the CNN theory of such PTAs: whereas multilateralism produces photo-op only for the big players such as Sir Leon Brittan and Ambassador Mickey Kantor, and nothing for the others, matters are more balanced at the regional trade meetings. Bureaucrats also get prestigious positions with each new PTA their nations sign up to. Then, of course, PTAs multiply by imitation: if MERCOSUR is going ahead, ASEAN follows suit, then SAPTA (among the South Asian nations) is started, and Gresham’s Law takes over.\textsuperscript{23} And finally, we must not forget that politicians, and much of the media, often do not understand the distinction between Free Trade and Free Trade Areas: they hold preVinerian ‘all-trade-is-good’ views for the simple reason that they have not been taught the distinction.

There is enough evidence therefore for continuing concern on the part of the economists who oppose PTAs. A rollback, such as through a ‘sunset clause’ in Article 24 which would bring external tariffs down to the internal tariffs within a definite time period for existing and future PTAs, is improbable even though many economists have suggested it. More PTAs will doubtless be added meanwhile to the many already in existence or in conception, including the substantial FTAA.

The strategy of the anti-PTA economists has to be three-pronged: first, continue the education and the agitation to alert policymakers to the dangers posed by PTAs; second, join forces to accelerate reform of Article 24 and of other disciplines at the WTO such that PTAs can do less harm; and third, push for worldwide freeing of trade at the border, preferably through the embrace of a visionary terminal date and a corresponding practical agenda to get there, so that the preferences implied by the PTAs are rendered void because preferences relative to zero are zero.\textsuperscript{24}

\textit{Columbia University}

\textit{University of Nottingham}

\textit{University of Maryland}

\textsuperscript{22} A complete list of preferential trade agreements under Article 24 terms, as also of partial ones under waivers and Enabling Clause exclusions, is included in the Appendix in Bhagwati and Panagariya (1996a).

\textsuperscript{23} In fact, the expansion of PTAs just when MFN trade liberalisation has brought down trade barriers worldwide has made economists wonder whether there is also a deeper link between the two. See, for instance, the recent theoretical analysis by Freund (1997), in her Columbia University dissertation, which neatly produces PTAs as an outcome of reduced MFN barriers.

\textsuperscript{24} Such a target was first proposed by Martin Wolf in The Financial Times and has been repeatedly endorsed by several economists.

© Royal Economic Society 1998
References


Bhagwati, Jagdish. (1993) 'Regionalism and multilateralism: an overview,' in (Melo and Panagariya, eds.).


© Royal Economic Society 1998


