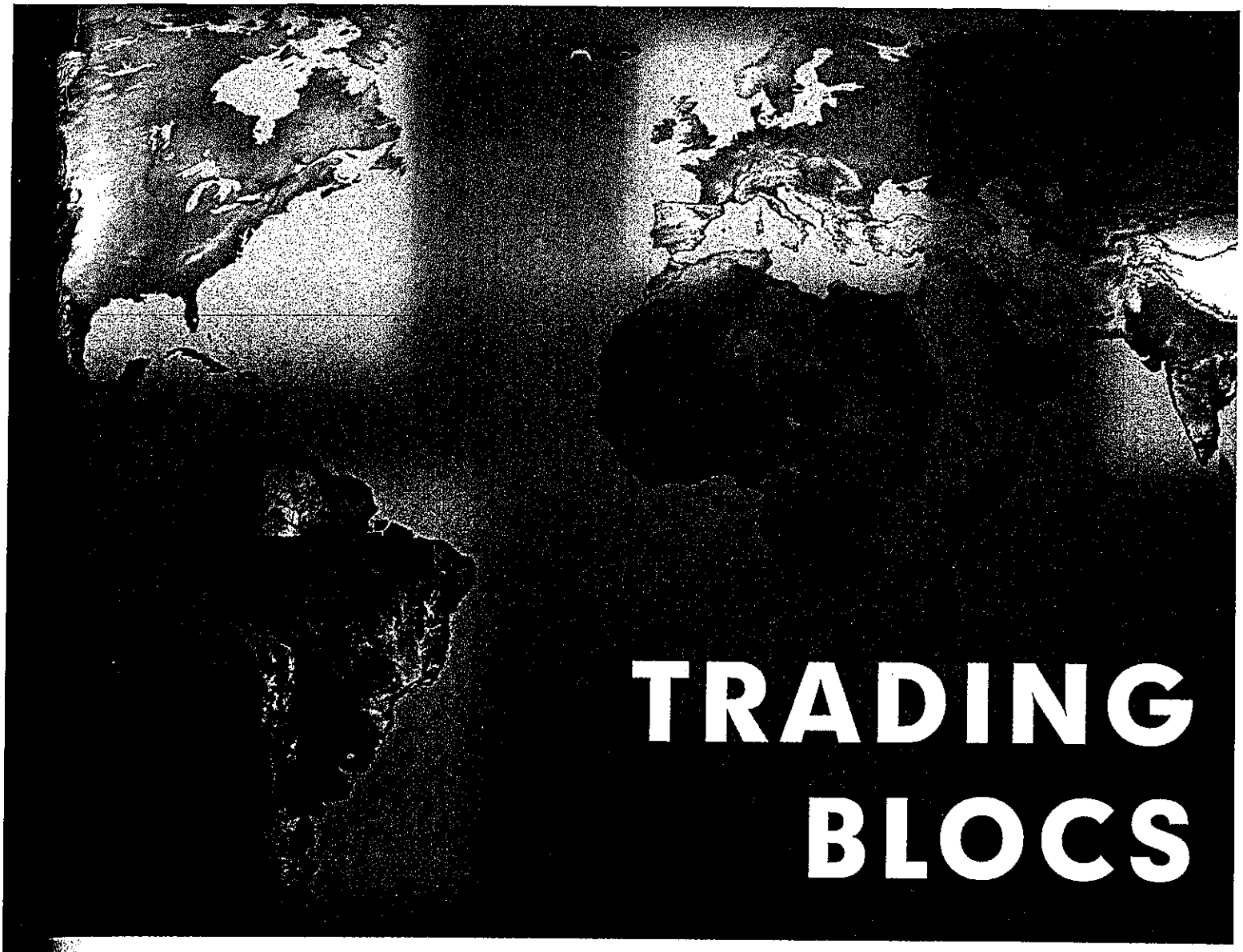


TRADING BLOCS

ALTERNATIVE APPROACHES TO ANALYZING PREFERENTIAL TRADE AGREEMENTS

edited by

JAGDISH BHAGWATI

PRAVIN KRISHNA

ARVIND PANAGARIYA

"BKP"

1 Regionalism and Multilateralism: An Overview

Jagdish Bhagwati

The question of "regionalism," defined broadly as preferential trade agreements among a subset of nations, is a longstanding one. As with all great issues, economists have long been divided on the wisdom of such arrangements. So have policymakers.

While this may not be evident to the many economists who are not inhibited by lack of comparative advantage from pronouncing on these matters, and whose pronouncements are a testimony to the enduring value of the theory of comparative advantage, preferential trade arrangements were debated by economists, as such, during the very formation of the General Agreement on Tariffs and Trade (GATT). The context, of course, was the difference between the British, led by Keynes, who were devoted to the continuation of the discrimination in Britain's favour through Imperial Preference, and the Americans, with Cordell Hull to the fore, who were strongly opposed, and embraced multilateralism and most favoured nation status (MFN) instead. Keynes, quite characteristically, at first denounced the American attachment to non-discrimination in trade and then later, when the British had virtually capitulated, celebrated its virtues with equal passion. I will juxtapose the two positions in Keynes' inimitable words, quoted in Bhagwati (1991a):

> My strong reaction against the word "discrimination" is the result of feeling so passionately that our hands must be free... [T]he word calls up and must call up ... all the old lumber, most-favored-nation clause and all the rest which was a notorious failure and made such a hash of the old world. We know also that it won't work. It is the clutch of the dead, or at least the moribund, hand.

> [The proposed policies] aim, above all, at the restoration of multilateral trade ... the bias of the policies before you is against bilateral barter and every kind of

Originally published in *New Dimensions in Regional Integration*, ed. Jaime de Melo and Arvind Panagariya (Cambridge, UK: World Bank and Cambridge University Press, 1993), 22–51. Reprinted with permission.

discriminatory practice. The separate blocs and all the friction and loss of friendship they must bring with them are expedients to which one may be driven in a hostile world where trade has ceased over wide areas to be cooperative and peaceful and where are forgotten the healthy rules of mutual advantage and equal treatment. But it is surely crazy to prefer that.

Closer to our times, the question of customs unions (CUs) and free-trade areas (FTAs), both permitted under GATT Article XXIV, became a major topic of theoretical research. The focus, however, since Jacob Viner's (1950) classic treatment, distinguishing between trade diversion and trade creation, was on showing that CUs and FTAs were not necessarily welfare-improving, either for member countries or for world welfare: in other words, the case for preferential trade arrangements was different from the case for free trade for all. The latter, enshrined in Adam Smith and Ricardo, and rigorously proved later by Samuelson (1939), Kemp (1972), and Grandmont and McFadden (1972), is a first-best case. The former, by contrast, reflects second-best considerations and was argued by Lipsey and Lancaster (1956–7), Lipsey (1957), Meade (1956), Johnson (1958a, 1958b) and others.[1]

But if the main focus of these analyses was on disabusing the faith in regionalism as being desirable (on static immediate-impact grounds) by analogy with the different and legitimate case for multilateralism (in the sense of free or freer trade for all), and thus could be seen as reinforcing the case for multilateralism, the effect could also go the other way, and did at times. One could thus argue, from the opposite counterfactual, that if you believed that regionalism, in being discriminatory, was necessarily inferior to non-discriminatory reduction of trade barriers, then this too was wrong. Ironically, in view of the later shift of his views to multilateralism and free trade, reflecting perhaps the changed intellectual environment in Cambridge and Chicago and also further reflection, it is interesting to quote Johnson (1967, pp. 163–4) in the context of proposals for trade preferences, for and among developing countries, for manufactured goods:

Both proposals violate the non-discrimination principle of the General Agreement on Tariffs and Trade and the GATT ban on new preferential arrangements other than customs unions and free trade areas embracing the bulk of the trade of the participating countries. This, however, does not mean that the proposed trading arrangements would necessarily be economically disadvantageous. The postwar development of the theory of customs unions and of commercial policy changes, culminating in the theory of second best has shown that in a tariff-ridden world economy there is no a priori reason for believing that nondiscrimination among import sources is economically superior to discriminatory trading arrangements. It

has demonstrated also that the question of whether a discriminatory tariff reduction improves or worsens the efficiency and economic welfare of the countries involved and the world as a whole depends on the empirical circumstances of the particular case.

Both the theory of second best and modern welfare economics (as well as ordinary common sense) indicate that policy changes that secure desirable results in terms of income distribution or other objectives at the cost of reduced economic efficiency may constitute improvements on a balance of gain and loss, and may legitimately be recommended if no more efficient method of achieving the same objectives is feasible or acceptable.

In fact, Johnson was an active proponent of NAFTA, an acronym which then stood for the North *Atlantic* Free-Trade Area (inclusive of the United Kingdom) rather than for the present North *American* Free-Trade Area which is predicated on a conceptually narrower, geographically-defined regional basis. As it happened, the concept of NAFTA failed to get off the ground, though the ideas concerning regional blocs and trading arrangements remained seductive through much of the 1960s, only to be abandoned thereafter until the recent 1980s' revival.

The recent revival of regionalism, which I describe as the "Second Regionalism" in contrast to, and because it is a sequel to, the "First Regionalism" of the 1960s, raises several of the old issues anew. But the historically changed situation which has resurrected regionalism equally provides the context in which it must be analysed, raising several new issues.

In this chapter, I address these manifold questions, dividing the analysis into a discussion of six areas:

- Article XXIV of the GATT, which sanctions CUs and FTAs (section 1.1);
- the "First Regionalism," briefly reviewing the factors that led to it and the reasons why, in the end, it failed (section 1.2);
- the "Second Regionalism," the reasons for its revival and its differential prospects (section 1.3);
- the key issues that this renewed regionalism raises, distinguishing among two main questions (section 1.4);
- the first, relating to the static impact effect of regional trade blocs (section 1.5);
- the second, concerning the dynamic time-path that regionalism offers, in itself and vis-à-vis multilateralism when the objective is to reach (non-discriminatory) free trade for all, so that one asks "whether multilateralism is the best way to get to multilateralism," therefore distinguishing between "process multilateralism" and "outcome multilateralism" (section 1.6).

In the light of this analysis, I conclude by examining the current US trade-policy shift to regionalism and arguing for a change in its focus from "piecemeal" to "programmatic" regionalism, less antithetical to reaching the "outcome-multilateralism" objective of eventual free trade for all (section 1.7). Some final observations conclude the chapter (section 1.8).

1.1 Article XXIV of the GATT: Rationale

The principle of non-discrimination is central to the final conception of the GATT, signed on 30 October 1947 by representatives from 23 countries in Geneva. Article I embodies the strong support for non-discrimination, requiring (unconditional) MFN for all GATT members.

Aside from "grandfathering" provisions, the only significant exception to MFN is made in Article XXIV, which permits CUs and FTAs and therefore sanctions preferential trade-barrier reductions among a subset of GATT members, as long as they go all the way to elimination.[2]

It is an intriguing question as to why Article XXIV was accepted, and it is a question that also has significance for some of the issues raised by the "Second Regionalism." It is a bit odd that an exception to MFN should be allowed as long as it is total (going all the way to 100 percent) rather than partial (say, 20 percent preference for one's favoured friends): it is as if your cardinal told you that petting is more morally reprehensible than sleeping around. In fact the post-Vinerian theory of preferential trade areas suggests that 100 percent preferences are less likely to increase welfare than partial preferences.[3]

The rationale for Article XXIV's inclusion in the GATT must therefore be explained in other ways. Perhaps there was an inchoate, if strong, feeling that integration with 100 percent preferences was somehow special and consonant with the objective on multilateralism. Thus, Dam (1970, pp. 274–5) quotes the prominent US official Clair Wilcox as follows:

> A Custom union (with 100% preferences) creates a wider trading area, removes obstacles to competition, makes possible a more economic allocation of resources and thus operates to increase production and raise planes of living. A preferential system (less than 100%) on the other hand, retains internal barriers, obstructs economy in production, and restrains the growth of income and demand ... A customs union is conducive to the expansion of trade on a basis of multilateralism and nondiscrimination; a preferential system is not.

Wilcox's statement was little more than assertion, however. But the rationale for inclusion of Article XXIV in the GATT appears to have been threefold, as follows:

• Full integration on trade, that is, going all the way down to freedom of trade flows among any subset of GATT members, would have to be allowed since it created an important element of single-nation characteristics (such as virtual freedom of trade and factor movements) among these nations, and implied that the resulting quasi-national status following from such integration in trade legitimated the exception to MFN obligation toward other GATT members.

• The fact that the exception would be permitted only for the extremely difficult case where all trade barriers would need to come down seemed to preclude the possibility that all kinds of preferential arrangements would break out, returning the world to the fragmented, discriminatory bilateralism-infested situation of the 1930s.

• One could also think of Article XXIV as permitting a supplemental, practical route to the universal free trade that GATT favoured as the ultimate goal, with the general negotiations during the many Rounds leading to a dismantling of trade barriers on a GATT-wide basis while deeper integration would be achieved simultaneously within those areas where politics permitted faster movement to free trade under a strategy of full and time-bound commitment. This is an argument that is not at centre stage: is regionalism truly a building, rather than a stumbling, bloc towards multilateral free trade for all: in other words, will it fragment, or integrate, the world economy?

The clear determination of 100 percent preferences as compatible with multilateralism and non-discrimination, and the equally firm view that anything less was not, meant that when Article XXIV was drafted, its principal objective was to close all possible loopholes by which it could degenerate into a justification for preferential arrangements of less than 100 percent; paragraphs 4–10 of Article XXIV were written precisely for this purpose. But, as is now commonly conceded, their inherent ambiguity and the political pressures for approval of substantial regional groupings of preferences of less than 100 percent have combined to frustrate the full import of the original desire to sanction only 100 percent preferences.

This tension between intention and reality has a direct bearing on the important question of strengthening Article XXIV today beyond even what its original drafters intended. I will therefore sketch briefly the important respects in which the original intention of Article XXIV was reasonably clear but was occasionally violated in spirit, to the point where the great expert on GATT law, Professor John Jackson, has gone so far as to observe that the accommodation of the European Common Market's

imperfect union in disregard of the legal requirements of Article XXIV was the beginning of the breakdown of the GATT's legal discipline, which we now seek to repair.[4] Two issues suffice to demonstrate this contention.

First, in regard to the elimination of internal barriers down to 100 percent, there was enough scope within the language of Article XXIV, paragraph 8, for its intent to be successfully avoided. Ambiguities could be exploited on two main fronts.

The first ambiguity lay in the directive that "duties and other restrictive regulations on commerce" were (with specified exceptions permitted under Article XI, XII, XIII, XIV, and XX) to be "eliminated with respect to substantially all the trade between the constituent territories." Skilful lawyers and representatives of governments could work wonders with the concept of "substantially all the trade," and then, even if a percentage cutoff point was accepted for this purpose (for example, 75 percent of all initial trade), important issues remained ambiguous, such as whether across-the-board (75 percent) cut on everything were required or whether substantial sectors could be left out altogether from the scope of the cuts—the latter being evidently at variance with the intent of those who favoured (100 percent) CUs but opposed (less than 100 percent) preferential arrangements. With both interpretations possible, sectorally non-uniform preferential arrangements could evidently not effectively be ruled out.

An ambiguity of equal importance arose in regard to the problem of the speed with which the "100 percent preferences" would be implemented. Evidently, if they were stretched out over very long periods, one was de facto sanctioning "less than 100 percent" preferential arrangements. In GATT jargon, this was the problem of "interim arrangements." Paragraph 5 therefore addressed this issue, requiring "a plan and schedule," and asking for the CU or FTA to be fully consummated "within a reasonable length of time." Paragraph 7, in turn, laid down specific procedures for such interim arrangements to be approved. Needless to say, this nonetheless left the door open for substantial laxity in conception and execution of the CUs and FTAs under Article XXIV.

Dam's (1970, p. 290) overall judgement of the outcome is perhaps too harsh, but is certainly in the ballpark:

> The record is not comforting ... Perhaps only one of the more than one dozen regional arrangements that have come before the GATT complied fully with Article XXIV criteria. That was the recent United Kingdom/Ireland Free-Trade Area, and even in that case certain doubts were expressed before the working party. In some

cases, the regional arrangements were very wide off the mark. The European Coal and Steel Community, covering only two major product lines, could not even qualify for the special regional-arrangement waiver of Article XXIV: 10 but required a general waiver under Article XXV: 5. The New Zealand/Australia Free-Trade Agreement, although not purportedly an example of "functional integration," provided for the liberalization of an even smaller percentage of intermember trade. A strong tendency has also been manifested for interim agreements to provide for an even longer transitional period and to contain increasingly fewer detailed commitments for eventual completion of the customs union or free-trade area.

1.2 The "First Regionalism": Failure in the 1960s

In any event, one can correctly assert (based on the acceptance of Article XXIV into the GATT) that regionalism, in the shape of (100 percent) CUs and FTAs, was not generally considered, by the architects of the GATT or by the United States, which was the chief proponent of multilateralism and non-discrimination, as antithetical to the GATT and to these principles.

1. Nonetheless, the United States, long suspicious of discriminatory trade arrangements, restrained itself from resorting to Article XXIV. The formation of the European Community in 1958 marked a partial watershed. The United States puts its shoulder to the wheel and saw the Common Market through, negotiating around the different hoops of Article XXIV, emasculating the Article somewhat so as to seek GATT approval of an imperfect union (especially in regard to discriminatory preferences for the eighteen ex-colonies in Africa that the Europeans insisted on retaining, requiring therefore a waiver of GATT rules), all in the cause of what it saw as a *politically* beneficial union of the original six nations that formed the Community. But despite the enthusiasm of many to follow the European Community with a NAFTA, and even a Pacific Free-Trade Area (PAFTA), centred on the United States, nothing came of it: the United States remained indifferent to such notions.[5]

2. There was an outbreak of FTA proposals in the developing countries as well. While stimulated by the European examples, they were motivated by the altogether different economic rationale formulated by Cooper and Massell (1965a, 1965b), Johnson (1965) and Bhagwati (1968). This was that, given any targeted level of import-substituting industrialisation, the developing countries with their small markets could reduce the cost of this industrialisation by exploiting economies of scale through preferential opening of markets with one another.[6] By the end of the 1960s, however, the attempts at forming regional FTAs and CUs along these lines had also

collapsed. The problem was that, rather than use trade liberalisation and hence prices to guide industry allocation, the developing countries attempting such unions sought to allocate industries by bureaucratic negotiation and to tie trade to such allocations, putting the cart before the horse and killing the forward motion.

Thus, while the world was indeed filled with proposals for NAFTA, PAFTA, LAFTA (the Latin American Free-Trade Area, replaced by LAIA, the Latin American Integration Agreement, in 1980), and ever more in the 1960s, until one could be forgiven for imagining that a veritable chemical revolution had broken out, regionalism had virtually died by the end of the decade, except for the original European Community and EFTA.

1.3 The "Second Regionalism": Revival in the 1980s

But regionalism (i.e., preferential trade liberalisation) is now back. Those who do not know the history of the "First Regionalism" are doomed to extrapolate from the current political ferment in favour of FTAs and CUs and assume uncritically that regionalism is here to stay. Those who know the history may make the reverse mistake of thinking that regionalism will again fail. I believe that careful analysis of the causes of the resurrection of regionalism suggests that regionalism this time is likely to endure.

The main driving force for regionalism today is the conversion of the United States, hitherto an abstaining party, to Article XXIV. Beginning with the FTA with Israel (a reflection of the special relationship between the two nations and hence of little general value), the FTA with Canada marked a distinct change. Now the NAFTA is being negotiated with Mexico, and the Enterprise for the Americas' Initiative (EAI) envisages more FTAs with the nations of South America, with Chile at the head of the line.

The conversion of the United States is of major significance. As the key defender of multilateralism through the postwar years, its decision now to travel the regional route (in the geographical *and* the preferential senses simultaneously) tilts the balance of forces at the margin away from multilateralism to regionalism. This shift has taken place in the context of an anti-multilateralist ethos that has reflected alternative but nonetheless eventually reinforcing views:

- The "Memorial Drive" school[7] holds that the GATT is dead (Thurow: Davos) or that the GATT should be killed (Dornbusch).[8] Regionalism is

then presented in effect as an *alternative* to multilateralism. This school, aptly named in view of its funereal approach to multilateralism, has influence in Democratic circles and plays to the prejudices that one finds in Congressional circles that mistakenly identify multilateralism with America's postwar altruism and regionalism (with its connotation of "exploiting our own markets for ourselves") with the presumed current necessity finally to "look after one's interests."

• An alternative view is that regionalism is a useful *supplement*, not an alternative, to multilateralism. "We are walking on only two legs" is the popular argument. That we may wind up walking on all fours is ignored.

• It is also often asserted that regionalism will not merely supplement multilateralism. It will also *accelerate* the multilateral process: the threat of going (unilateral and) regional will produce multilateral agreements that may otherwise be held up. (However, this may be an optimistic view since threats that have to be implemented and repeatedly made, as has been the case with US regionalism, are not efficient threats; and they change external perceptions about what US trade policy priorities are, quite regardless of what the United States asserts to be its true intentions. In fact, the taking of two roads simultaneously can affect adversely the travel down one, as I argue below at length.)

• The panic over the continuing payments deficit has also fed demands for "quick" results on trade (although the two issues are broadly delinkable: payments surpluses and deficits are macroeconomic phenomena that are not influenced in any predictable way by trade policy changes whose impact on the difference between domestic savings and investment, if any, can come in different ways that can go in opposing directions). Associated with this has been impatience with the pace of the multilateral trade-negotiating process and the non sequitur (examined below) that regionalism necessarily works faster.

• In addition, "Europe 1992" and the impending integration of Eastern Europe into the European Community have reinforced, as the formation of the Common Market did with many three decades ago, those in North America who feel that a countervailing bloc must be formed there as well. Indeed, the fear that European investments would be diverted to Eastern Europe, once it is integrated with the European Community, was cited by President Salinas of Mexico as a factor decisively pushing him toward the Mexico–US FTA: this would, he felt, enable Mexico to get the investment needed from America and Japan.

• There are strong non-economic, political and cultural factors also driving Mexico toward an FTA with its northern neighbour. Just as the Turks since Ataturk have tried to seek a European rather than an Arab (or Islamic) identity, the Mexicans clearly now seek an American future rather than one with their southern neighbours. The Hispanic (economic) destiny that many in America fear from illegal immigration and integration with Mexico has its flip side in the American (economic) destiny that Mexico's reforming elite, trained in the top universities in the United States, hope for.

• The offer in June 1990 by President Bush to get more nations from South America to join the United States in an FTA, as part of a general package of economic initiatives to assist these nations, reflects the compulsions that the debt crisis there imposes on American policy to respond in a regional framework to ensure that this crisis remains manageable and does not engulf the United States, whose banks are principally endangered by it.

• Then again, the response of South American nations to the prospect of FTAs with NAFTA, and in some cases with one another first and then joining up with NAFTA, has been enthusiastic. This time around, the prospects are better that in the 1960s. Quite simply, there is now a marked shift in economic thinking towards trade liberalisation and market forces. The macroeconomic crisis of the 1980s has fed the movement to microeconomic reforms, much as it is currently doing in India. The changed economic and political attitudes are comforting to those of us who went into the trenches to fight these battles as early as the 1960s. It is also amusing to see those who dismissed our arguments as "reactionary" or "ideological" then, now embracing these ideas and policies and the leaders who are implementing them, with no apologies to us and with a facade of independently-obtained wisdom. But, frankly, it is good to have them finally on the right side; and it is good to have been in the right.

• Finally, the conjunction of the two dramatic events, "Europe 1992" and the US–Canada FTA, even though fortuitous and prompted by different motivations and historical circumstances, certainly has created a sense elsewhere that regionalism is the order of the day, and that others must follow suit. In the Far East, for instance, there has been a sense that a Japan-centred regional bloc may be necessary in a bloc-infested world, and Malaysia has actively sought a Japan-centred Asian bloc to rival and confront the US-led Americas bloc.

1.4 Regionalism versus Multilateralism: Key Questions

I suspect therefore that the "Second Regionalism" will endure: it shows many signs of strength and few points of vulnerability. But if so, those of us who see virtue in a rule-based, open and multilateral trading system must ask searching questions as to its compatibility with such discriminatory trading arrangements. In particular, two major questions must be answered:

- Is the immediate impact effect of such preferential trade blocs, whether CUs or FTAs, to reduce rather than increase world welfare?
- Regardless of the immediate impact effect, will regionalism lead to non-discriminatory multilateral free trade for all, through continued expansion of the regional blocs until universal free trade is reached, or will it fragment the world economy? And will, in any event, such a dynamic time-path show that regionalism will get us closer to the *goal* of multilateral free trade for all than multilateralism as the *process* of trade negotiation will?

I shall now treat each of these two important, and distinct (if at times analytically interrelated), questions in turn.

1.5 The Static Impact-Effect Question

The question of the static impact effect of preferential trade arrangements such as FTAs and CUs is, quite simply, the question raised by Viner (1950): would not such discriminatory arrangements be trade-diverting rather than trade-creating?[9]

It is important to raise this question because, as Viner taught us, FTAs and CUs are two-faced: they liberalise trade (among members), but they also protect (against outsiders). The important issue therefore is: which aspect of an FTA or a CU is dominant? Or, to put it in the economist's language: is a particular FTA or CU trade-diverting (that is, taking trade away from efficient outside suppliers and giving it to inefficient member countries) or trade-creating (that is, generating trade from one more efficient member at the expense of another less efficient member)?

Sadly, one might have scanned the leading articles, the editorials, and the Congressional testimony when the renewal of fast-track authority for the extension of NAFTA to Mexico was being debated in 1991, looking for references to trade diversion—and find scarcely any. Astonishingly, it

was not just the politicians and lawyers for Mexico's lobby who equated the FTA with (non-discriminatory) free trade; reputed economists did so too.[10]

What can we say about this issue? In particular, what can we propose to ensure that, if CUs and FTAs are to flourish, they do not become trade-diversionary? Article XXIV's injunction not to raise the CU's or the FTA's average external tariff can be interpreted as a precaution against trade diversion and harm to outside GATT members, though (as argued below) this is not a satisfactory way to do it.

In essence, there are three approaches to containing the fallout of trade diversion from CUs and FTAs.

1.5.1 Converting Preferential CUs and FTAs into (Geographically) Regional Blocs

It is occasionally argued that we should encourage geographically proximate countries to form CUs and FTAs, discouraging geographically distant countries from doing so since the latter would be more likely to be trade-diverting.[11] This is a misguided prescription in my view, for several reasons.

To see this, it must be first appreciated that it rests on a syllogism. The first premise is that a CU or FTA is more likely to create trade and thus raise welfare, given a country's volume of international trade, the higher is the proportion of trade with the country's CU or FTA partners and the lower is this proportion with the non-member countries. The second premise is that countries sharing borders, or closer geographically to one another, have higher proportions of trade with one another than countries further apart do.

The first premise is, of course, well known to trade economists from the early post-Vinerian theory, as developed by Lipsey (1958). But Lipsey's argument focuses on the relative sizes of imports from each source vis-à-vis expenditure on domestic goods as the key and decisive factor in determining the size of losses and gains from the preferential cuts in trade barriers.[12]

While the likelihood argument is valid within the Lipsey model, it must be noted that it is only that. Thus, for specific CUs and FTAs, the *actual* welfare effects will depend, not merely on the trade and expenditure shares à la Lipsey but also on the *substitution* at the margin between commodities. Thus, for instance, the substitution between non-member goods and domestic goods may be very high, so that the costs of discrimination

would tend to be high as well, ceteris paribus. In short, it is important to guess at substitution elasticities among goods *as well as* trade shares, with and between members and non-members of CUs and FTAs, to arrive at a better picture of the likely effects of *specific* CUs and FTAs that may be proposed.

As for the second premise, I have problems with this too, as a policy guideline. If I had access to captive research assistance and funds, I could examine whether, for all conceivable combinations of countries and distances among them, and for several different time periods, the premise is valid. I do not, so I must rely on casual empiricism and a priori arguments. Compare, for instance, the trade throughout the 1960s between India and Pakistan with that between India and the United Kingdom or the then USSR. The former trade has been smaller than the latter. Borders can breed hostility and undermine trade, just as alliances among distant countries with shared causes can promote trade (Gowa and Mansfield, 1991). The flag follows trade; and trade equally follows the flag which, at least in the 19th-century European expansion, was not directly across the European nations' borders. Again, even if the premise is statistically valid for any set of observations, it may be a result of trade diversion itself: proximity may have led to preferential grant of concessions such as OAP and GSP at the expense of countries elsewhere.

In short, prescriptions to confine CUs and FTAs only to geographically proximate countries are not defensible because both premises have problems: the former is, at best, a likelihood proposition that should not be applied to specific situations where the welfare impact depends critically on other variables as well, whereas the latter does not have a firm empirical or conceptual basis.

But possibly the most damaging criticism that one can make of such a prescription is that it concentrates, at best, on the static impact-effect question and ignores the more important dynamic time-path question. By prescribing that we must rule out "distant" country unions, as between the United States and Israel and Chile, we would make the CUs and FTAs more exclusive and less open to new members, undercutting the objective of moving speedily towards the shared objective of (non-discriminatory) multilateral free trade for all. That would be tragic indeed.

1.5.2 Designing Disciplines to Minimise Trade Diversion

A different, and my preferred, approach is not to pretend to find rules of thumb to exclude CUs and FTAs "likely" to be trade-diversionary, but

rather to examine the different ways in which trade diversion could arise and then to establish disciplines that would minimise its incidence.

1.5.2.1 Article XXIV In a sense, Article XXIV (paragraph 5) seeks to do this by requiring that CUs, which must have a common external tariff, should ensure that this common tariff "shall not on the whole be higher or more restrictive than the general incidence of the duties and regulations of commerce applicable ... prior to the formation of such a union." For FTAs, the rule is that the 'duties and other regulations of commerce' are not to be "higher or more restrictive" than those previously in effect.

Evidently, when tariffs change, as in CUs, and some increase and others fall, the scope for skulduggery arises again, since Article XXIV leaves the matter wholly ambiguous. As Dam (1970, p. 217) has noted: "these ambiguities plagued the review by the CONTRACTING PARTIES to the EEC Treaty of Rome—The Six, having used an arithmetic average, refused to discuss the best method of calculation, because in their view paragraph 5 did not require any special method."

Besides, it is evident to trade economists that maintaining external tariffs unchanged is, in any event, not the same as eliminating trade diversion. What *can* be said is that, the lower the external barriers, the less is the scope for diverting efficient foreign supplies to member countries. A desirable discipline to impose on CUs and FTAs would thus be to require, for Article XXIV sanction, that one price to be paid must be the simultaneous reduction of the external tariff (implicit and explicit), pro rata to the progressive elimination of internal trade barriers.

Possible ways of ensuring this may be indirect disciplines. One way would be to modify Article XXIV to rule out FTAs with diverse tariffs by members[13] and to permit only CUs with common external tariffs (CETs). With most tariffs bound, this would ensure that for the most part a substantial downward shift in tariffs would be a consequence—that, say, Argentina or Brazil would be lowering her trade barriers, *not* that the United States would be raising hers. Since regionalism is probably going to be a matter of low trade barrier hubs such as the United States and Japan, joining with their respective regional spokes, this insistence on CUs could perhaps produce excellent results.

An alternative, and surer, way would be to insist on CUs but also write into Article XXIV the requirement that the *lowest* tariff of any union member on an item *before* the union must be part of the CET of the union.

1.5.2.2 Articles VI and XIX: AD and VERs But none of this is enough today. For the trade economists who work in a sustained way on the problems of the world trading system are aware that protection today takes the form of unfair capture of fair trade mechanisms such as anti-dumping (AD) actions and of voluntary export restraints (VERs); countries today thus have access to selective and elastic instruments of protection.[14] Given this reality, even the modification of Article XXIV, to ensure that the external (implicit and explicit) tariff barriers come down as a price for CUs to be allowed under GATT rules, will leave open a gaping hole that would be tantamount to an open invitation to trade diversion by these preferential arrangements. In fact, trade creation can degenerate rapidly into trade diversion, when AD actions and VERs are freely used.

Imagine that the United States begins to eliminate (by outcompeting) an inefficient Mexican industry once the FTA goes into effect. Even though the most efficient producer is Taiwan, if the next efficient United States outcompetes the least efficient Mexico, that would be desirable trade creation (though the best course would be free trade so that Taiwan would take more of the Mexican market instead).

But what would the Mexicans be likely to do? They would probably start AD actions against Taiwan, which would lead to reduced imports from Taiwan as the imports from the United States increased, leaving the Mexican production relatively unaffected: trade diversion from Taiwan to the United States would have occurred. Similarly, the effect of Mexican competition against the United States could well be that the United States would start AD actions and even VERs against Taiwan.

My belief that FTAs will lead to considerable trade diversion (because of modern methods of protection, which are inherently selective and can be captured readily for protectionist purposes) is one that may have been borne out in the European Community. It is well known that the European Community has used AD actions and VERs profusely to erect "Fortress Europe" against the Far East. Cannot much of this be a trade-diverting policy in response to the intensification of internal competition among the member states of the European Community?[15]

Two conclusions follow: (1) If inherently discriminatory regionalism is to flourish, as seems likely, then we need greater discipline for AD actions and VERs; Article VI needs reform and Article XIX needs compliance alongside the elimination of VERs (as the Dunkel draft on the MTN recommends). (2) This also implies that regionalism means, not the redundancy of the GATT, but the need for a stronger GATT. Those who think

of the two as alternatives are prisoners of defunct modes of thinking, based on the days when protection was a different beast.

1.5.2.3 Judging Trade Diversion Case by Case While the foregoing analysis embraces a set of policy-framework and incentive-creating reforms to minimise trade diversion, an alternative approach to the problem could be in terms of a case-by-case approach where the approval by the GATT of a proposed CU or FTA would depend on the evaluation of its trade-creating and trade-diverting effects and the requirement that the net anticipated effect be trade-creating.

McMillan (1991) has argued this in an ingenious paper[16] which proposes a simple test of admissibility: "does the bloc result in less trade between member countries and outsider countries?" Based on the welfare economics of CU theory, this is an aggregative test and therefore has some obvious analytical problems. It is also subject to the problem of computing plausible trade outcomes. It is hard enough to apply it ex post; ex ante, as a test of admissibility, I see little prospect of its being effectively used to exclude any proposed CU or FTA.

Its main merit is its apparent simplicity and its better grounding in economic theory. I therefore endorse the advisability of *some* version of the McMillan test replacing in Article XXIV the current requirement not to raise the average external tariff. But I see it as doing little *in practice* to avoid trade diversion. For this, we will have to rely on changing the incentive structure, including through suitable constraints imposed by stricter discipline on selective and elastic targeting of foreign suppliers. The issue of constraining trade diversion from proliferating preferential groupings is so important that it may not be a bad idea to *combine* the proposals made by McMillan and myself, rather than to treat them as alternatives.

1.6 The Dynamic Time-Path Question

The question of the dynamic time-path is particularly difficult: it is almost virgin territory.

Perhaps the theoretical approach to CU theory that appears to be most relevant to this problem is that of Kemp and Wan (1976). In contrast to the Vinerian approach, Kemp and Wan make the external tariff structure endogenously determined for the CU such that it improves the CU members' welfare while maintaining the outsiders' welfare unchanged.

This restores the pre-Vinerian intuition that a CU should be welfare-improving. The problem with the operational significance of the Kemp–Wan argument is that it really is an existence argument, without any structure being put on it within the context of a specific model so that we can develop intuition about what the external tariff structure for such a Kemp–Wan CU would be.[17] But that *any* subset of countries *could* form an unambiguously (world) welfare-improving union is definitely established by Kemp and Wan.

This also implies that the time-path to U^* (see figure 2.9, this volume) achieved under multilateral free trade as the optimum optimorum, can be made monotonic.[18] But what it does *not* say is that the union will necessarily expand and, if so, in a monotonically welfare-improving manner. For *that* answer, we must turn to the *incentive structure* that any CU provides to relevant "groups" for further expansion of the CU.

The incentives in question need not be *economic* incentives. In fact, it is hard to imagine that the arbitrary groupings of countries that seek FTAs and CUs are dependent on economic arguments as their key determinants. Often, politics seems to drive these choices of partners, as in the case of the European Community, and now in the case of FTAs throughout the Americas. This also accounts for the occasional non-regionally proximate choices of partners in such blocs: e.g., the United States and Israel, and Pakistan, Iran and Turkey in the early 1960s. But that economic factors contribute to the incentives for such blocs to be formed is not implausible. Thus, for instance, Edward Mansfield, a Columbia University political scientist, has suggested that trade blocs will tend to be formed by security-driven allies because the gains from trade from them will accrue to friends rather than foes.[19]

A meaningful examination of the incentives to form and to expand trade blocs will therefore have to be in the new and growing field of political economy-theoretic analysis. I believe that the models within which we investigate these issues will have to distinguish among at least three kinds of "agents," which I will detail below with illustrations of the kinds of arguments which we would find relevant:

- *Governments of member countries:* whether a CU will expand or not will depend partly on the willingness of the CU authorities to do so. This will be affected by ideas and ideology. Here I worry that CUs will be under pressure *not* to expand because one possible reaction to a CU will be: "we are already a large market, so what do we really stand to gain by going

through the hassle of adding more members?" This is what I call the "Our Market Is Large Enough" syndrome. I think, as Martin Wolf has often noted, that large countries tend to be more inward-looking for precisely this type of reason.

In addition, the expansion of the CU to include any specific set of outside countries will imply differential aggregate-welfare effects for current members, implying in turn differential incentives for member countries for and against the expansion.[20] In this context, a CU (which generally includes transfers among members) may be more expansionary (à la Kemp–Wan argumentation) than an FTA, though a CU that simultaneously seeks *political* integration may be less willing to expand.

• *Interest groups in member countries:* We need also to consider how interest groups, who lobby for or against CU expansion, will behave. Again, since CUs are a balance of trade-creating and trade-protecting forces, it is possible that the protectionists who profit from the diversion of trade away from efficient suppliers abroad to themselves will line up against CU expansion to include those suppliers. The problem then will be the "These Are Our Markets" syndrome.

This syndrome is not absent from the NAFTA scene, as many leader articles and media quote from business groups testified during the fast-track renewal. In fact, this syndrome was also present in the Eastman Kodak pamphlet (Dornbusch et al., 1989) that I cited earlier. It is also a sentiment that was beautifully expressed by Signor Agnelli of Fiat: "The single market must first offer an advantage to European companies. This is a message we must insist on without hesitation."[21] It is, of course, fine or Signor Agnelli to express such sentiments: after all, Fiat has run for years, not on gas, but on VERs against the Japanese. But should economists also embrace such sentiments?

• *Interest groups and governments of outside countries:* The third set of "agents" has to be the outside countries. Here, the example of a CU may lead others to emulate and seek entry. Otherwise, the fear of trade diversion may also induce outsiders to seek entry: Irwin's marvellous study on the historical experience with trade liberalisation in the 19th century (1993) shows that the Anglo–French Treaty may well have served this purpose. If so, this acts as an incentive to expand the CU.

This is clearly an uncharted area that is evidently the most interesting for further analysis.[22] I might add just one empirical–econometric study, by Mansfield (1992), which takes trade data for 1850–1965, estimates an

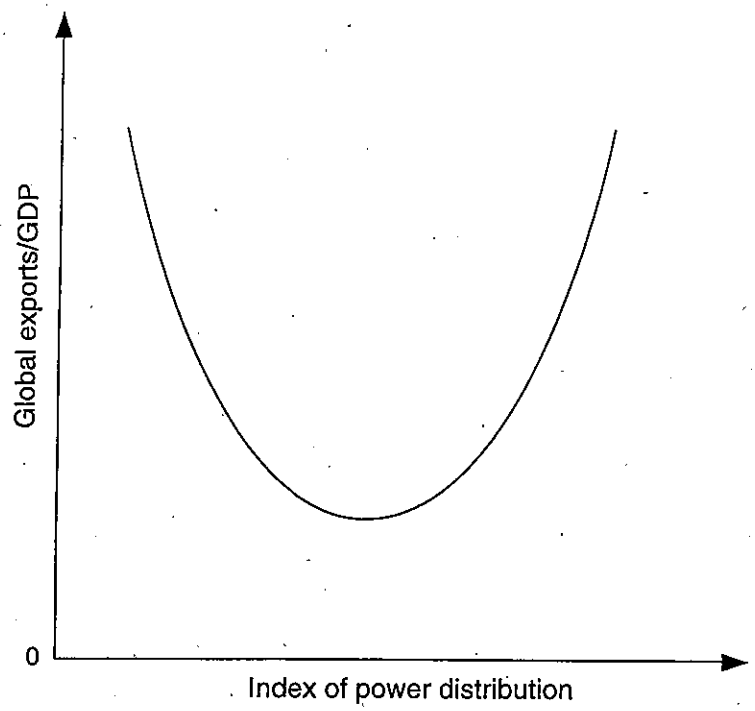


Figure 1.1
Concentration of power in the world and global exports. *Source*: Mansfield (1992).

index of "power distribution" (reflecting, among other things, trade blocs and economic power distribution) and comes up with figure 1.1. When power was centred in hegemons, during periods of British and American hegemony, and when there was "anarchy," the world economy was relatively liberalised (in the sense that global exports: GDP ratio was high); when there were a few middle-sized powers, as could happen with trade blocs, the result was a smaller ratio of trade to GDP (figure 1.1).

If Mansfield's analysis is accepted, and if it is considered to be a reasonable approximation to the question whether CUs will have expansionist or protectionist outlooks (mapping perhaps also into their attitudes to CU expansion or stagnation), then the presumption would be that historical experience suggests that trade blocs will fragment the world economy, not go on to unify it. Of course, history does not always repeat itself. But Mansfield's work certainly suggests caution in place of the gung-ho regionalism that has been urged by the Memorial Drive School.

To conclude, consider the following popular assertions by the regionalists:

- regionalism is quicker;
- regionalism is more efficient; and
- regionalism is more certain.

1.6.1 Is Regionalism Quicker?

The regionalists claim that the GATT is the "General Agreement to Talk and Talk," whereas regionalism proceeds quicker. But is this really so?

1. Historically, at least, the "First Regionalism" failed whereas the GATT oversaw the effective dismantling of prewar tariffs in the OECD countries and the enlargement of disciplines over NTBs at the Tokyo Round and beyond. A little caution, to say the least, is necessary before celebrating regionalism's quickfootedness.

2. For those who believe that regionalism offers a quick route to effective trade liberalisation, Dam's analysis quoted above needs renewed attention. There is a world of difference between announcing an FTA or a CU and its implementation, and the comparison is not pleasing if you are in the regional camp.

3. As for speed, even the best example of regionalism, the European Community, started almost four decades ago (1957) and is now into 1992. The "transition" has not therefore been instantaneous any more than negotiated reductions of trade barriers under the GATT Rounds. And this, too, despite the enormous political support for a united Europe.

4. Take agriculture. The record of regional trade blocs dealing with agricultural trade liberalisation is either non-existent or dismal; the CAP is not exactly the European Community's crowning achievement. In fact, if it were not for multilateralism (i.e., the Uruguay Round and the coalition of the Cairns Group that crystallised around the MTN), it is difficult to imagine that the process of unravelling the CAP could even have begun.

5. The (actual or potential) exercise of the regional option can also affect the efficacy of the multilateral one. The unwillingness of the European Community to start the MTN in 1982 and its largely reactive, rather than leadership, role at the Uruguay Round, are in some degree a reflection of its being less hungry for multilateralism given its internal market size and preoccupations. Then again, is it not evident that, were it not for the European Community, the capacity of the French (for whose political predicament one can only have sympathy, much as one deplores its consequence for the willingness to liberalise agriculture) to slow down the reform of the CAP and the liberalisation of world agriculture would have been significantly less?

6. Moreover, if regionalism is available as a realistic option, it will encourage exit rather than the seeking of voice and even the manifestation of loyalty to multilateralism:

• This may happen at the level of the bureaucrats who wind up preferring small-group negotiations among friends (code phrase: "like-minded people") to the intellectually and politically more demanding business of negotiating with and for the larger community of trading nations.

• Or else it may happen that, just as public choice theory à la Olson tells us in regard to the diffusion of consumer losses and concentration of producer gains that favour protectionist outcomes, the proponents of regionalism tend to be better focused and mobilised (they are often regional "experts" and partisans who ally themselves with the preferred policy options of the countries whose FTA cause they support), whereas the support for multilateralism is often more diffused and less politically effective and therefore takes second place when regionalism is on the political scene.

• Then again, regionalism may appeal to politicians since it translates more easily into votes: the wooing of the Hispanic voters, by urging them to identify with the FTA, was quite evident during the renewal of the fast-track authority in 1991 for the NAFTA negotiations with Mexico.

• The support of business groups for multilateralism may also erode with regional alternatives because of two different reasons: (i) If one can get a deal regionally, where one may have a "great deal of trade," then one may forget about the multilateral arena. Thus, if Canada could get the United States to agree to a fairer operation of the unfair trade mechanisms (a matter on which many Canadians today feel they were mistaken, with Prime Minister Mulroney and Mr. Riesman talking about Americans being "thugs" or like "third world dictators"[23]), why bother to fight the battles at the Uruguay Round where the powerful American manufacturing lobbies, zeroing in with the European Community against the Far East, seek instead to weaken the GATT rules? (ii) Again, one may get better protectionist, trade-diversionary deals for oneself in a preferential arrangement than in the non-discriminatory world of the GATT: e.g., Mexico's textile interests should benefit in the NAFTA relative to Caribbean and other external competitors in the US market, weakening the Mexican incentive to push for reform in the MFA forthwith.

7. Finally, it is true that the free-rider problem looks difficult as the number of GATT members increases steadily. Yet recent theoretical work on

GATT-style trade negotiations (Ludema, 1991) suggests that the free-rider problem may not be an effective barrier to freeing trade. Moreover, as Finger (1979) has pointed out, and as experience of inadequate GSP concessions underlines, developing countries have not been able to free-ride as much as their exemption from reciprocity under Special & Differential (S & D) treatment would imply: the trade concessions on commodities of interest to them have not gone as far as the concessions on commodities of interest to other GATT members without such an exemption. (Unconditional) MFN does not work in practice as well as it should from the free-riders' perspective.

1.6.2 Is Regionalism More Efficient?

Occasionally, one finds the regionalists arguing that regionalism is also more *efficient*: it produces *better* results. A typical argument is that, as part of the NAFTA negotiations, Mexico has accepted virtually all the US demands on intellectual property (IP) protection. A story, told in developing country circles, serves to probe this assertion critically: Ambassador Carla Hills was on a tour of South America, extolling the virtues of Mexico's "capitulation." At a dinner in her honour in Caracas, she apparently claimed: "Mexico now has world-class IP legislation." At this point, President Carlos Peretz supposedly turned to his left and remarked: "But Mexico does not have a world-class parliament."

The true moral of the story, however, is that, as part of the bilateral quid pro quos in an FTA or a CU, weak states may agree to specific demands of strong states,[24] in ways that are not exactly *optimal* from the viewpoint of the economic efficiency of the world trading system. In turn, however, these concessions can distort the outcome of the multilateral negotiations.

This may well have happened with TRIPs and TRIMs at the Uruguay Round.[25] As is now widely conceded among economists, the case for TRIPs for instance is *not* similar to the case for free trade: there is no presumption of mutual gain, world welfare itself may be reduced by any or more IP protection, and there is little empirical support for the view that "inadequate" IP protection impedes the creation of new technical knowledge significantly.[26] Yet the use of US muscle, unilaterally through "Special 301" actions, and the playing of the regional card through the NAFTA carrot for Mexico, have put TRIPs squarely and effectively into the MTN.

Again, a distorting impact on the multilateral trade rule from NAFTA negotiations can be feared from the fact that, as a price for the latter to be accepted by the Congress during the delicate renewal of fast-track authority, the US Administration had to accept demands for harmonisation in environment and labour standards by Mexico towards US standards. In political circles this effectively linked the case for free trade with the demands for "level playing fields" or fair trade (extremely widely interpreted),[27] legitimating these demands and weakening the ability of economists and of governments negotiating at the GATT (multilaterally for arm's length free trade) to resist this illegitimate constraint on freeing trade.[28]

1.6.3 Is Regionalism More Certain?

Much has been made, in the Mexican context, of the argument that the FTA will make trade liberalisation irreversible. But something needs to be added here:

- GATT also creates commitments: tariffs are bound. (This does not apply to concessions made under conditionality, of course, by the IMF or the IBRD.) Mexico *is* a member, if recent, of the GATT.
- Recall Dam (quoted above): Article XXIV is so full of holes in its discipline that almost anything goes. Reductions of trade barriers can be slowed down, as "circumstances" require, other bindings can be torn up by mutual consent (an easier task when there are only a few members in the bloc but more difficult under the GATT), etc.
- Recall, too, that regional agreements have failed (LAFTA) and stagnated (ASEAN) as well. The current mood in Canada over NAFTA is sour and the MTN looks better in consequence.[29] The sense, however, that the United States has let Canada down and failed to live by the spirit of the FTA agreements will probably not endure. But who knows?

1.7 The United States: From "Piecemeal" to "Programmatic" Regionalism

Let me conclude by considering more specifically the US shift to regionalism for the Americas in the perspective of the objective of arriving at (non-discriminatory) free trade for all.

US regionalism, when presided over by Ambassador William Brock, then the US Trade Representative, was *not* geographically-circumscribed

regionalism. Rather, it was truly open-ended. Brock was known to have offered an FTA to Egypt (along with the one to Israel) and to the ASEAN countries; indeed, he would have offered it to the moon and Mars if only life had been discovered there with a government in place to negotiate with. This regionalism was evidently motivated by a vision, even if flawed,[30] that saw regionalism as clearly the route to multilateralism: it would go on expanding, eventually embracing many, preferably all.

By contrast, today's regionalism, confined to the Americas by President Bush's men, lacks the "vision thing." In fact, when allied with Secretary Baker's recently reported admonition to the Japanese not to encourage an Asian trade bloc, as suggested by Malaysia as a necessary response to the European Community and US regionalism, the US policy appears to Asia also to be self-contradictory and self-serving: "regional blocs are good for us but not for you." And it simply won't wash, though Japan, fearing further bashing, will be deterred for a while.

If America's regionalism is not to turn into a piecemeal, world trading system-fragmenting force, it is necessary to give to it a programmatic, world trade system unifying format and agenda. One possibility is to encourage, not discourage, Japan to line up the Asian countries (all the way to the Indian subcontinent) into an AFTA, with the US lining up the South Americans into the NAFTA, on a schedule, say, of 10 years. Then, Japan and the United States, the two "hubs," would meet and coalesce into a larger FTA at that point,[31] finally negotiating with the European Community and its associate countries to arrive at the Grand Finale of multilateral free trade for all in Geneva.

Only such "programmatic" regionalism, in one of several possible variants, would ensure that US regionalism was not perceived by Asia to be hostile and fragmenting.[32] It alone would make regionalism less harmful to the MTN and the GATT and more supportive of the cause of multilateral free trade for all.

1.8 Concluding Remarks

The question of regionalism is thus both a difficult and delicate one. Only time will tell whether the revival of regionalism since the 1980s will have been a sanguine and benign development or a malign force that will serve to undermine the widely-shared objective of multilateral free trade for all.

My judgement is that the revival of regionalism is unfortunate. But, given its political appeal and its likely spread, I believe that it is important to contain and shape it in the ways sketched here so that it becomes max-

imally useful and minimally damaging, and consonant with the objectives of arriving at multilateral free trade for all.

Notes

This chapter reflects my personal views and bears no relationship to my position as Economic Policy Adviser to the Director-General, GATT. Thanks are due to Robert Baldwin, James Benedict, Richard Blackhurst, Christopher Bliss, Don Davis, Sunil Gulati, Douglas Irwin, John McMillan, Arvind Panagariya, T. N. Srinivasan and John Whalley for helpful conversations and suggestions.

Editorial Note: A figure and corresponding discussion that appeared in the original chapter have been omitted from section 1.4 to avoid overlap with chapter 2 (this volume).

1. The Vinerian approach to customs union theory has been carried forward by others more recently, chiefly by Berglas (1979) and Corden (1976). In addition, three alternative theoretical approaches can be distinguished: by Kemp and Wan (1976): by Cooper and Massell (1965a, 1965b), Johnson (1965) and Bhagwati (1968); and by Brecher and Bhagwati (1981). All four approaches are distinguished and discussed in the graduate textbook by Bhagwati and Srinivasan (1983, Chapter 27) and in Bhagwati (1991a). Each is touched upon later in this chapter.

2. Two points should be noted. First, there is a difference between intention and reality: as argued below, the Article XXIV-sanctioned FTAs and CUs have never gone "all the way." Second, GATT's MFN is universal only for its members, so it falls short of total universalism. but the important point to remember is that the GATT is open to membership to all who meet the criteria for admission, and has generally been inclusive rather than exclusive.

3. Of course, this theory developed *after* the incorporation of Article XXIV into the GATT. So its inconsistency with Article XXIV, on its own terms, is perhaps only an amusing observation. Note, however, that James Meade was a main actor in both. The argument is developed in two alternative ways in Lipsey (1960, p. 507) and in Johnson (1967, p. 203).

4. A substantially improved and more effectively functioning dispute settlement mechanism, aimed at restoring GATT's legal discipline, is an important part of the 1992 "Dunkel draft" of what the Uruguay Round should conclude.

5. Japan in fact, appears to have probed the possibility of going into such an arrangement with the United States as one of its partners in the 1960s but to no avail.

6. The question of "multilateralism" versus "regionalism" surfaced at a different level even within this preferential trade liberalisation among the developing countries. Thus, in the early 1960s, we were discussing whether the Cooper–Massell–Johnson–Bhagwati argument should not be considered on a G-77-wide basis rather than for much smaller groups of developing countries. This was the main issue before a 1962 UNCTAD Expert Group in New York, of which I was a member, which met over three weeks to draft the recommendation that preferential trade liberalisation among the developing countries be "multilateral," i.e., G-77-wide, rather than narrowly focused. Unfortunately, the preferential arrangements that were contemplated took the latter, narrower focus.

7. The MIT Economics Department is at 50 Memorial Drive in Cambridge, Massachusetts. I obviously exclude the diaspora, including myself! If the views expressed with Dornbusch in an Eastman Kodak publication (Dornbusch et al., 1989) are a guide. Krugman may hold one of the positions described above. This pamphlet makes somewhat odd and untenable

statements about what the GATT does and does not do. Cf. Finger's (1989) rather blunt analysis of these assertions in *The World Economy* and my own complaints about the confusions following from loose writing on trade-policy issues, and the resulting prostitution of an important debate, in Bhagwati (1991b). Whether the Memorial Drive school has by now under fire shifted its anti-multilateral stance and joined the more common view that regionalism is a useful supplement, not an alternative, to multilateralism is anyone's guess, given the conflicting reports one hears of its many oral pronouncements on the lecture circuit from its peripatetic members. But if it truly has abandoned its early vitriolic anti-GATT position, I would be delighted in its demise.

8. I rely upon oral presentation at the 1988 annual meeting of the American Enterprise Institute in Washington, DC.

9. Defined in Vinerian fashion, a trade-diverting FTA can still improve a member country's welfare but will generally harm outside countries. The focus below is on the impact on others, as is presumably the intention also of Article XXIV's injunction not to raise the average external tariff.

10. Aside from obfuscating the distinction between preferential and non-discriminatory trade liberalisation, the pro-FTA economists got carried away by the "battle for Mexico." Thus, while it is perfectly possible for Mexico to gain much while the United States gains little, a *Wall Street Journal* article by Dornbusch (1991) argued that trade with Mexico was already largely free because of OAP provisions and the GSP (so that the union fears of job losses, etc. were exaggerated), and simultaneously that Mexico would achieve prosperity thanks to the FTA. It is, of course, possible to argue each position separately in the "segmented markets" of Mexico City and Washington, DC, turning arguments on their head as necessary for one's case. But it takes chutzpah to make the contradictory arguments in the same article.

11. I must confess that I had not come across this prescription earlier. But in a report in the *Economic Focus* column in *The Economist* in 1991, of a Jackson Hole Conference on FTAs, it was attributed to Paul Krugman and Lawrence Summers.

12. See Lipsey (1960, pp. 507–8): "As far as the prices of the goods from a country's union partner are concerned, they are brought into equality with rates of transformation vis-à-vis domestic goods, but they are moved away from equality with rates of transformation vis-à-vis imports from the outside world. These imports from the union partner are thus involved in both a gain and a loss and their size is per se unimportant; what matters is the relation between imports from the outside world and expenditure on domestic commodities: the larger are purchases of domestic commodities and the smaller are purchases from the outside world, the more likely is it that the union will bring gain. Consider a simple example in which a country purchases from its union partner only eggs while it purchases from the outside world only shoes, all other commodities being produced and consumed at home. Now when the union is formed the 'correct' price ratio (i.e., the one which conforms with the real rate of transformation) between eggs and shoes will be disturbed but, on the other hand, eggs will be brought into the 'correct' price relationship with all other commodities—bacon, butter, cheese, meat, etc., and in these circumstances a customs union is very likely to bring gain, for the loss in distorting the price ratio between eggs and shoes will be small relative to the gain in establishing the correct price ratio between eggs and all other commodities. Now, however, let us reverse the position of domestic trade and imports from the outside world, making shoes the only commodity produced and consumed at home, eggs still being imported from the union partner, while everything else is now bought from the outside world. In these circumstances the customs union is most likely to bring a loss: the gains in establishing the

correct price ratio between eggs and shoes are indeed likely to be very small compared with the losses of distorting the price ratio between eggs and all other commodities."

13. In any event, by encouraging rules of origin because the trade-barrier walls everywhere are not equally high, FTAs encourage in turn the bureaucratic-cum-industry capture of the essentially arbitrary "local content" rules for protectionist purposes.

14. VERs are evidently selective by countries; AD actions are selective down to the level of the firm, as Brian Hindley has often noted.

15. Brian Hindley and Patrick Messerlin are investigating this hypothesis for the GATT Secretariat as part of a set of studies to support the 1992 GATT *Annual Report* on *Regionalism and Multilateralism,* following the 1991 *Annual Report* on *Trade and the Environment.*

16. This paper has also been commissioned by the GATT Secretariat for its 1992 *Annual Report.*

17. Christopher Bliss (1990) has recently made a valuable stab at this problem.

18. Such time-paths are clearly not unique. Thus, for instance, any number of such paths could be generated by relaxing the requirement that, at each stage, the non-union outside countries be left only as well off as before the new expansion of the CU.

19. This argument is being investigated in Mansfield's forthcoming paper for the 1992 GATT *Annual Report.*

20. This analysis must use the Brecher and Bhagwati (1981) approach to theorising about CUs since it relates to analysing the effects of changes in domestic and external policies and parameters on the distribution of income and welfare among member states.

21. Quoted by Wolf (1989).

22. Again, at the instance of the GATT secretariat, this question will be investigated in depth for the 1992 GATT *Annual Report,* by Bernhart Hoekman with Michael Leidy, and by Edward Mansfield.

23. Those who think that much of Japan-bashing is not prejudiced may want to think about the differential and exaggerated reaction in the United States to the far more innocuous remarks of Prime Minister Miyazawa and Speaker Sakarauchi.

24. In Mexico's case, President Salinas's political stake in getting an FTA with the United States is vastly disproportionate to President Bush's.

25. TRIPs are trade-related IP provisions and TRIMs are trade-related investment measures. The weakness of the case for their inclusion in the GATT, at least in the forms canvassed by many lobbies, is discussed in Bhagwati (1991a).

26. It is not surprising therefore that the spokesman for TRIPs have shifted from utilitarian methods of argumentation to "rights": they talk now of "theft" and "piracy."

27. That the environmental and labour standards' negotiations in NAFTA will be "parallel" rather than "integrated" is of no consequence, any more than running the services negotiations parallel to other negotiations at the Uruguay Round has been.

28. The danger posed by the proliferating demands for "level playing fields" or fair trade, chiefly in the United States but elsewhere too, is extremely serious. It is analysed, and the theoretical questions raised by it are noted, in Bhagwati (1992). The environment issue, in particular, has been discussed in this context in the 1991 GATT *Annual Report.*

29. Whalley's splendid study of the US–Canada FTA (1992) supports the sceptical views that I have advanced of the prospect and wisdom of the "Second Regionalism."

30. For reasons that I have already indicated above, regionalism is not quite the benign trade policy that it is now popularly believed to be.

31. This would require discarding the extreme Japanophobia that characterises the so-called "revisionists" who are really "regressionists" twice over: they use simple-minded regressions to condemn Japan for its "closed markets" (e.g., that Japan's manufactures' import share is stagnant and/or low compared to others') and they also wish to return the United States to the Japan-bashing of the prewar period that had given way to sense and sensitivity in the postwar years: cf. Bhagwati (1991a).

32. Saxonhouse's excellent study (1992) only complements and underlines what I argue here. I should add that while the United States signals a world trading system fragmenting message to Asia through NAFTA, Mexico by contrast signals a pro-world trade message. In joining in free trade with the colossus to its north, President Salinas boldly and effectively tells the developing countries that free trade is good and not to be feared.

References

Berglas, E. (1979) "Preferential Trading Theory: The *n* Commodity Case," *Journal of Political Economy*, **87**, pp. 315–31.

Bhagwati, J. N. (1968) "Trade Liberalization Among LDCs, Trade Theory and GATT Rules," in J. N. Wolfe (ed.), *Value, Capital, and Growth: Papers in Honour of Sir John Hicks*, Oxford: Oxford University.

——— (1991a) *The World Trading System at Risk*, Princeton: Princeton University Press and Harvester Wheatsheaf.

——— (1991b) "Revealing Talk on Trade," *The American Enterprise*, **2(6)**, pp. 72–8.

——— (1992) "Fair Trade, Reciprocity and Harmonization: The New Challenge to the Theory and Policy of Free Trade," Columbia University, Economics Department, *Discussion Paper Series*, **604** (April).

Bhagwati, J. N. and T. N. Srinivasan (1983) *Lectures on International Trade*, Cambridge, Mass.: MIT Press.

Bliss, C. (1990) "The Optimal External Tariff in an Enlarging Customs Union," CEPR, *Discussion Paper*, 368 (February), London: CEPR.

Brecher, R. and J. N. Bhagwati (1981) "Foreign Ownership and the Theory of Trade and Welfare," *Journal of Political Economy*, **89(3)**, pp. 497–512.

Cooper, C. A. and B. F. Massell (1965a) "A New Look at Customs Union Theory," *The Economic Journal*, **75**, pp. 742–7.

——— (1965b) "Towards a General Theory of Customs Unions for Developing Countries," *Journal of Political Economy*, **73(5)**, pp. 461–76.

Corden, W. M. (1976) "Customs Union Theory and the Nonuniformity of Tariffs," *Journal of International Economics*, **61(1)**, pp. 99–107.

Dam, K. (1970) *The GATT: Law and Intenational Economic Organization*, Chicago: University of Chicago Press.

Dornbusch, R. (1991) "If Mexico Prospers, So will We," *The Wall Street Journal* (April 11) Op.-Ed. section.

Dornbusch, R. et al. (1989) *Meeting World Challenges: United States Manufacturing in the 1990s*, Rochester, NY: pamphlet issued by the Eastman Kodak company.

The Economist (1991) Economics Focus column (July).

Finger, J. M. (1979) "Trade Liberalization: A Public Cholice Perspective," in R. C. Amachen, G. Haberler and T. Willett (eds.), *Challenges to a Liberal International Economic Order*, Washington, DC: American Enterprise Institute.

——— (1989) "Picturing America's Future: Kodak's Solution of American Trade Exposure," *The World Economy*, **12(3)**, pp. 377–80.

Gowa, J. and E. Mansfield (1991) "Allies, Adversaries, and International Trade," paper presented to the American Political Science Association Meetings, Washington, DC (mimeo).

Grandmont, J. M. and D. McFadden (1972) "A Technical Note on Classical Gains from Trade," *Journal of International Economics*, **2**, pp. 109–25.

Irwin, D. (1993) "Multilateral and Bilateral Trade Policies in the World Trading System: An Historical Perspective," chap. 4 in J. de Melo and A. Panagariya (eds.), *New Dimensions in Regional Integration*, Cambridge, UK: World Bank and Cambridge University Press.

Johnson, H. G. (1958a) "The Gains from Free Trade with Europe: An Estimate," *Manchester School of Economic and Social Studies*.

——— (1958b) "The Economics Gains from Free Trade with Europe," *Three Banks Review*.

——— (1965) "An Economic Theory of Protectionism, Tariff Bargaining, and the Formation of Customs Unions," *Journal of Political Economy*, **73** (June), pp. 256–83.

——— (1967) *Eonomic Policies Toward Less Developed Countries*, Washington, DC: Brookings Institution.

Kemp, M. C. (1972) "The Gains from International Trade," *The Economic Journal*, **72**, pp. 803–19.

Kemp, M. C. and H. Wan (1976) "An Elementary Proposition Concerning the Formation of Customs Unions," *Journal of International Economics*, **6** (February), pp. 95–8.

Lipsey, R. G. (1957) "The Theory of Customs Unions: Trade Diversion and Welfare," *Economica*, **24**, pp. 40–6.

——— (1958) "The Theory of Customs Unions: A General Equilibrium Analysis," University of London, Ph.D. thesis.

——— (1960) "The Theory of Customs Unions: A General Survey," *The Economic Journal*, **70**, pp. 498–513.

Lipsey, R. G. and K. J. Lancaster (1956–7) "The General Theory of Second Best," *Review of Economic Studies*, **24**, pp. 33–49.

Ludema, R. (1991) "International Trade Bargaining and the Most Favoured Nation Clause," *Economics and Politics*, **3(1)**, pp. 1–41.

Mansfield, E. (1992) "The Concentration of Capabilities and International Trade," *International Organization* (forthcoming).

McMillan, J. (1991) "Do Trade Blocs Foster Open Trade?," University of California at San Diego (mimeo).

Meade, J. E. (1956) *The Theory of Customs Unions*, Amsterdam: North-Holland.

Samuelson, P. A. (1939) "The Gains from International Trade," *Canadian Journal of Economics and Political Science*, **5(2)**, pp. 195–205.

Saxonhouse, G. R. (1992) "Trading Blocs and East Asia," chap. 12 in J. de Melo and A. Panagariya (eds.), *New Dimensions in Regional Integration*, Cambridge, UK: World Bank and Cambridge University Press.

Viner, J. (1950) *The Customs Union Issue*, New York: Carnegie Endowment for International Peace.

Whalley, J. (1993) "Regional Trade Arrangements in North America: CUSTA and NAFTA," chap. 11 in J. de Melo and A. Panagariya (eds.), *New Dimensions in Regional Integration*, Cambridge, UK: World Bank and Cambridge University Press.

Wolf, M. (1989) "European Community 1992: The Lure of the *Chasse Gardée*," *The World Economy*, **12(3)**, pp. 373–6.

2

Preferential Trading Areas and Multilateralism—Strangers, Friends, or Foes?

Jagdish Bhagwati and
Arvind Panagariya

The question of Preferential Trading Areas, as we should call them in preference to Free Trade Areas and customs unions, phrases that falsely equate them in the public mind and discourse with nonpreferential free trade, has not been distant from international economists' thoughts and concerns since the beginning of the postwar period when the architects of the General Agreement on Tariffs and Trade had to confront PTAs and accommodate them into the GATT via Article XXIV.[1]

Their wisdom became a center of analytical attention, especially at the time of the steps taken to form the European Community by the Treaty of Rome in 1957 and when, in what Bhagwati (1991) has called the period of First Regionalism, other Article XXIV-sanctioned PTAs were considered and even attempted in other areas.[2] The theory of PTAs of Viner (1950)—to which Meade (1955), Lipsey (1957, 1960), and other international economists at the time made important contributions—while preceding the formation of the European Community, developed more fully as a result of that singular event. The attempts at providing a more realistic rationale for the extension of such PTAs to developing countries, on the other hand, as a way of reducing the cost of any targeted level of industrialization, came from Cooper and Massell (1965a, 1965b), Johnson (1965), and Bhagwati (1968) at the time.[3]

It must be said that the First Regionalism was stillborn; beyond the European Community (and its offshoot, the European Free Trade Association), there was practically no successful emulation of the European developments elsewhere. At the same time, given the fact that it arose over the concerns that such PTAs were not the same as nondiscriminatory freeing of trade, the Vinerian theory was "static," concerning itself simply with the

Originally published, as chapter 1, in *The Economics of Preferential Trade Agreements*, ed. Jagdish Bhagwati and Arvind Panagariya (Washington, DC: American Enterprise Institute, 1996), 1–78. Reprinted with permission.

issue as to when such PTAs would be trade-diverting or trade-creating, thus diminishing or increasing welfare.

The recent revival of interest in the theory of Preferential Trading Areas, marking what Bhagwati (1991) has christened the Second Regionalism, has come instead from the conversion of the United States to preferential trading arrangements, starting with the Canada-U.S. Free Trade Agreement (CUFTA) and the later extension to include Mexico under the North American Free Trade Agreement (NAFTA). This time around, the movement has extended equally to other areas, involving again developing countries on their own, as in the Southern Cone Common Market (MERCOSUR), but with success rather than failure.

In 1982, the United States could not get multilateral trade talks started at Geneva and hence turned to ever-expanding PTAs as an alternative way of getting eventually to worldwide free trade. This has given the theory of PTAs a "dynamic time-path" dimension (Bhagwati 1993a). When would such an approach lead to a progressive freeing of trade barriers through expanding membership (and/or accelerating multilateral trade negotiations in a benign symbiosis)? This is also a political economy-theoretic question, fitting nicely into the modern preoccupation of economic theorists with questions relating to what policies emerge (that is, with "public choice") rather than what they should be (that is, with "social choice").

From a policy viewpoint also, this revival of PTAs is an important development. It was fed (if not led) by the U.S.-centered NAFTA and its proposed extension to Chile and beyond, and by Asia-Pacific Economic Cooperation (APEC), which some in the United States would like to see turn into another PTA, and by the call of European politicians such as Foreign Minister Klaus Kinkel of Germany at the outset, and by many others subsequently, to form TAFTA (a Transatlantic Free Trade Area). With WTO jumpstarted and multilateralism functioning, the theoretical and policy questions then must be confronted: should these proposals for proliferating PTAs, especially when inclusive of hegemonic powers such as the United States, be encouraged by economists?

In this chapter we undertake the following tasks. After reviewing key phrases and concepts, we extend the "static" analysis of PTAs. This enables us to examine several recent claims in favor of PTAs and persuades us to discard them as unpersuasive.

Specifically, our analysis enables us to examine and reject the much-cited claim that it is wrong to worry about trade diversion and that PTAs are generally as good as nonpreferential trade liberalization.

Our analysis gives added insight into why the usual argument made these days is mistaken. This is the argument that when countries joining a PTA have large shares of their trade with one another and are thus "natural trading partners," they need not fear losses. The nonhegemonic countries that are liberalizing with a hegemon that is generally open and offering few new reductions of trade barriers, as is the case with Mexico and with other potential NAFTA members outside of the United States and Canada, could face the prospect of significant "static" welfare losses.

Next, we turn to the dynamic time-path question. In the policy context, this necessitates our considering arguments as to why a proliferation of PTAs, despite their creating a harmful "spaghetti-bowl" phenomenon in the world economy, may be beneficial because of their helpful consequences for the progressive freeing of trade and moving the world economy to worldwide free trade.

We systematize the current analytical contributions on this problem and evaluate the current policy developments.[4] It is our view that PTAs that are hegemon centered, as NAFTA is, are not the desirable way to advance the cause of worldwide freeing of trade barriers and that it is better to focus on WTO-centered MFN trade liberalization. By contrast, we consider intradeveloping-country, non-hegemon-centered PTAs, such as MERCOSUR, in a more favorable light. First, however, we need to clarify a few central phrases and concepts.

Phrases and Concepts

Two phrases are frequently used: PTAs and regionalism. The two significant concepts are, first, trade creation and diversion and, second, stumbling and building blocks.

Preferential Trading Areas

This term refers to FTAs, CUs (which also have a common external tariff), and Common Markets (which additionally have freedom of internal factor movement within the area defined by member states). All these arrangements fall within the purview of GATT Article XXIV. Lesser forms have traditionally been permitted for developing countries and come under Economic Cooperation among Developing Countries (ECDC). We will have something to say about that too, though our chapter will be almost exclusively focused on Article XXIV-sanctioned PTAs and, within that category, on FTAs in particular.

Regionalism

This term has been loosely used by many, including us, as synonymous with PTAs. Strictly speaking, however, regionalism refers to PTAs defined by a geographic region. There is a school of thought (to which Lawrence Summers and Paul Krugman have subscribed) that considers regional PTAs to be a priori less likely to lead to static trade diversion than non-regional PTAs and such regional PTAs to be therefore ipso facto acceptable. This is a substantive issue that we will consider, as did Bhagwati (1993a) in a preliminary way. Our focus, however, will be on PTAs, not regional PTAs.

Trade Creation and Trade Diversion (Viner)

The concepts of trade creation and trade diversion as two possibilities that define the second-best nature of the static analysis of PTAs go back to Viner (1950), of course. While there are various ways in which these two concepts have subsequently been defined, we will use them (in the theoretical analysis below) in the original Vinerian sense to mean a shift of imports from an efficient to an inefficient source under trade diversion, and a shift from an inefficient to an efficient source under trade creation.[5]

"Stumbling Blocks" and "Building Blocks" (Bhagwati)

The phraseology and conceptualization of PTAs that, in a dynamic time-path sense, contribute to the multilateral freeing of trade either by progressively adding new members (down the PTA path to worldwide free trade) or by prompting accelerated multilateral trade negotiations and are thus *building blocks* toward the multilateral freeing of trade and those that do the opposite and hence are *stumbling blocks* to the goal of worldwide multilateral freeing of trade, owes to Bhagwati (1991, 77) and has been adopted by Lawrence (1991) and others.[6] Insofar as Viner's trade creation and trade diversion concepts were designed to divide PTAs into those that were good and those that were bad in the static sense, Bhagwati's building block and stumbling block concepts are designed to divide PTAs into those that are good and those that are bad in the dynamic, time-path sense.

Rethinking Static Welfare Analysis

We now begin with the static analysis. Frankly, so much has been written on the static analysis since Viner's pioneering 1950 contribution, indeed by virtually every important international economist, that one may think that there is little to add.

The Issues Examined

Yet there is something to be gained by another, close look at the conventional static analysis in view of several presuppositions, mostly favorable to PTAs, which have recently been made by policy analysts.

It has been forcefully argued by Summers (1991, 299) in an influential paper that international economists should not be preoccupied by trade diversion: "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."

Our first reaction is to deny the premise of his analogy: economists, faced with a second-best problem, typically *do* worry about that aspect of the problem. Indeed, if the world was first best, market prices would reflect social opportunity costs, and there would be no need for cost-benefit analysis for projects. The World Bank, where Summers served with distinction, would then have to close down most of its project-lending research and analysis aimed at determining the shadow prices to be used in judging the acceptability of projects.

Second, the problem of preferential trade liberalization is indeed an inherently second-best problem since nondiscriminatory trade liberalization is being ruled out. Ignoring this aspect is unwarranted.

Third, one should not confuse "second best" with "primary impact." First-best problems also are characterized by primary and total effects.

Fourth, if Summers implies that trade-diverting PTAs are a minor nuisance, he is misled presumably by the fact that efficiency losses are Harberger triangles and "small." But such PTAs impose losses on member countries also through tariff-revenue-*redistribution*, and these can be large: they are rectangles, while the efficiency effects are triangles.

We also consider the contention in the recent policy debate that countries that trade with each other in larger volume than with other nations are "natural" trading partners and hence that PTAs among them are likely to be welfare enhancing to their members for that reason.

This contention is further linked with the argument that "regional" PTAs are desirable (in the sense of being more likely to create welfare gains for their members) because geographically contiguous countries (particularly if they share common borders) have larger volumes of trade with one another than with others.

Our analysis here challenges the premise that large volumes of initial trade lessen the likelihood of loss from PTAs. Consequently, it also undermines the associated contention that regional PTAs are more desirable.

We also question the alternative but related "natural trading partners" hypothesis that regional PTAs are likely to improve welfare by conserving on transport costs. We show that transport costs by themselves do not provide a reason for discriminatory PTAs.[7]

The Theoretical Analysis

Since Viner's classic work in 1950, PTAs have been considered to be harmful both to member countries (whose imports are the subject of the trade diversion) and to the world when trade diversion arises, and to be welfare enhancing when trade creation occurs instead. This ambiguity of outcomes, depending on the relative strengths of the two effects when a PTA is formed, has been the principal reason for the debate among economists as to whether a specific PTA is desirable.

We will begin the theoretical analysis below by showing, however, that the conventional trade creation and trade diversion are not the entire story in deciding on the welfare outcome for an *individual* member of a PTA. Even if trade creation effects are larger than trade diversion effects so that the union as a whole benefits, an individual member could lose on account of adverse income distribution effects arising from tariff revenue redistribution.

The redistribution of tariff revenue between member countries arises, of course, from a shift in the terms of trade within the union. When a member country lowers its tariff on the partner without lowering it on the rest of the world, within-union terms of trade shift in favor of the partner (for both existing and new imports from it). The extent of the unfavorable redistributive effect on a member country is obviously determined by the degree of preferential access it gives to the partner country in relation to the preferential access it receives from the latter: the greater the margin of preference the country gives, the more it stands to lose. This implies that when a country with a high degree of protection forms a PTA with a country with relatively open markets, as in the case of Mexico and the

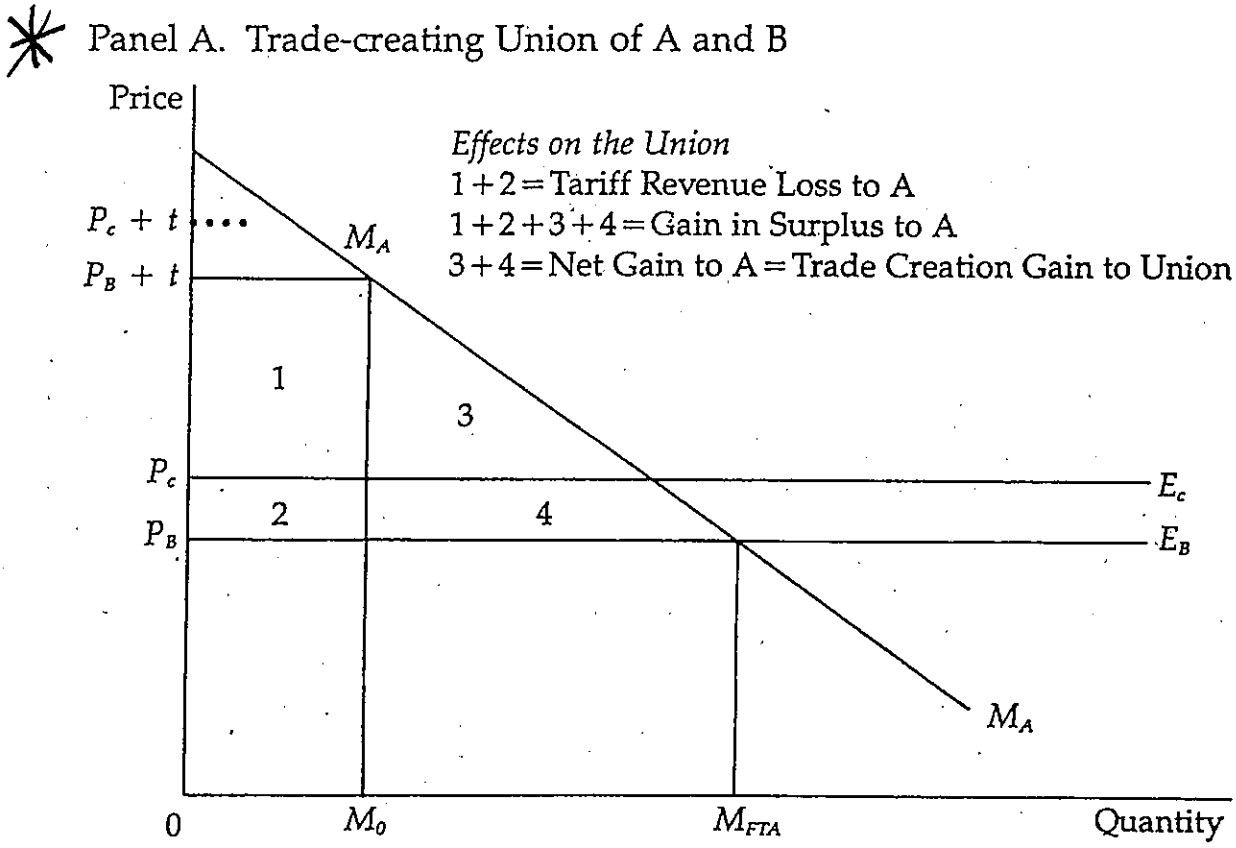
United States, the former may well be faced with a net welfare loss. We develop this theme and its ramifications, in the following analysis, using simple models from the literature and distinguishing clearly among the effects on member country and world welfare.

The Viner Model—Constant Costs

The natural starting point for explaining the economics of regional integration is Viner's partial equilibrium model. This model does not fully capture the effects noted above but is, nevertheless, an important step toward understanding them. Assume that there are three countries, A, B, and C. Countries A and B are potential union partners and C represents the rest of the world. In figure 2.1, panels a and b, let M_AM_A represent A's import demand for a specific product and P_BE_B and P_CE_C the (export) supplies of the same product available from B and C, respectively. Following Viner, it is assumed that the supply prices of B and C are constant at P_B and P_C, respectively. In panel a, the supply price of C exceeds that of B and in panel b the opposite is true.

In panel a illustrating the case of a *trade-creating union*, with an initial nondiscriminatory specific tariff t, A imports OM_0 quantity of the good.[8] All imports come from B so that A raises areas 1 and 2 in tariff revenue. If A now forms an FTA with B, imports from B expand from OM_0 to OM_{FTA}. The tariff revenue disappears, but the price facing consumers declines by t; A's consumers capture the entire revenue in the form of increased surplus. Because B is the lower cost source of the product, there is positive trade creation and no trade diversion.[9] Working like nondiscriminatory free trade, the FTA yields to A and to the union a net gain represented by areas 3 and 4.[10]

Panel b illustrates the case of a *trade-diverting union*. Here B is the higher cost source of the product with the result that, given a nondiscriminatory tariff in A, all imports come from C. A imports OM_0 and collects areas 1 and 2 in tariff revenue. If A and B now form an FTA, imports expand to OM_{FTA}, but the source of their supply switches from C to B. Though the reduction in A's domestic price leads to some trade creation—increased imports lead to a displacement of some inefficient domestic production and increase in consumption in A—the switch to the higher cost source, B, leads to a large trade diversion of OM_0 quantity of imports from C to B. Thus, panel b shows a case where the union diverts trade from C, but it also creates some trade. The gains to A are given by area 3 and the losses by area 2. The loss of area 2 results from a deterioration in A's terms of



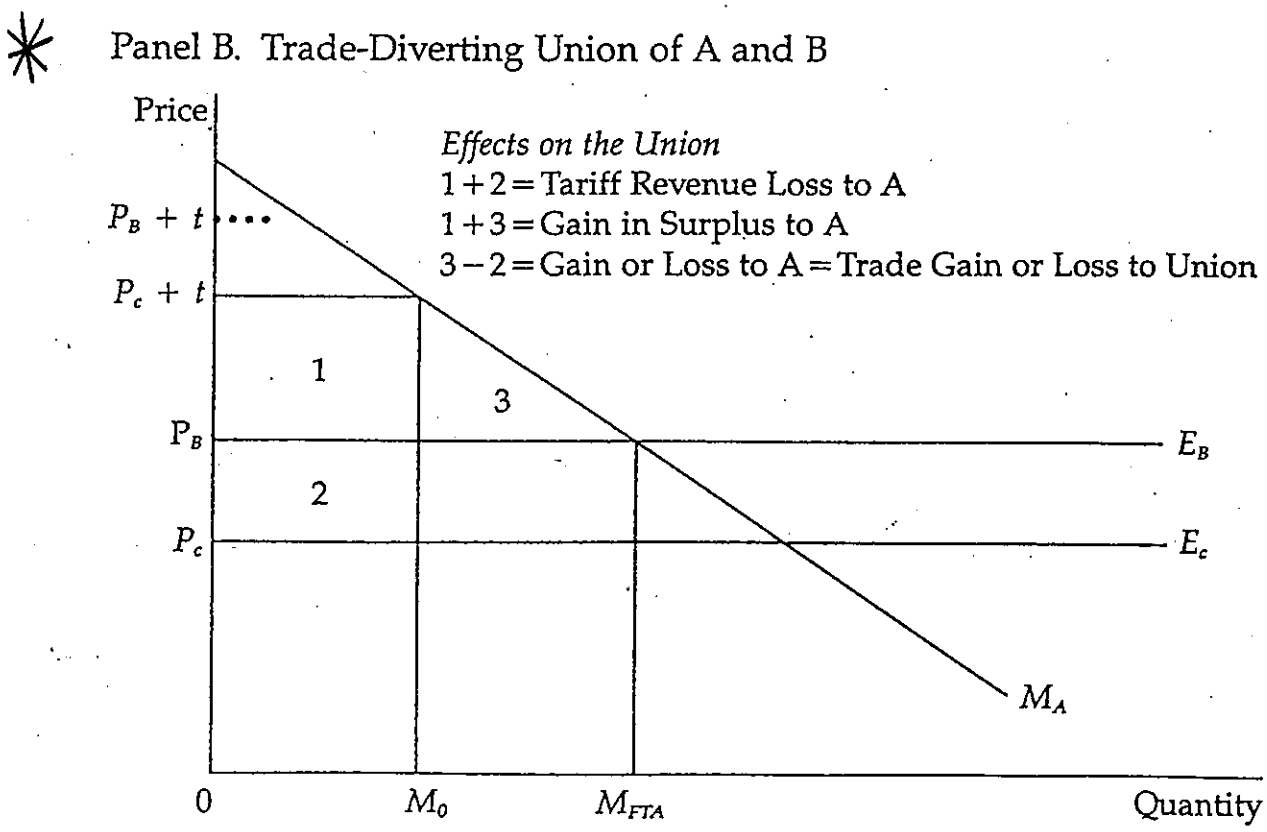


Figure 2.1 (panels A, B, and C)
Constant costs, according to strictly Vinerian analysis.

Panel C. General Equilibrium Analysis: Viner-Lipsey Model of Trade-Diverting Union

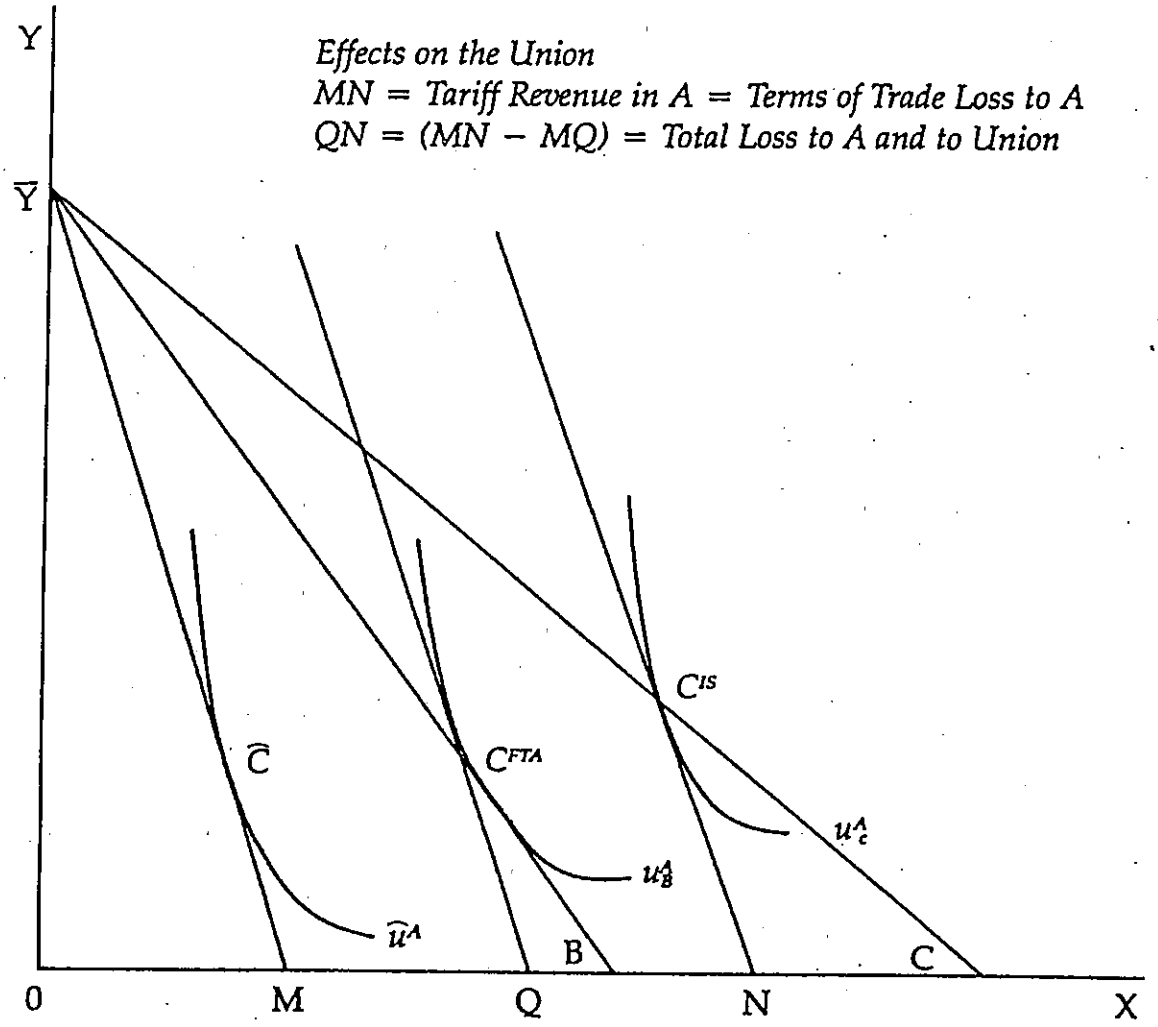


Figure 2.1 (continued)

trade from P_C to P_B and takes the form of the excess of the loss of tariff revenue (1 + 2) over that which is captured partially (1) by A's consumers. Area 2 goes to pay for the higher cost of production in B than in C.

Now, unless cost differences between B and C are small, areas similar to area 2 will be large in relation to the triangular areas of gain. The welfare loss to A from the loss of revenue on diverted imports applies to the *entire* initial quantity of imports, whereas the gain applies only to the *change* in the quantity of imports. The FTA will be associated with trade creation in some sectors and trade diversion in others. But since losses the likely to be large in cases involving trade diversion, trade diversion in even a few sectors can more than offset the gains arising from trade creation in a large number of sectors.

The trade-diverting case of panel b can also be illustrated in general equilibrium by using the Lipsey (1958) version of the Viner analysis as in

panel c. There, the economy of A is specialized in producing at $\bar{Y}$, with $\bar{Y}C$ and $\bar{Y}B$ the given, fixed terms of trade with C and B, respectively. With an initial nondiscriminating tariff, A trades with C and consumes at C^{IS}, winding up with welfare at U_C^A. With the FTA between A and B, the trade shifts to B. A winds up consuming at C^{FTA} and its welfare is reduced to U_B^A. The welfare loss QN can then be seen as the difference between the tariff-revenue or terms-of-trade loss MN and the gain MQ that comes from the ability to shift consumption from $\hat{C}$ to C^{FTA}.[11] (OM is the income at domestic prices in the initial situation, and tariff revenue is MN, the sum of the two yielding ON as national expenditure.)

A final and obvious point may be stressed concerning nondiscriminatory trade liberalization by country A. In both the cases shown in figure 2.1, A obtains maximum trade gains and its welfare is improved relative to the initial as well as the FTA equilibrium by a nondiscriminatory liberalization. Such liberalization leads to the same equilibrium in the trade-creating union in panel a (as a limiting case) and eliminates trade diversion in the case in panel b, amounting to free trade with the most efficient supplier for each commodity.

Partner Country's Supply Curve Is Upward Sloped

Because of the assumption that the export-supply curves of both B and C are perfectly elastic, the model in figure 2.1 leads to at least two unrealistic outcomes.[12] First, imports into A come from either B or C but not both. Second, in the trade diversion case, the losses of A represented by area 2 are used up entirely to finance B's higher costs of production: the partner country B makes no gain whatsoever. The model thus captures only one side of the possibly "mercantilist" nature of trade-diverting FTAs: country A can lose from its own (discriminatory) trade liberalization, but country B does not gain from it.

A more realistic model allows the supply curve of one or both countries B and C to slope upward. In the interest of simplicity, we will allow for an upward-sloped supply curve for only one country at a time. Figure 2.2 takes up the case when the partner country B's supply curve slopes upward and that of the outside country C is horizontal. This case captures the essence of the more general model in which the outside country's supply curve also slopes upward but is more elastic than the partner's. Figure 2.3 shows instead the case when the union is between A and C so that the partner country's supply curve is more elastic than that of the outside world.

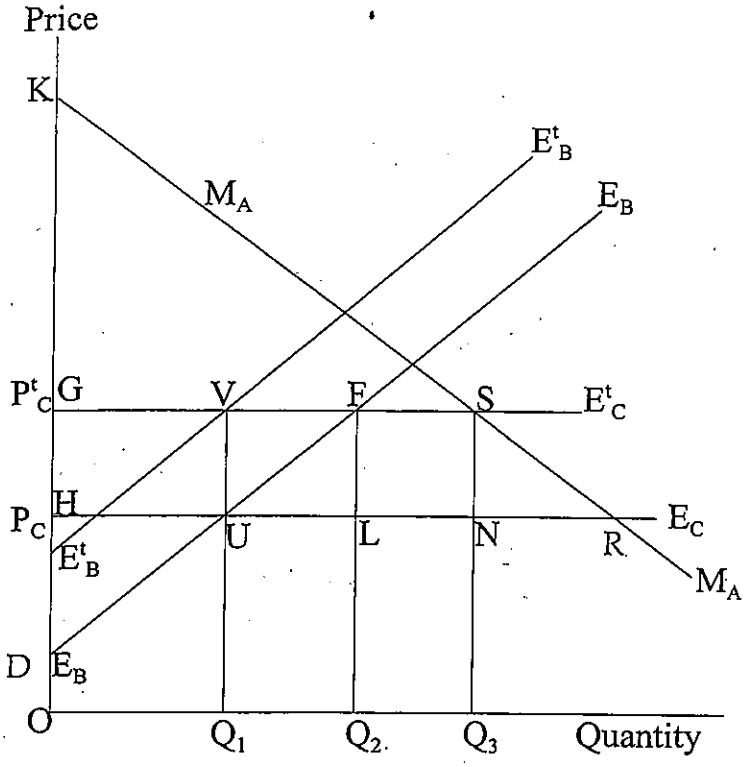


Figure 2.2
Effect of union (A + B) with rising costs from partner country.

In both figures 2.2 and 2.3, as before, we then let M_AM_A represent the import demand for a product imported by A. The supply curve of the product available from B is upward sloped and is represented by E_BE_B. Country C's export supply curve, represented by P_CE_C, is horizontal. The tariff continues to be specific. Consider then figure 2.2 and three cases: an initial nondiscriminatory tariff, free trade, and an FTA.

Under a *nondiscriminatory tariff* at rate t per unit, supplies from B and C, as perceived by buyers in A, are given by $E^t_BE^t_B$ and $P^t_CE^t_C$, respectively. Total imports into A equal OQ_3 of which OQ_1 come from B and Q_1Q_3 from C. Country A collects tariff revenue equivalent to rectangle GHNS. The gains from trade for A amount to the area under the import-demand curve and above the domestic price plus the tariff revenue, that is, triangle KSG plus rectangle GHNS. For country B, the gains from trade equal the area above E_BE_B and below the net price received, P_C, that is, area HUD. Country C neither gains nor loses from trade. Table 2.1 summarizes this information in column 1.

Suppose instead the A decides to adopt a policy of *free trade* by eliminating the tariff on a nondiscriminatory basis. The price in A declines to P_C, imports from B do not change, and imports from C rise by NR. Tariff revenue disappears, but the gains from trade rise to KGS + GHNS + RSN:

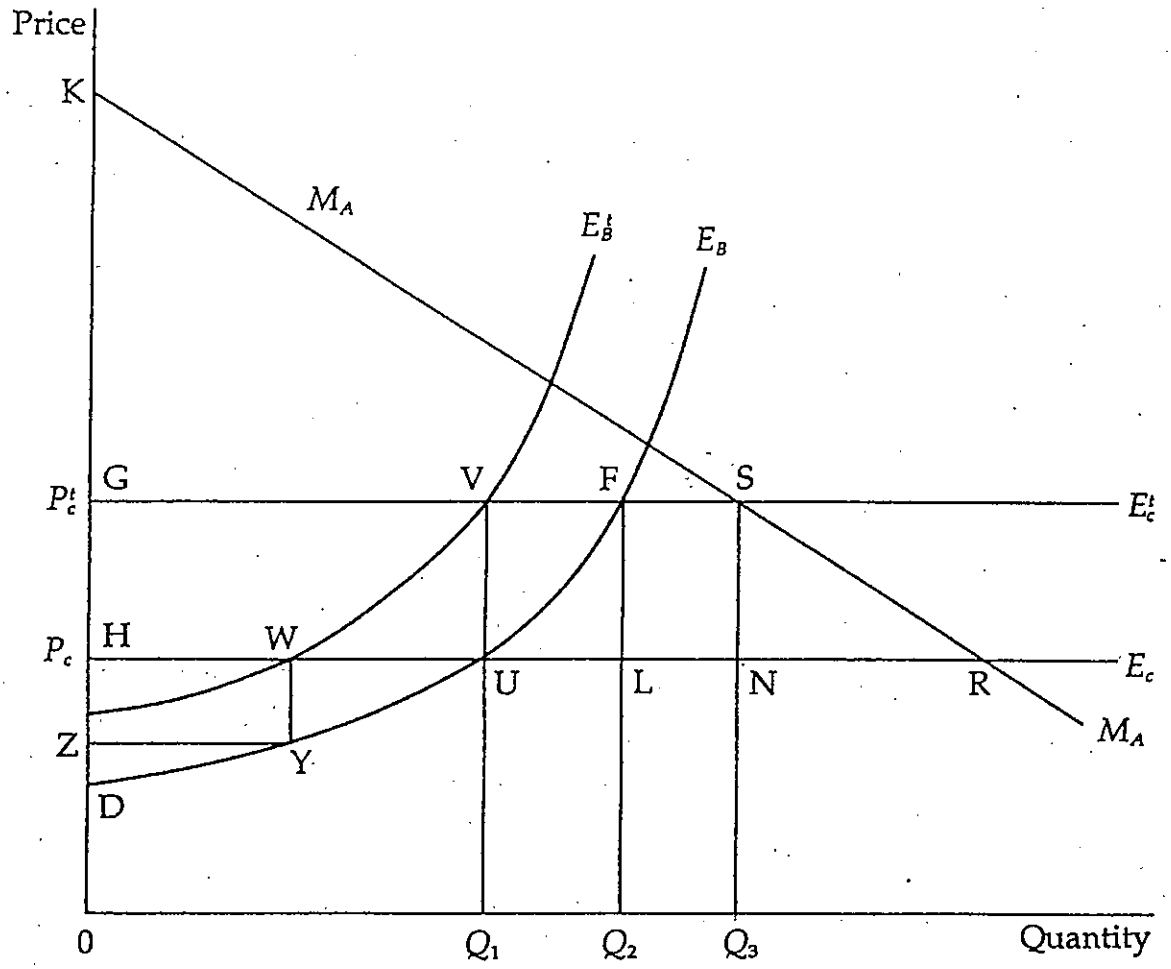


Figure 2.3
Effect of union (A + C) with rising costs from outside country.

there is a net welfare gain to A of RSN. The extra gain comes from increased benefits to consumers and producers in A. The gains to country B remain unchanged at HDU. Because of the perfectly elastic supply, country C neither gains nor loses from trade before or after trade liberalization by A. Therefore, the world as a whole benefits by area RSN. These changes are summarized in column 2 of table 2.1.

Next, assume that A forms an FTA with B by eliminating entirely the tariff on B but retaining it on C. Imports from B rise to OQ_2, and those from C decline to Q_2Q_3. Now B gains from the FTA due to an improvement in its terms of trade. The net price received by the exporters of B increases from P_C to P_C^t, and the gains from trade to B rise to HDU + GFUH. Country B gains from A's liberalization.

Because imports continue to come from C before as well as after the FTA and C's supply is perfectly elastic, the price in A is unchanged. But now that there is no tariff revenue on goods coming from B, A' gains from trade decline by GFLH. Stated differently, A's within-union terms of

Table 2.1
Gains from Trade under Unilateral Liberalization and Free Trade Area (Country A plus Country B)

Country	Nondiscriminatory tariff (Initial situation) (1)	Free trade (FT) (2)	Free trade area (A and B) (FTA) (3)
A	KGS + GHNS	KGS + GHNS + RSN (A gains)	KGS + GHNS − GFLH (A loses)
B	HDU	HDU (no change)	HDU + (GFLH − FLU =) GFUH (B gains)
C	0	0 (no change)	0 (no change)
World	KGS + GHNS + HDU	KGS + GHNS + HDU + RSN (World gains)	KGS + GHNS + HDU − FLU (World losses)

Note: This table relates to figure 2.2 in the text.
a. World Welfare Loss from FTA compared with FT: FLU + RSN.
b. World Welfare Loss from FTA compared with Initial Situation: FLU.

trade worsen by the full amount of the tariff liberalization country A loses from its own liberalization. Because the FTA diverts imports Q_1Q_2 from the more efficient C to the less efficient B, A's loss exceeds B's gain by the area FLU. The world as a whole loses by the same area FLU. The last column in table 2.1 shows these changes.

It is now evident that Summers' earlier-cited argument that international economists should embrace PTAs because second-best "trade diversion" worries are "laughable," and that primary effects must be considered to be dominant, is misplaced when impacts on the welfare of specific countries are considered. The loss to A from its own preferential liberalization arises primarily from the primary effect of these.[13] If we assume that the initial imports from the union partner are large, the loss to A in this wholly trade-diverting union is substantial and reflects the tariff revenues lost on the original imports (plus these diverted imports) from the partner country B.

Clearly, FTAs can give rise to large redistributive effects (on original imports) between countries. The amount of trade diverted (Q_1Q_2) may be small, and the loss to the union from this trade diversion is small because it is a triangle, but all this really has no relevance to our critique of Summers as just concluded.

Next, our analysis casts doubt on the recent presumption that countries that trade with each other in large volume are "natural" trading partners and regional arrangements among them must therefore be beneficial to

them. It is not entirely clear from the literature what it means to be "natural" trading partners.[14] A quotation from Summer[14] (1991, 297), however, should help: "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."[15]

Later we consider this entire question of natural trading partners and their desirability. But our analysis so far already provides a devastating critique of the presumption advanced in favor of such natural trading blocs. It is evident from figure 2.2 that the larger the initial quantity of imports from a trading partner, the greater (not smaller) the loss to the country liberalizing preferentially, ceteris paribus. That is to say, the more natural the trading partner according to Summers' definition, the larger the loss from a discriminatory trade liberalization with it!

Finally, it has been frequently argued that, given today's low levels of trade restrictions, preferential trading arrangements are unlikely to be harmful: trade creation effects should dominate the outcome, making PTAs as good as FT (free trade). But this argument, plausible as it sounds, is contradicted by our analysis. Thus in figure 2.2, if the initial nondiscriminatory tariff is sufficiently high, an FTA between A and B can eliminate C as a supplier of the product. In this case, the FTA lowers the internal price in A and gives rise to trade creation. Under some (admittedly strong) conditions, this trade creation can outweigh the tariff-revenue loss and may improve welfare. By contrast, if the initial tariff is low, the chances are poor that the formation of the FTA will eliminate imports from C and lower the internal price.

The Outside Country's Supply Curve Is Upward Sloped

The conclusion that A's preferential liberalization hurts itself and benefits its union partner has been derived under the assumption that the supply of B is less than perfectly elastic and that of C is perfectly elastic. In this setting, the union partner is a less efficient supplier of the product than is the outside world. What will happen if the situation was reversed such that B's supply curve was perfectly elastic and C's less than perfectly elastic?

This case can be analyzed by letting A form a union with C rather than B. In this case, analyzed in figure 2.3, an FTA lowers the price in A to P_C. Though there is no gain to the union partner, A's gain from the FTA

Table 2.2
Gains from Trade under Free Trade Area (Country A plus Country C)

Country	Nondiscriminatory tariff (1)	FTA (A and C) (2)
A	KGS + GHNS	KGS + GHNS + RSN + HWYZ (A gains)
B	HDU	ZYD = HDU − WYU − HWYZ (B losses)
C	0	0 (no change)
World	KGS + GHNS + HDU	KGS + GHNS + HDU + RSN − WYU (World may gain or lose according as RSN >< WYU)

Note: This table relates to figure 2.3 in the text. Column 1 is identical to column (1) in table 2.1 and is reproduced here to facilitate comparison. The results under Free Trade are identical between the two FTAs, (A and B) and (A and C).

(= RSN + HWYZ) exceeds that under nondiscriminatory liberalization (that is, free trade) by the amount of tariff revenue (= HWYZ) collected on imports from the outside country.[16] This case brings us back to the conventional presumption that A's liberalization should benefit it (though the presumption that others should gain from the liberalization does not carry through for the outsider country B that loses). The precise welfare results, based on analysis of figure 2.3, are drawn together in table 2.2.

This case clearly undercuts the arguments about the dangers of PTAs to country A that were made in the previous section. Therefore, it is important to ask how relevant this case is empirically. It is perhaps reasonable to assert that a union partner is likely to resemble B in some products and C in other products, and therefore the effect of the FTA will be ambiguous in general.

A common claim has been that NAFTA is likely to benefit Mexico because the United States the Canada are very large and therefore the most efficient suppliers of a majority of Mexico's products. Our analysis suggests, however, at least two reasons why this conclusion is not warranted.

First, given that the outside world includes the European Union, Japan, China, Korea, Hong Kong, and numerous other outward oriented and highly competitive countries, the conclusion that the United States and Canada are the most efficient suppliers of a large majority of Mexico's products is highly suspect. Indeed, if it were true we would be hard-pressed to explain the persistent demands for antidumping and other forms of protection in the United States.[17] Second, recall that if the union partner is a large supplier of imports, the tariff redistribution losses to A in the case of trade diversion are large. Therefore, even if the union partner is the most efficient supplier of the majority of A's imports, losses

may outweigh any gains. In the case of NAFTA, the United States does account for a sufficiently large proportion of Mexico's imports for us to conjecture plausibly that the tariff-redistribution losses in trade-diversion cases could outweigh the gains in trade-creation cases.

This analysis has an important qualification that will be discussed in the next section. Before doing so, we mention two additional possibilities that are worthy of brief consideration: first, export-supply curves are upward sloped for both B and C; second, the products of A, B, and C are imperfect substitutes. In either of these cases, the small-country and small-union assumptions are violated, and a complete elimination of the tariff by A, whether on a discriminatory or nondiscriminatory basis, is not the optimal policy.[18] We will look at the second case in detail [a bit later].

But here we note that our conclusions remain valid under the following circumstances. In case one, if the elasticity of supply of the outside country is high in relation to that of the union partner, B, a discriminatory tariff reduction is likely to hurt A itself while benefiting B. In case two, analyzed later in the chapter, if B's goods are poor substitutes for A's goods but not C's, as seems entirely plausible, discriminatory liberalization by A will hurt A itself and benefit the union partner, B, even at constant terms of trade, whereas the terms of trade effects will reinforce this outcome. Before we present this analysis in detail, an important qualification to figures 2.2 and 2.3 must be noted.

A Qualification and Modification

Figures 2.2 and 2.3 capture the essence of a large body of the literature on regional arrangements that emerged in the 1950s and 1960s. But these figures have an important limitation that has been ignored entirely in the literature until recently. They implicitly assume that either (1) the partner country maintains the same tariff as A on the product under consideration (that is, the arrangement is a Customs Union), or (2) the product is not consumed in the partner country. Let us explain why.

Consider first the case depicted in figure 2.2. The common practice in the literature, as in our foregoing analysis, has been to assume that post-FTA prices in a member country are determined by the price in the outside country, C, plus the country's own tariff. As Richardson (1994) has noted recently, however, this assumption is incorrect in general. It implies that, if tariffs in A and B are different, producer as well as consumer prices in A and B are different. But given duty-free movement of goods pro-

duced within the union, producer prices between A and B must equalize under an FTA.

Assume that the tariff on the product under consideration is lower in B than in A, violating condition one above. In figure 2.2, recall that E_BE_B is B's supply curve for exports, that is, output supply net of domestic consumption. Under a nondiscriminatory tariff in A, B's producers sell OQ_1 in A. Because the net price received by exporters on sales in A is P_C, the domestic price in B will also be P_C. If A and B now form an FTA and the price in A remains P_C^t, producers in B have no incentive to sell anything in their domestic market unless the price there also rises to P_C^t. But given that the tariff in B is lower than that in A, the price in B cannot rise to P_C^t, and the entire quantity of the product previously sold in B is diverted to A. The rules of origin can forbid the diversion of goods *imported* from C to A but not of goods *produced* in B.[19] Unless domestic consumption of the product in B is zero (assumption two above), B's export-supply curve shifts to the right by the quantities demanded in B at each price, that is, B's export-supply curve coincides with its output-supply curve.

Figure 2.4 lays out how the allowance for the diversion of B's domestic sales to A after the formation of the FTA affects our conclusions. It reproduces figure 2.2, omitting $E_B^tE_B^t$. In the initial equilibrium, with a nondiscriminatory tariff in A, imports from B are OQ_1 as in figure 2.2. After the FTA is formed, the expansion of exports is larger than that given by point F. How much larger it is will depend on where B's total supply curve lies. There are three possibilities.[20]

First, if the total supply curve intersects M_AM_A above point S as shown by S_BS_B, the results of the previous section hold with a vengeance.[21] Exports from B now expand more than in figure 2.2, and losses to A from the transfer of tariff revenue are larger. In this case, B's producers sell all of their output in A and receive the same price as A's producers, namely, P_C^t. The entire quantity consumed in B is imported from C, with consumers paying a price lower than P_C^t. Country A imports from countries B and C.

Second, suppose that B's supply curve intersects M_AM_A between S and W, where the height of W is P_C plus the tariff in B. In this case, the price in A is determined by the height of the point of intersection of B's supply curve and M_AM_A. Because this price is below P_C^t, a part of the lost tariff revenue is now captured by A's consumers. But we still have a tariff-revenue transfer to firms in B. The transfer is larger the closer the intersection point of the two curves to S. Producers in B sell all their output in A, A does not import anything from C, and B imports everything from C.

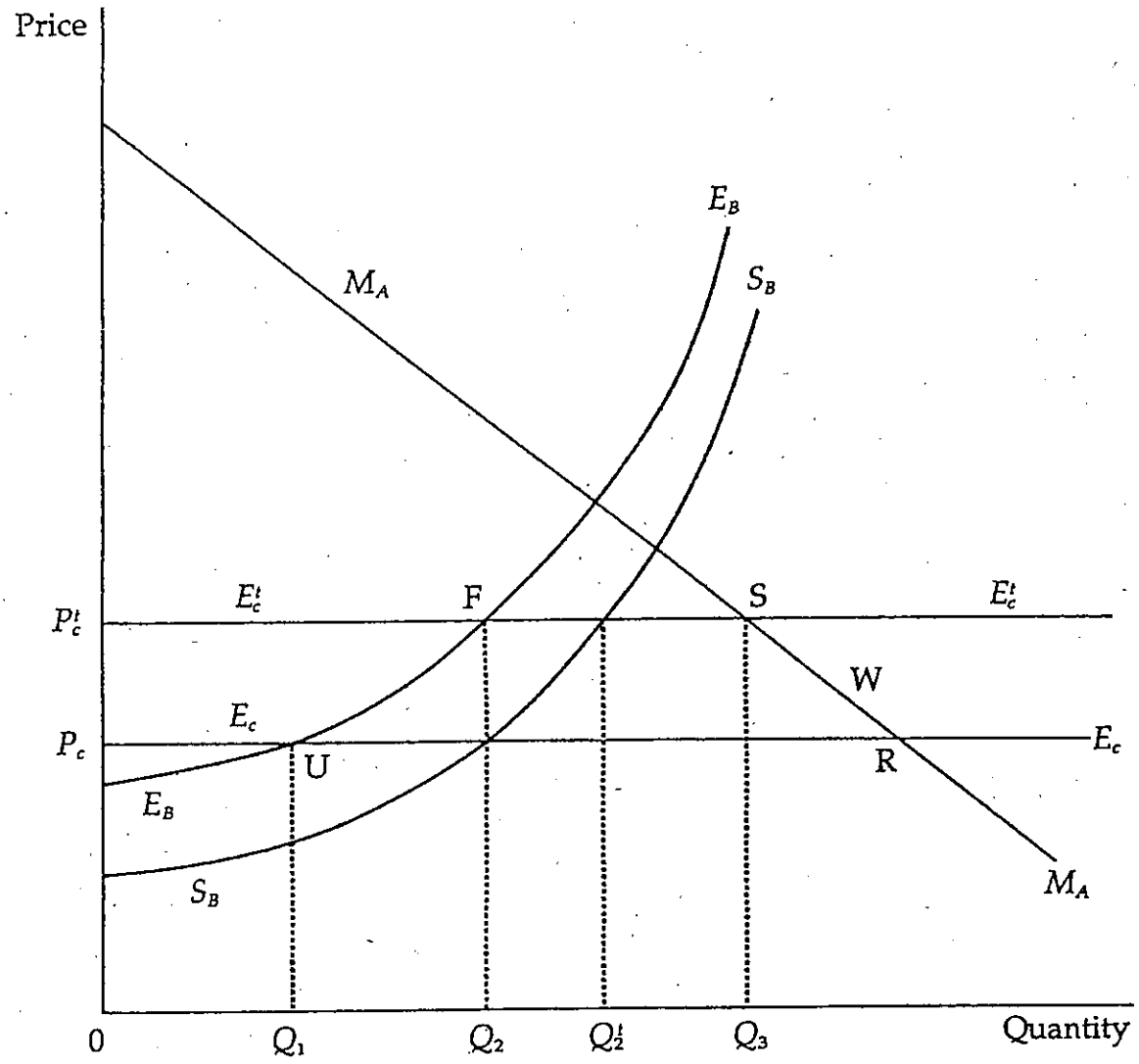


Figure 2.4
The consequences of differing external tariff rates in members of a free trade area.

Finally, if B's supply curve intersects $M_A M_A$ below point W, the price in A drops to the tariff-inclusive price in B given by the height of point W. All of A's imports come from B with producers in B selling in A as well as B. Both consumer and producer prices equalize between A and B. In this case, the redistributive effect is a declining function of the tariff in B. In the limit, if the external tariff in B is zero, the FTA leads to free trade in A (just as in B).

The case depicted in figure 2.3 is also modified along the lines of figure 2.4 if the good in question is consumed in the partner country (C) and the latter levies a tariff lower than that of A. To illustrate, assume that the tariff in C is zero, and the demand for the product in C at P_C is larger than B's supply at that price. Then, B can sell all it wants to export at P_C to C. In the post-FTA equilibrium, A's imports come entirely from C, while B

sells all its exports to C. The tariff revenue raised by A on imports from B in figure 2.3 is no longer available, and A's gains from the FTA with C are reduced to triangle RSN, the same as under unilateral, nondiscriminatory liberalization.

An Imperfect-Substitutes Model

An unrealistic implication of the model just explored in figure 2.4 is that, under an FTA, either producers of B must sell all their output in A and none in their domestic market (the first two cases) or consumers in A must import everything from B and none from C (the last case). This conclusion does not require a complete FTA; it can hold true even in the presence of a small tariff preference as long as external tariffs in the two countries are different. A quick examination of the direction of trade data of member countries of preferential trading arrangements such as the Association of Southeast Asian Nations (ASEAN) and NAFTA shows that this outcome is inconsistent with reality.

A natural way to avoid these extreme results is to cast the analysis in terms of a model with product differentiation. A fully satisfactory model of this type requires the introduction of economies of scale and monopolistic competition or oligopoly. Such an elaborate model is beyond the scope of this chapter. But taking recourse to the Armington structure whereby products are distinguished by the country of origin and drawing on the Meade (1955) model, we take a first stab at the problem.

An important point to note at the outset is that when products are differentiated by the country of origin, the small-country or small-union assumption must be abandoned.[22] If the product originating in a country is not produced anywhere else, by definition, the country is a monopolist for that product and cannot be a price taker in the world market.[23] Our approach below is to first consider the implications of FTAs at constant border prices and then bring in the effects of changes in the terms of trade.

Assume that there are three products denoted a, b, and c. Countries A, B, and C specialize in and export a, b, and c, respectively. Choose the units of each product so that its international price is unity in the initial equilibrium. Focus as before on country A's welfare. In the initial equilibrium, let A impose a uniform tariff t per unit on imports from B and C.

In figure 2.5 we measure A's consumption of b to the right and that of c to the left of the origin, O. Because b and c are not produced in A, the demand curves also represent import demands. Given the tariff t on b and

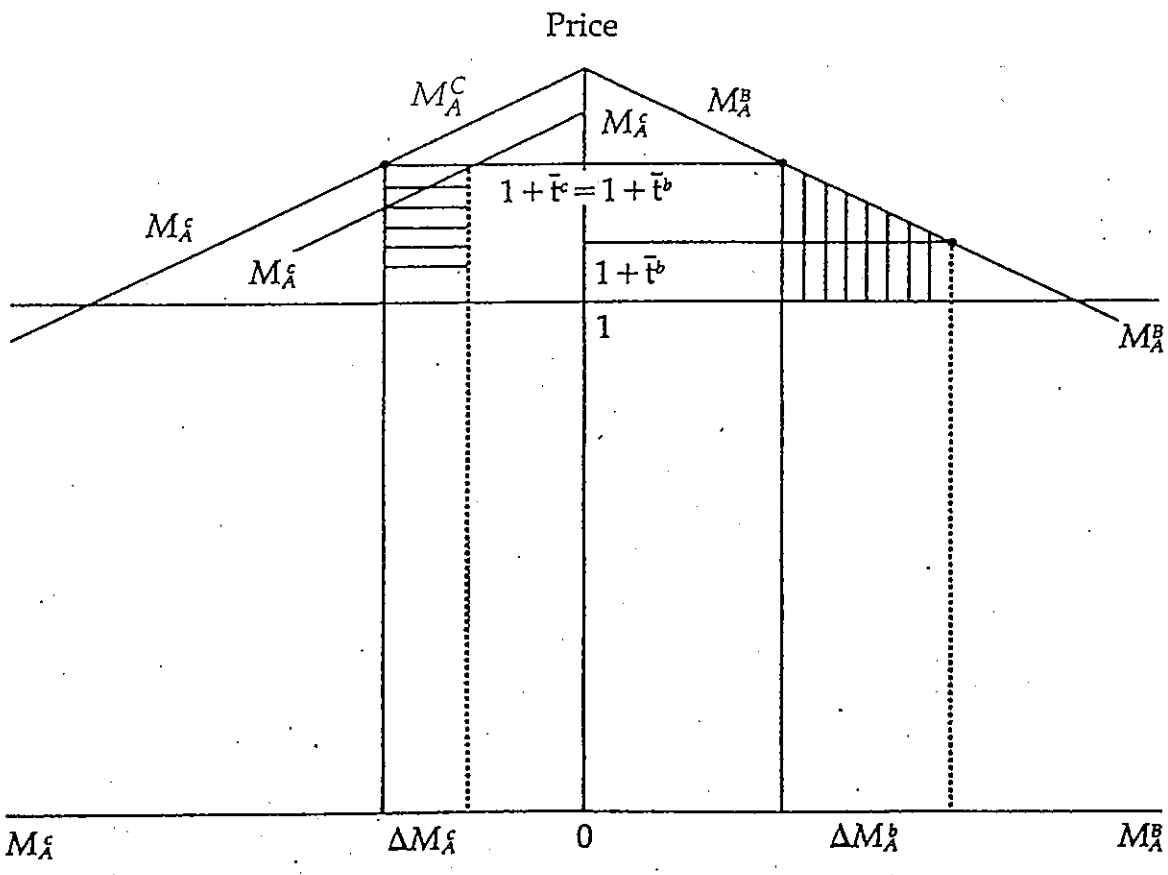


Figure 2.5
Effect of a small tariff preference by country A to country B.

c, (import) demand curves are represented by $M_A^b M_A^b$ and $M_A^c M_A^c$. The demand curve for each product is drawn given the tariff rate on the other product. Assuming substitutability, a reduction in the tariff on one product shifts the demand curve for the other product toward the vertical axis.

Let us now introduce preferential trading through a small reduction in the tariff on imports from B. Imports from B expand and generate a gain equal to $t_b \Delta M_A^b$ and approximated by the vertically shaded area in figure 2.5. This is trade creation. But the reduction in the tariff on b also causes an inward shift in the demand curve for c as shown by the dotted curve. There is trade diversion and a corresponding loss equal to $t_c \Delta M_A^c$ and approximated by the horizontally shaded area.[24]

Is there a net gain or loss to country A? The answer depends on the relative sizes of the two shaded areas. For a small change in the tariff, these areas are approximated by rectangles whose height equals t. Therefore, the gain is larger than the loss if and only if the increase in the value of imports of b at world prices is larger than the reduction in the value of imports of c.[25] If we now assume that the partner country's good, b, and A's export good, a, are substitutes in A's demand, the preferential reduc-

tion in the tariff lowers the consumption of good a and allows an expansion of exports. Working through the trade balance condition, we can see that the expansion of exports must expand total imports valued at world prices. That is to say, imports of b expand more than imports of c contract. The area associated with trade creation in figure 2.5 exceeds the area associated with trade diversion; the *introduction* of preferential trading is beneficial.

This result is attributed to Lipsey (1958) and hinges critically on substitutability between demands for the partner country's goods and exportables and constancy of the terms of trade. For the moment, let us make these assumptions and ask what happens as we continue to lower the tariff on good b, holding the tariff on good c unchanged. For each successive reduction in the tariff, the height of the rectangle associated with trade creation declines but that of the rectangle associated with trade diversion remains unchanged. Sooner or later, before the tariff on b goes to zero, the gain from extra trade creation becomes smaller than the loss from extra trade diversion. Further reductions in the tariff lead to a *reduction* in welfare.[26]

In sum, assuming constant terms of trade and substitutability between imports from B and exports, a preferential reduction in the tariff on B's goods first improves welfare and then lowers it. This relationship is shown in figure 2.6. As drawn, the level of welfare with a complete FTA is lower than that in the initial equilibrium. But in general, we cannot tell whether welfare rises or falls upon the establishment of an FTA.

The natural question then is whether we can establish a presumption one way or the other. To answer it, let us examine the second-best optimum tariff in the Meade model on B's goods given the tariff on C's goods. As shown in Panagariya (1996b), this tariff can be written as

$$\frac{t^b_{opt}}{1+t^b_{opt}} = \frac{\bar{t}^c}{1+\bar{t}^c} \cdot \frac{1}{1+\frac{\eta_{ba}}{\eta_{bc}}}, \tag{1}$$

where η_{ba} and η_{bc} are country A's compensated, crossprice elasticities of demand for the partner country's good with respect to the price of its own good and that of the outside country's good, respectively. These elasticities respectively measure the degree of substitutability between the partner's and A's own goods and that between the partner's and outside country's goods.

If the two elasticities are equal, the optimum tariff on b is approximately half of the tariff on c. In applying this model and argument to the

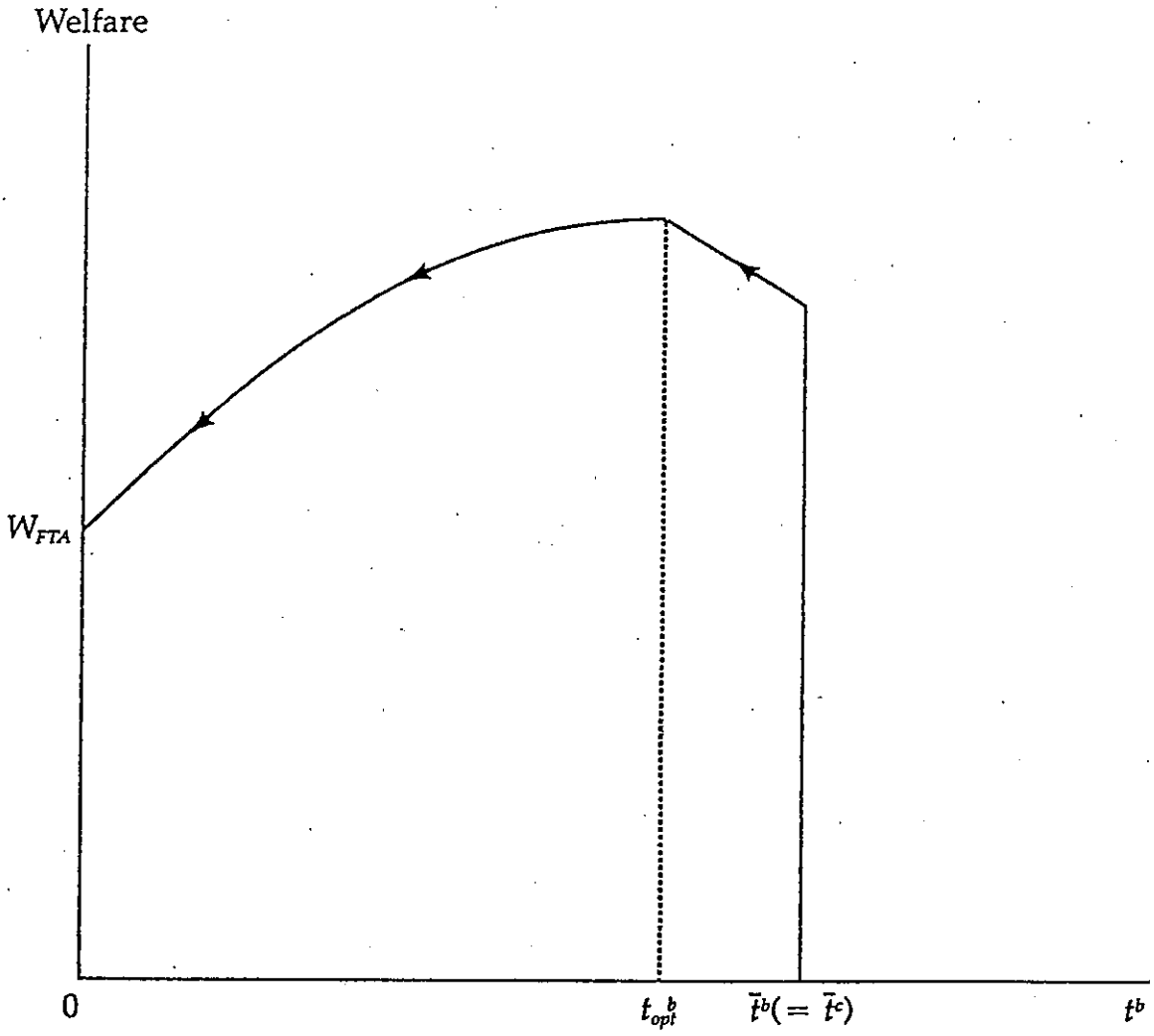


Figure 2.6
Effect of preferential tariff reduction and welfare.

real world, it is reasonable to expect, however, that the degree of substitutability is substantially higher between the imports from the two sources, B and C, than that between imports from B and A's exportables. For instance, Chile's imports from North America are likely to exhibit a much greater degree of substitutability with goods from the European Union or East Asia than with its own exports. Given this fact, the optimum tariff on b is higher than one-half of the tariff on the outside country's goods. In the limit, if the cross-price elasticity of demand for B with respect to the price of A is zero, the optimum tariff on b equals the initial tariff on c. In terms of figure 2.5, exports do not change at all when preferential trading is introduced and trade diversion exactly offsets trade creation. In terms of figure 2.6, welfare falls monotonically as we lower the tariff on b while holding the tariff on c constant.

The analysis up to this point has assumed that the terms of trade are constant, and it does not allow for the tariff-revenue-redistribution effect discussed earlier.[27] As already noted, with goods differentiated by the country of origin, the terms of trade cannot be assumed constant. The derivation of the effects of preferential trading on the terms of trade in the three-good model is complicated. Fortunately, in a neglected but important paper, these effects were worked out by Robert Mundell. To quote him,

> 1. A discriminatory tariff reduction by a member country improves the terms of trade of the partner country with respect to both the tariff reducing country and the rest of the world, but the terms of trade of the tariff-reducing country might rise or fall with respect to third countries.
> 2. The degree of improvement in the terms of trade of the partner country is likely to be larger the greater is the member's tariff reduction; this establishes the presumption that a member's gain from a free-trade area will be larger the higher are initial tariffs of partner countries (Mundell 1964, 8).

Not surprisingly, once the terms-of-trade changes are brought back into the analysis, the "mercantilist" bias in results noted earlier (that is, that A loses from its own liberalization) comes back even in the Meade model.[28] We are once again driven to the conclusion that a high-protection country (Mexico) forming an FTA with a low-protection country (United States) is likely to lose from the FTA. Observe that the terms-of-trade effects are in addition to the likely losses from second-best considerations at fixed terms of trade as discussed in figures 2.5 and 2.6.

Revenue Seeking

The conclusion that a country is likely to lose from its own preferential liberalization can break down in the presence of 100 percent, perfectly competitive, resource-using revenue-seeking activities.[29] Given this type of revenue seeking, each dollar's worth of tariff revenue will be matched by a dollar's worth of real resources used unproductively. The tariff revenue is represented by the rectangle GHNS in figure 2.2, where A and B form the FTA. This revenue is now lost in revenue seeking and will not contribute to the country's welfare. The introduction of preferential trading will then lead to a loss of tariff revenue in the amount GFLH, but it will generate an exactly equivalent gain from a release of resources employed in revenue seeking, leaving A's welfare unchanged. For the union as a whole, however, the reduced revenue seeking will generate a net gain equal to GFLH. A large part of this gain, trapezium GFUH, will

go to the partner country B, while the remaining part, triangle UFL, pays for the cost of trade diversion. In sum, country A's welfare does not change while that of B rises.

Next, consider the case in figure 2.3 where A and C have the FTA instead. Once again, the rectangle GSNH now will not contribute to the country's welfare in the initial equilibrium. But when preferential liberalization is introduced, the internal price of A falls to the level shown by point R and the rectangle (plus triangle SNR) becomes a part of the consumers' surplus and hence A's welfare rises. Country B's welfare does not change.

Combining the two cases, we obtain the conclusion that, in the presence of 100 percent perfectly competitive revenue seeking, each partner benefits unambiguously (or at least does not lose) from preferential trading. This conclusion undermines our argument that preferential liberalization by a country with respect to its major trading partner is likely to hurt itself and benefit its partner.

We suggest, however, that there are at least two reasons why we should not take this conclusion seriously. First, even though revenue seeking is an important phenomenon in certain contexts and worthy of analysis in its own right, it is hardly invoked when making major policy decisions. We are not aware of a single reference to revenue seeking as a major reason for NAFTA in the public debate in either Mexico or the United States and Canada preceding its approval. Indeed, if we are to take revenue seeking seriously, we should take it and other types of directly unproductive profit-seeking (DUP) activities arising from all other policies into account as well. Second, the twin assumptions of 100 percent and perfectly competitive revenue seeking are unrealistic. Empirically, revenue seeking is likely for several reasons to be a small fraction of the total revenue. In particular, the operation of the "Brother-in-Law Theorem" and of settled rules for allocation of revenues will often turn potential DUP activities into transfers.

"Natural Trading Partners" Hypothesis and Regional PTAs

We now turn to the question of natural trading partners.[30] As we noted earlier, the "natural trading partners" phrasing and hypothesis (that PTAs among them are more likely to beneficial) originated in Wonnacott and Lutz (1989). Based on the work of Viner (1950), Lipsey (1960), and Johnson (1962), these authors provided detailed criteria for determining whether or not a given set of countries constituted natural trading partners:

Trade creation is likely to be great, and trade diversion small, if the prospective members of an FTA are natural trading partners. Several points are relevant:

- Are the prospective members already major trading partners? If so, the FTA will be reinforcing natural trading partners, not artificially diverting them.
- Are the prospective members close geographically? Groupings of distant nations may be economically inefficient because of the high transportation costs (Wonnacott and Lutz 1989, 69).

Wonnacott and Lutz offered two further criteria, one based on complementarity versus competitiveness of the economies and the other on the countries' relative levels of economic development. They noted, however, that these characteristics are "much more difficult to evaluate." Because subsequent advocates of FTAs have not included these criteria in defining natural trading partners, we will not discuss them.

For clarity, we will refer to the first two criteria spelled out in the above passage from Wonnacott and Lutz as the "volume-of-trade" and "transport-cost" criteria and examine them in turn.

The Volume-of-Trade Criterion The volume-of-trade criterion for choosing natural trading partners and treating them as likely therefore to be welfare enhancing to their members seems plausible at first glance but is, in fact, treacherous for several reasons.

First, the criterion is neither symmetric nor transitive. A lack of symmetry implies that country A may be a natural trading partner of country B, but the reverse may not hold true. A lack of transitivity implies that even if A is a natural trading partner of B, and B is a natural trading partner of C, A may not be a natural trading partner of C. Lest this be viewed as a purely academic point, we note that the United States is Mexico's largest trading partner, but the reverse is not true. Similarly, the United States is the largest trading partner of both Canada and Mexico, but Canada and Mexico have little trade with each other.

Second, the volume-of-trade criterion is premised on the view that a larger initial volume of trade between potential partners implies a lower likelihood of loss because of trade diversion. In terms of figure 2.2, this implies that the larger is OQ_1, the smaller is Q_1Q_2.

This is, however, an unsupported inference from the fact that, for any given volume of initial imports (OQ_3), the higher is the partner country's initial share, the lower is the outside country's share and hence the smaller is the *scope* for diverting trade. Instead, what one needs to determine is how likely is the *actual* trade diversion. Thus, for example, between two alternate situations, one where Q_1Q_3 (the scope for trade diversion) is

twice as large as in the other, Q_1Q_2 (the actual trade diversion) could still be only half as much.

The underlying model that defines the trade volumes in different equilibriums may well imply then that the relationship between the initial volume of imports from the partner country and the trade to be diverted to it may be altogether tenuous.

Thus, consider the Lipsey (1958) analysis of the question, based on the small-union version of the Meade model we have discussed.[31] Lipsey, as Bhagwati recalled in his earlier critique of the volume-of-trade criterion, focused not on the initial volume of imports but "on the relative sizes of imports from each source vis-à-vis expenditure on domestic goods as the key and decisive factor in determining the size of losses and gains from the preferential cuts in trade barriers" (Bhagwati 1993a, 34). Of course, on the basis of equation 1 and the discussion of it, we can also conclude that, in general in this model, the higher is the compensated crossprice elasticity of its demand for the partner's good with respect to the price of its own good *relative* to the crossprice elasticity of its demand for the good with respect to the price of the outside country's good, the higher is the likelihood that an FTA improves a country's welfare. This general conclusion reduces to the Lipsey argument when the liberalizing country's preferences are of the CES variety.[32]

For a country such as Mexico joining the NAFTA with the United States, we may well expect in fact the former elasticity to be lower than the latter so that the welfare presumption for this "natural trading partner" of the United States from NAFTA is ironically likely to be in favor of trade-diversion effects dominating the outcome.

There is a further subtle point to be noted. In figure 2.6, starting from a nondiscriminatory tariff, as country A lowers the tariff on B, trade share shifts in favor of B at the expense of country C. That is, A and B become more natural trading partners according to the volume-of-trade criterion. Yet, once the tariff on B attains the second-best optimum, t^b_{opt}, further preferential liberalization is accompanied by a *reduction* in the welfare of A. Thus, to the left of t^b_{opt}, A and B are more natural trading partners than to the right of it, but preferential tariff reductions in that range reduce welfare.

Third, even this conclusion understates the folly of regarding a large initial volume of imports as a benign phenomenon. It ignores the crucial tariff-revenue-redistribution effect that we have highlighted. In FTAs involving countries with asymmetric levels of protection and a high volume of trade initially, the country with higher protection is likely to

lose even if trade-creation effects dominate trade-diversion effects. Under such circumstances, the net gain from trade-creation and trade-diversion effects could likely be swamped by the loss from the tariff-revenue-redistribution effect. The case for Mexico gaining from joining NAFTA thus looks dismal on this account as well.

While, therefore, the volume-of-trade criterion for judging FTAs to be benign is clearly to be rejected, linking it to *regionalism* and thus declaring regional FTAs to be more benign than nonregional FTAs is additionally wrong. There is no evidence at all that pairs of contiguous countries, or countries with common borders, have larger volumes of trade with each other than do pairs that are not so situated or that trade volumes of pairs of countries, arranged by distance between the countries in the pair, will show distance to be inversely related to trade volumes.[33]

This is evident from the somewhat aggregated destination-related trade volume statistics for major regions in 1980, 1985, and 1990 in table 2.3.[34] Then again, take just one telling example.[35] Chile shares a common border with Argentina, but in 1993 it shipped only 6.2 percent of exports to Argentina and received only 5 percent of its imports from her (Panagariya 1995b, tables 3 and 4). By contrast, the United States does not have a common border with Chile but in 1993 accounted for 16.2 percent of her exports and 24.9 percent of her imports. The volume-of-trade criterion then would make the United States, *not* Argentina, the natural trading partner of Chile, clearly controverting the claim that the volume-of-trade criterion translates into a regional criterion.

As contended by Bhagwati (1993a), the equation by Krugman (1991a) and Summers (1991) of the two concepts of volume of trade and regionalism (whether of the distance or the common border or contiguity variety) is therefore simply wrong.

Nonetheless, Frankel and Wei (1995) have recently argued otherwise, claiming that their empirical work favors the Krugman-Summers assertion. They use the gravity model as their basic tool to conclude that "proximity is in general an important determinant of bilateral trade around the world, notwithstanding exceptions like India-Pakistan and other cases."

But this misses the point at issue. What is at stake is not whether distance, interpreted through the gravity model and/or common border modeled through a dummy, matters.[36] There does seem to be a *partial* correlation between distance, proximity, common border, and so on, on the one hand, and trade volumes on the other.[37] But what we have to look at is the *total* initial volume of trade, and this does not correlate simply with distance as the right-hand side variable, as required by the

Table 2.3
Direction of Exports by Major Regions, 1980, 1985, and 1990

		Partner							
Exporter	Year	North America	Western Europe	Europe	East Asia[a]	Latin America	Africa	Middle East	South Asia
North America	1980	33.5	25.2	27.4	15.8	8.9	3.3	4.2	1.0
	1985	44.4	19.3	21.0	15.5	5.9	2.5	3.2	1.0
	1990	41.9	22.3	23.4	20.4	5.0	1.7	2.6	0.8
Western Europe	1980	6.7	67.1	71.9	2.9	2.4	7.2	5.5	0.7
	1985	11.3	64.9	68.9	3.6	1.6	5.2	5.0	0.9
	1990	8.3	71.0	74.4	5.3	1.1	3.3	3.3	0.7
Europe	1980	6.3	63.7	72.7	2.7	2.3	6.9	5.5	0.7
	1985	11.0	63.5	69.2	3.4	1.6	5.1	5.0	0.9
	1990	8.2	70.6	74.5	5.2	1.1	3.3	3.3	0.7
East Asia	1980	26.0	16.8	18.9	29.9	4.1	4.4	7.4	1.8
	1985	37.8	13.6	15.5	25.3	2.8	2.2	5.1	2.0
	1990	31.9	19.8	20.7	32.3	1.9	1.6	3.0	1.5
Latin America	1980	27.9	26.5	35.1	5.4	16.6	2.7	1.9	0.5
	1985	35.8	25.9	30.4	7.1	12.1	3.7	3.0	0.7
	1990	22.9	25.3	27.6	10.3	14.0	2.1	2.4	0.4
Africa	1980	27.4	43.6	46.1	4.3	3.2	1.8	1.7	0.3
	1985	14.8	64.9	69.3	1.8	4.2	5.1	2.2	0.7
	1990	3.0	66.0	68.0	4.6	0.6	12.8	4.4	3.6
Middle East	1980	11.5	40.3	41.5	28.7	5.0	1.5	4.1	2.5
	1985	6.2	15.0	17.7	1.5	0.3	1.4	8.7	0.4
	1990	17.8	48.6	53.0	9.1	1.2	3.6	8.5	0.9
South Asia	1980	10.9	24.6	39.4	14.5	0.5	6.8	14.5	5.6
	1985	18.4	20.8	37.0	16.4	0.4	4.6	11.0	4.4
	1990	17.1	30.1	46.6	18.3	0.3	2.7	6.5	3.2

Source: Panagariya (1993). He cites U.N. COMTRADE data.
Note: This table broadly underlines the point that *total* trade volumes that matter do not show any relationship to proximity of countries geographically.
a. East Asia does not include China.

"natural trading partners" assertion of the volume-of-trade criterion for forming PTAs.

Next, we have the difficult problem of endogeneity of initial trade volumes with respect to preferences. If the large volumes are themselves attributable, in significant degree, to preferences granted earlier, then they are not "natural," nor is it proper to think that additional preferences are "therefore" harmless. The point is best understood by thinking of high trade barriers by a country leading to a larger within-country trade relative to external trade. To deduce that added barriers are harmless is to compound the harm done by existing barriers that are, of course, preferences in favor of trade within the country.

This is not an idle question. Offshore assembly provisions between the United States and Mexico and the longstanding GATT-sanctioned free trade regime in autos between Canada and the United States are certainly not negligible factors in pre-NAFTA U.S. trade with these NAFTA members. In granting preferences under the Generalized System of Preferences, the United States, EC, and Japan have all concentrated on their regions. Thus, the partial correlation between distance and trade volumes (in gravity models) may be a result of preferences granted to proximate neighbors, rather than a "natural" phenomenon justifying (new) preferences.[38]

Finally, we need to raise a different objection to the argument that a high initial volume of imports from a partner country will work to protect a country against trade diversion. Quite aside from the fact that aggregate volumes shift significantly in practice over time, the comparative advantage in specific goods and services often changes in different locations.[39] Consistent with a given aggregate trade volume, its composition may shift so as to yield greater trade diversion when a PTA is present.

Consider a case, based on constant costs for simplicity, in which the United States imports a product from Canada under a nondiscriminatory tariff. if a PTA is formed between the two countries, the product will continue to be imported from Canada. But suppose that, on a future data, Canada loses its comparative advantage to Taiwan ever so slightly so that the preferential advantage enjoyed by her outweighs this loss. There will be trade diversion, and imports into the United States will continue to come from Canada with the volume of trade remaining unchanged.[40] Observe that there is an asymmetry here between a shift in comparative advantage away from the partner and that toward it. If Canada experiences a reduction in the cost of production of a product imported by the United States from Taiwan under a PTA, there can still be trade diversion. Because of the preference, Canada will replace Taiwan as the supplier of this

product even before Canada's costs fall below those of Taiwan. The volume of trade will rise, and at the same time there will be trade diversion.[41] The proponents of the complacent "high volume of imports" thesis are thus trapped in a static view of comparative advantage that is particularly at odds with today's volatile, "kaleidoscopic" comparative advantage in the global economy.

Transport-Cost Criterion But if the volume-of-trade criterion is conceptually inappropriate and must be summarily rejected, what about the transport-cost criterion? This criterion maps directly into distance and hence into regionalism. However, the question to be analyzed is: should PTA partners be chosen on the basis of lower transport costs, and hence greater proximity, to maximize gains to members or to minimize losses to them?

The earliest reference we could find to transport costs in the context of trade liberalization is from Johnson (1962, 61): "If the separate markets of various members are divided by serious geographical barriers which require high transport costs to overcome them, the enlargement of the market may be more apparent than real." All he seemed to be arguing was the truism that trade liberalization may be meaningless if high transport costs prevented trade from breaking out.

But the natural trading partners hypothesis is altogether different and incorrect. There is, in fact, no reason to think that greater proximity increases the likelihood of gain for members in a PTA. This can be seen simply by constructing a counterexample where a union with a country (C) that is more distant produces more gain (for A) than a union with the country (B) that is less distant but otherwise identical (to C).

First note that as long as country A in figure 2.2 imports the good from both B and C in the pre- and post-FTA equilibrium, the presence of transportation costs has no effect whatsoever on the analysis based on that figure. All we need to do is to imagine that the supply price of C is inclusive of transport costs, while such costs are absent for the partner, B. This introduction of transport costs leaves the remainder of the analysis entirely unchanged.

To construct the counterexample noted above, consider a world consisting of three countries: A, B, and C. Country A has the option to form an FTA with either B or C. Countries B and C are identical in all respects except that the latter is located farther away. If the supply curves of B and C were horizontal, we would be in a world represented in panel a of figure 2.1 with $(P_C - P_B)$ representing transportation costs from C to A.

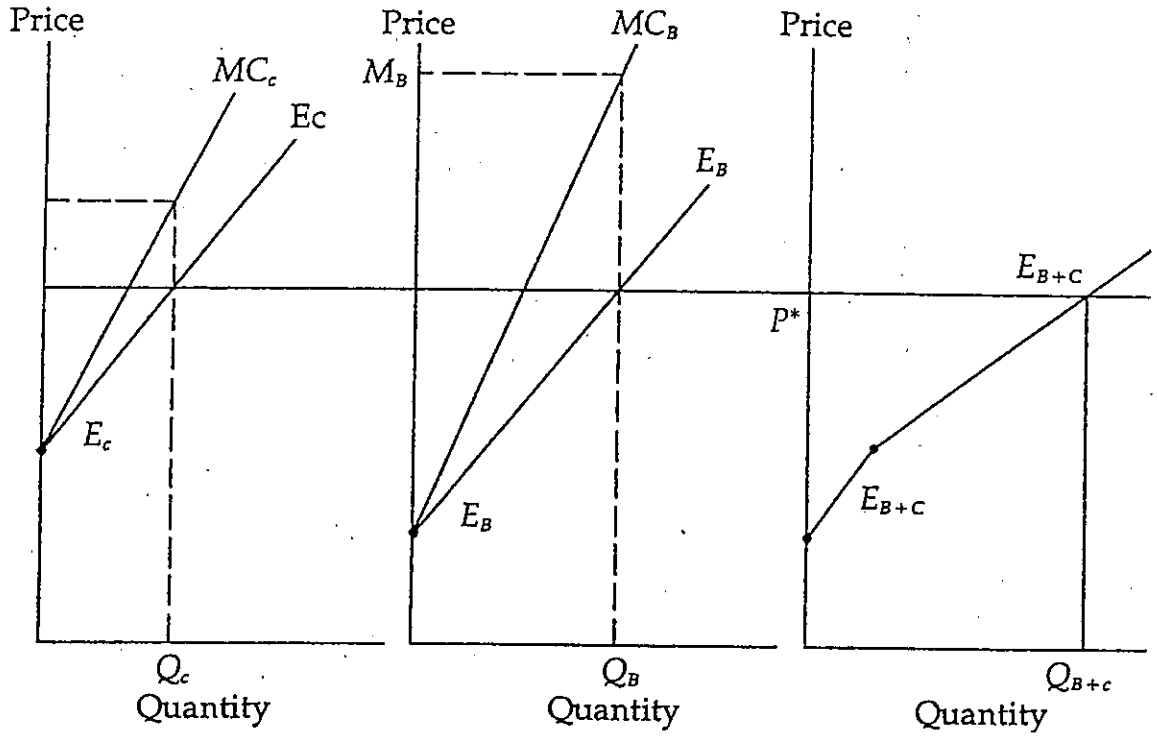


Figure 2.7
Example of positive effects of union with a more distant country.

Technically, in this case an FTA with the geographically proximate B improves A's welfare. But recall the limitation of such an FTA: country A does not trade with C before or after the union is formed; and in the post-FTA equilibrium, the external tariff does not matter so the FTA is really equivalent to nondiscriminatory free trade.

To make the example substantive, we must therefore assume that supply curves of B and C are upward sloped.[42] In figure 2.7, we draw three panels. In the first two panels, we show the export supply curves of C and B as E_CE_C and E_BE_B, respectively. In the third panel, we have their combined supply obtained by summing horizontally the individual supplies from the first two panels. The supply curves of C and B are identical in all respects except that C's supply price includes a constant per-unit transportation cost. Thus, for each quantity, C's supply price exceeds that of B by the per-unit transportation cost.[43]

To avoid clutter, we do not draw A's demand curve. Instead, imagine that there is an arbitrary nondiscriminatory tariff initially that yields the total demand for imports as represented by point Q_{B+C}. The price paid for this quantity to B and C is P^*. Individual supplies of B and C can be obtained by intersecting their supply curves with P^* and are shown by Q_B and Q_C. Not surprisingly, imports are larger from the geographically proximate country B than from C.

Now consider the introduction of preferential trading. To see which way preferences should be given, draw the *marginal* cost curve associated with each supply curve. These are shown by MC_B and MC_C. It is then immediate that, at the initial nondiscriminatory tariff, the marginal cost of imports is higher on imports from B than from C. We then obtain the dramatic conclusion that if A wants to give a tariff preference, it should opt for the distant partner C rather than the proximate B! The transport-cost criterion for choosing partners in a PTA is exactly wrong in this instance.

The explanation of this result is straightforward. The discriminating monopsonist model says that for any quantity of total purchases, the supplier with higher elasticity should be paid a higher price. In the present problem, this prescription translates into a lower tax on the supplier with higher elasticity. And transportation costs make C's supply curve more elastic than that of B.

Endogenous Tariffs on the Outside Country

So far, we have assumed that when an FTA is formed, the tariff on the outside country is held at its original level. But this may not always be true. When an FTA begins to take a bite, lobbies representing declining domestic industries may be able to reassert themselves. Because the FTA ties the authorities' hands with respect to the union partner, they will have to respond by raising protection against outside countries. This, indeed, happened recently following the [1994] Mexican crisis when the country raised external tariffs on 502 products from 20 percent or less to 35 percent!

This possibility had been anticipated by Bhagwati (1993a, 36–37). He wrote:

> Imagine that the United States begins to eliminate (by out competing) an inefficient Mexican industry once the FTA goes into effect. Even though the most efficient producer is Taiwan, if the next efficient United States out competes the least efficient Mexico, that would be desirable trade creation....
>
> But what would the Mexicans be likely to do? They would probably start AD actions against Taiwan.

This possibility raises the questions whether, once we allow for endogenous policy response, welfare may actually decline relative to the FTA and, indeed, to the initial equilibrium. Answers to both questions are in the affirmative.

A simple example demonstrating welfare deterioration relative to the FTA can be given as follows. For a zero tariff on B, calculate A's optimum

tariff on C. Suppose that A sets the initial, nondiscriminatory tariff on B and C at this level. Then, by construction, an FTA with B, holding C's tariff unchanged, not only improves A's welfare but actually maximizes it. If now lobbying pressure leads to a rise in the external tariff, A's welfare will necessarily fall.

The more interesting is the possibility that A's welfare can decline relative to the initial, pre-FTA equilibrium. To demonstrate it, note that A's welfare can be written as

$$\begin{aligned} W &= CS + PS + t_B P_B^* M_B + t_C P_C^* M_C \\ &= CS + PS + (P - P_B^*)M_B + (P - P_C^*)M_C \\ &= CS + PS + P(M_B + M_C) - (P_B^* M_B + P_C^* M_C), \end{aligned} \tag{2}$$

where CS denotes A's consumers' surplus, PS its producers' surplus, P domestic price, $P_i^*(i = \mathrm{B}, \mathrm{C})$ border price on imports from i, t_i the *ad valorem* tariff on imports from i, and M_i imports from i. The last two terms in these equalities represent tariff revenue on imports. Given a nondiscriminatory tariff initially, $P_B^* = P_C^*$.

Take the case favorable to an FTA with B by assuming that at each world price, B's supply is more elastic than C's. Assume further that the initial, nondiscriminatory tariff is sufficiently high that the FTA with no change in the tariff on C is welfare improving for A.[44] We will now show that if, because of lobbying pressure, the FTA is accompanied by a rise in the tariff on C such that *total* imports are unchanged, it is possible for its welfare to decline. Given that the FTA with no change in the tariff on C is welfare improving, this result shows that the endogenous tariff response can turn a welfare-improving FTA into a welfare-reducing proposition.

With no change in imports, the domestic price in A does not change and neither do CS and PS. From equation 2, it is then clear that welfare will rise or fall as the cost of imports, represented by the last term in the last equality, falls or rises. This property allows us to analyze the impact of the endogenous choice of the tariff by focusing on import supplies from B and C only.

In figure 2.8, as assumed, B's export-supply curve is more elastic than that of C at each price. This means that under a nondiscriminatory tariff, A's *private* marginal cost of obtaining imports from B is lower than that from C. Therefore, *at the margin*, A benefits by switching imports from C to B.

Initially, with a nondiscriminatory tariff, A buys the product at P^* per-unit from both B and C. Imports from the two countries are given by M_B and M_C, respectively. The marginal cost of obtaining imports from B is

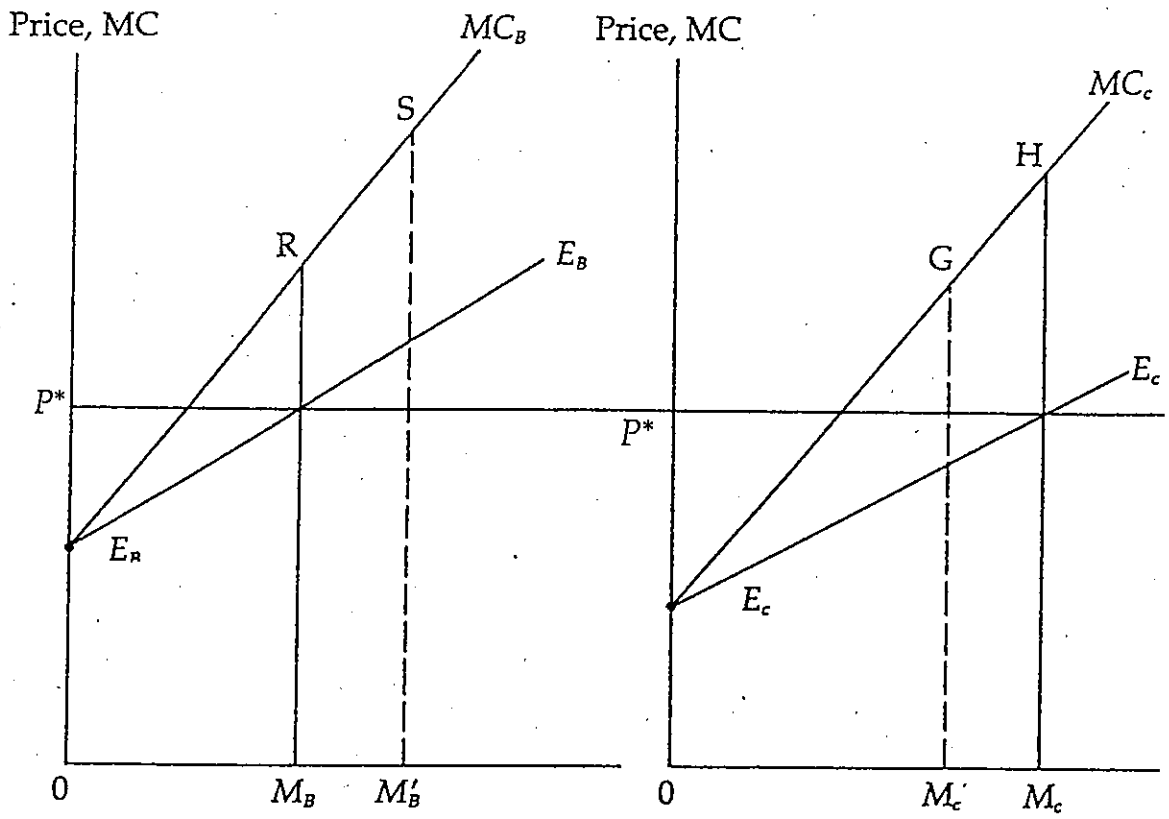


Figure 2.8
Welfare loss from endogenous increase in external tariff after formation of a free trade area.

less than that from C, $RM_B < HM_C$. As noted in the previous paragraph, at the margin, switching imports from C to B is beneficial to A: a small reduction in the tariff on B and increase in tariff on C which keeps total imports unchanged is welfare improving. But the FTA requires taking the tariff on B all the way to 0. As that is done, the marginal cost of obtaining imports from B rises, and as we correspondingly raise the tariff on C to keep the total imports unchanged, the marginal cost of imports from C falls. It is entirely possible that the two marginal costs cross and then reach levels such that the total cost of imports actually rises. Figure 2.8 is drawn on the assumption that the FTA increases imports from B by $M_BM'_B$. The tariff on C has to be raised to reduce imports from that country by an equivalent amount shown by $M_CM'_C$. As drawn, the net change in the cost of imports, $SRM_BM'_B$-$GHM_CM'_C$, is positive indicating that welfare declines.[45]

Welfare Loss without Trade Diversion

The general impression in the literature is that a welfare loss from an FTA can arise only if there is trade diversion. It is easy to show, however, that

a welfare loss to an *individual* member (though not to the union as a whole nor to the world) can arise even if there is no trade diversion. The simplest example of this phenomenon can be gleaned from figure 2.2. Starting from a nondiscriminatory tariff, marginal costs of production in B and C are equal. Given that *at the margin* both B and C are equally efficient suppliers of the product, there can be no trade diversion if we lower the tariff on B by an infinitesimally small amount. Yet because A's terms of trade with respect to B deteriorate by the full amount of the tariff reduction, it will lose from such a change.

In figure 2.2, because the domestic price does not change after the introduction of preferential trading, there is no trade creation. But if we allow C's supply curve to slope upward, the introduction of a small tariff preference for B will also generate a trade-creation effect. This is because the preference improves A's terms of trade with respect to C, lowers the domestic price, and displaces some inefficient domestic production. For reasons explained in the previous paragraph, there is no trade diversion, however. Yet it is possible for the loss from the accompanying deterioration in the terms of trade vis-à-vis B to more than offset the gain from trade creation as well as the improvement in the terms of trade vis-à-vis C (a result that can be derived algebraically, of course).

Concluding against PTAs

Our analysis of the static effects of PTAs is far less sanguine than is customarily assumed by several policy economists, bureaucrats, and politicians today. It also challenges and undermines the validity of the claims made in behalf of "regional" PTAs, whether the regions are defined in terms of countries with relatively high intraregional trade or in terms of proximity with or without common borders.

Therefore, if we were to assume that PTAs result from a variety of noneconomic factors, we need not be complacent about the possibility of their resulting in harmful effects.[46] Nor would there be any good reason to be complacent even if those PTAs were to be essentially regional in scope, when "regional" means geographic proximity or higher volumes of trade among, rather than outside, members.

We add three final observations. First, the common usage by journalists and politicians of the word "regional" frequently includes "common-ocean" arrangements such as APEC. Remember that APEC includes both South Korea and Chile, countries whose mutual trade is characterized by smallness of volume *and* largeness of distance, so that neither of the two

criteria of distance or volume of trade for sanctifying PTAs as desirable, inappropriate as we have shown it to be, holds for every member of APEC vis-á-vis every other.

Second, is the presence of common waters a new criterion for getting nations to form a PTA (the Pacific Ocean in the case of APEC)? We should not forget that the major oceans, and hence most of the trading nations of the world, are united by the world's water, and even more readily thanks to the Suez and Panama canals! In fact, the fullest-bodied common-waters "regional" area is clearly approximated by the membership of the WTO, as would have been appreciated by Ferdinand Magellan, who starting out from San Lucar in 1519 sailed from the Atlantic into the Pacific, an ocean unknown at the time.[47]

Third, the term "continental trading arrangements" has also been frequently used by Wei and Frankel (1995), who argue that "many [trading blocs] are along continental lines."[48] But this is at best misleading and at worst incorrect. Even if we confine ourselves to Article XXIV-sanctioned arrangements, we still must distinguish among PTAs that are continent-wide and hence "continental," those that cut across continents and are thus "intercontinental," and those that consist of members entirely *within*, but are not extended to *all* countries in, a continent and hence must be called "subcontinental."

Geographers and earth scientists divide the earth traditionally into four oceans (Arctic, Indian, Atlantic, and Pacific) and seven continents (Europe, Asia, Africa, Australia, North America, South America, and Antarctica). Only NAFTA and the PTA between Australia and New Zealand can then qualify as continental. And, the major new Article XXIV-sanctioned PTAs, which have been proposed by different groups in recent years (NAFTA extension into South America, APEC, and TAFTA) and which would clearly dwarf the continental PTAs clearly cut across continents.[49] Then again, MERCOSUR and ASEAN are clearly subcontinental. Of course, if one adds all the non-Article XXIV preferential trading arrangements, the matter looks even worse for those who claim that "many" of today's "trade blocs" are "continental."

Theoretical Analysis of the Dynamic Time-Path Question

Our analysis of the economics of PTAs would be seriously incomplete if, having analyzed the static effects, we did not go on to analyze the dynamic time-path question.

Formulating the Time-Path Question

Essentially, this question relates not to whether the immediate (static) effect of a PTA is good or bad, but whether the (dynamic) effect to the PTA is to accelerate or decelerate the continued reduction of trade barriers toward the goal of reducing them worldwide. This question may be formulated analytically in two separate ways.

Question I Assume that the time-path of MTN (multilateral trade negotiations) and the time-path of PTAs are separable and do not influence each other. The two policies are "strangers" to (that is, independent of) one another: neither hurts or helps the other. Will then the PTA time-path be characterized 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? A similar question can be raised for the MTN time-path. And 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 II Assume instead, as is more sensible, 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 impact on (be a "foe" of) the progress along the MTN time-path, or it will have a benign effect on (be a "friend" of) the MTN time-path.[50]

Question I can be illustrated with the aid of figure 2.9, which portrays a sample of possibilities for the time-paths in question. World (rather than individual member) welfare is put on the vertical axis and time along the horizontal axis. For the PTA time-paths drawn, an upward movement along the path implies growing membership; for the MTN (or what are described as "process-multilateralism") time-paths, it implies nondiscriminatory lowering of trade barriers among the nearly worldwide WTO membership instead. The PTA and MTN time-paths are assumed to be independent of each other; the PTA time-path neither accelerates nor decelerates the course of MTN (thus ruling out Question II-type issues). The goal can be treated as reaching U^*, the worldwide freeing of trade barriers on a nondiscriminatory basis at a specified time.

Question I can be illustrated by reference to the PTA paths I-IV. Thus, PTAs may improve welfare immediately, in the static sense, from U^0 to U_p^2 or reduce it to U_p^1. In either case, the time-path could then be stagnant

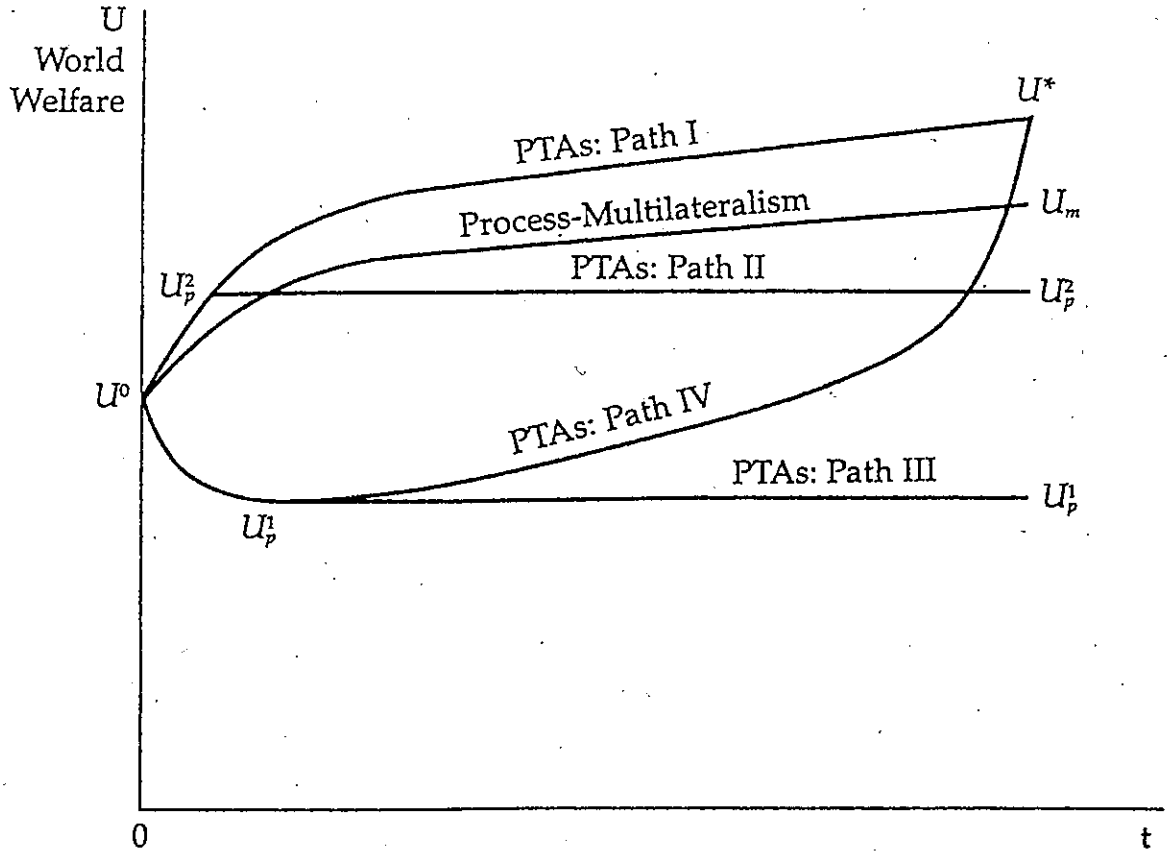


Figure 2.9
Alternative time-paths under multilateralism and under PTAs.

(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," that is, 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 II. In that case, the MTN time-path becomes a function of whether the PTA time-path is traveled simultaneously.

Question Originating in Policy

The dynamic time-path question has arisen, just as the static one did, in policy concerns and political decisions that ran ahead of the theory. The post-Vinerian, in-depth analysis of the static question coincided with the movement that eventually created the European Community through

the Treaty of Rome in 1957. The dynamic time-path question has arisen in the context of the U.S. failure to get an MTN Round started at the GATT Ministerial in 1982 and the U.S. decision to finally abandon its studied avoidance of Article XXIV-sanctioned PTAs. The policy choice made was initially Hobson's choice: if the MTN could not be used to continue lowering trade barriers, then PTAs would be used instead. If the turnpike could not be used, one had no option except to use the dirt road.

For several reasons that have been systematically explored in Bhagwati (1993b), the United States ended, however, becoming committed to "walking on both legs," embracing both the PTA and the MTN paths. Indeed, the United States has now become an active proponent of this view, continuing to do so even after the Uruguay Round of MTN had been successfully conducted and the WTO launched. And, in doing so, its spokespersons have frequently implied that PTAs will have a benign, beneficial impact on the worldwide lowering of trade barriers through induced acceleration of MTN.

The questions that we have posed above spring from this shift in U.S. policy, which has been manifest for several years, starting from the Bush administration and articulated as a distinct policy in the Clinton administration. In Bhagwati (1991, 1993), the challenge to international trade theorists to analyze these questions was identified and a preliminary set of arguments offered. We recapitulate briefly those arguments and then review the theoretical literature that has been developing since then on the dynamic time-path questions.

"Exogenously Determined" Time-Paths: A Diversion

First, however, it is necessary to consider and to turn aside certain theoretical approaches that are not meaningful for thinking about the dynamic time-path questions at hand, even though they have often been mistaken to be so.

Kemp-Wan The seminal approach of Kemp and Wan (1976) to Customs Union theory seems to be the most pertinent to our questions but, in fact, is not. Unlike the Vinerian approach, Kemp and Wan made the external tariff structure (of the Customs Union) endogenously chosen so that each member country's welfare would be improved, while that of the nonmembers was left unchanged. The beauty of this approach was that it restored, as it were, the commonsense intuition prior to Viner that a CU

should be welfare improving for members and for the world. This is, of course, a "possibility" theorem, no more and no less.[51]

It is then immediately apparent that the PTA time-path to U^* in figure 2.9 can be made monotonic, provided expanding membership of a PTA always satisfies the Kemp-Wan rule for forming a Customs Union. 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.

For that answer, to what is obviously Question I, we must turn to the *incentive structure* that any CU/PTA provides, through interests, ideology, and institutions, for expansion or stagnation of its membership.

Krugman The same argument applies to the theoretical approach to the question of PTAs recently introduced by Paul Krugman (1991a, 1991b, 1993). Again the expansion of membership is treated as exogenously specified, as in Viner, and the welfare consequences of the world mechanically dividing into a steadily increasing number of symmetric blocs—clearly demarcated countries are then not even the natural constituents of these "blobs"-cum-blocs—are considered and, for particular specifications, the monotonicity of world welfare examined, including even calculations concerning the "optimal" number of such symmetric PTAs/blocs! This, in turn, has led to critiques, as of the symmetry assumption by Srinivasan (1993), who essentially shows that the specific Krugman conclusions are easily reversed by abandoning symmetry, and to further variations by a few others.[52] Yet it is hard to see the analytical interest of this approach or, more important, its relevance to the compelling (incentive-structure) questions today concerning the membership expansion of PTAs. In short, it fails to throw light on the analysis of the dynamic time-path questions of the type introduced above. For that analysis which is currently, quite correctly, on the top of the theoretical agenda, we must turn elsewhere.

Incentive Structure Arguments

At the April 1992 World Bank Conference on Regional Integration, Bhagwati (1993a), having reiterated the need to analyze the dynamic time-path question, advanced several arguments concerning the incentive structure within specific PTAs, once formed, to expand or to stagnate. Before we discuss the theoretical modeling of such ideas by Baldwin (1993), Krishna (1993), and Levy (1994), among others, it is worth recapitulating the principal arguments distinguished by Bhagwati.[53]

We need to recognize, of course, that the incentives may be political rather than (narrowly) economic. A PTA may be formed, and even expanded, to seek political allies by using trade as foreign policy and to target the benefits of trade to politically favored nations.[54] Politics is not a negligible factor in the discriminatory trade arrangements implemented by the EU via Association Agreements with the smaller countries on its periphery and beyond; and it certainly cannot be ignored in the transformation of the original Canada-U.S. Free Trade Agreement into NAFTA with Mexico and then into the Enterprise of Americas Initiative.

But that is clearly not the whole story, and we can learn much by thinking carefully about the incentive structure for membership expansion in political-economy-theoretic terms. To do this, Bhagwati (1993a) distinguished among three different types of "agents" and offered the following analysis.

Governments of Member Countries PTAs will be under pressure not to expand because governments may feel that "we already have a large market, so what do we stand to gain by going through the hassle of adding more members?" This is the "our market is large enough" syndrome, emphasized by Martin Wolf, who has often noted that large countries have tended to opt for inward-looking trade and investment strategies, while the small ones have gone the outward-looking route.

Interest Groups in Member Countries The interest groups in member countries may be for or against new members. The internationally oriented exporting firms may be expected to endorse new members whose markets then become preferentially available to them vis-à-vis nonmember exporters to these new members.[55] On the other hand, the firms that are profiting from access to preferential markets in the member (partner) countries will not want new members whose firms are also exporters of the same or similar products in the member markets. Both incentives reflect the preferential nature of the PTAs.

The former incentive was clear in the NAFTA debate in the United States and reflected in many pronouncements, including that of pro-NAFTA economists (and even President Clinton, who played the Japanophobic card that the United States would have preferential access to Mexico vis-à-vis Japan). It is also evident in the statement of Signor Agnelli of Fiat: "The single market must first offer an advantage to European companies. This is a message we must insist on without hesitation."

Interest Groups in Nonmember Countries The third set of agents is in the nonmember countries. Here the example of a PTA may lead others to emulate, even to seek, entry. Then again, the fear of trade diversion may also induce outsiders to seek entry.[56]

Recent Theoretical Analyses

Subsequently, the analysis of the dynamic time-path question moved into formal political economy-theoretic modeling. We provide here a synoptic review of the few significant contributions to date, organizing the literature analytically in light of the two questions distinguished above and also in terms of whether the analysis models the incentives of nonmembers to join or those of the members to expand.[57]

Question I The single contribution that focuses on Question I (the incentive to add members to a PTA) is by Richard 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 tariffs that member countries' firms do not have to pay. These firms then lobby for entry, tilting the political equilibrium at the margin toward entry demands in their countries. The countries closest to the margin will then enter the bloc, assuming that the members have open entry. This enlarges the market and thereby increases the cost of nonmembership and pulls in countries at the next margin. Given the assumptions, including continuity, this domino model can take the PTA time-path to U^* in figure 2.9.

While Baldwin formalizes the incentive of nonmembers to get inside the PTA, interestingly there is no formalization of the incentives of members to add or reject new members that have been discussed in the literature, as by Bhagwati (1993a). Indeed, the Baldwin model itself shows, on the flip side, that member firms will gain from the cost advantage that they enjoy vis-à-vis the nonmember firms and hence will have an opposed interest in not admitting the nonmembers to the PTA: a full analysis of the political economy of both members and nonmembers in the Baldwin model could then lead to specific equilibrium outcomes that leave the PTA expansion imperiled.

Question II The rest of the theoretical contributions address Question II, that is, whether the PTA possibility and/or time-path helps or harms the MTN time-path. Pravin Krishna (1993) and Philip Levy (1994) address directly and quite aptly this question and reach the "malign-impact" conclusion, unfavorable to the exhortation to "walk on both legs."

Krishna models the political process in the fashion of the government acting in response to implicit lobbying by firms, what Bhagwati (1990) has called "clearinghouse"—government assumption where the government is passive, as in Findlay and Wellisz (1982). Krishna shows in his oligopolistic-competition model that the bilateral PTA between two member countries reduces the incentive of the member countries to liberalize 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 liberalization.

Levy models the political process instead in a median-voter model à la Mayer (1984); the government is not what Bhagwati (1990) has called "self-willed" with its own objectives but acts again as a clearing-house. Using a richer model with scale economies and product variety, Levy demonstrates 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 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. However, we might also note that there are contributions that take the more conventional view of governments, which act as agents maximizing social welfare (so that they may be regarded as acting as the custodians of the "general interest" as defined by economists), but then ask whether the effect of allowing PTAs to form affects outcomes concerning trade policy relating to the multilateral system. Rodney Ludema (1993) has analyzed the effect of PTAs on multilateral bargaining outcomes, arguing plausibly that the PTAs give strategic advantage to their members, whereas Kyle Bagwell and Robert Staiger (1993) have analyzed how the formation of a PTA—distinguishing between an FTA and a CU, as they yield different answers—will affect the (unbound) tariffs of the member countries on nonmembers.

The Sequential Bargaining Argument

In conclusion, we note that a different kind of model is implied, though not yet formalized, by the recent argument of Bhagwati (1994) that combines three separate notions.

The first is that even though a multilateral bargain *simultaneously* with a group of nonhegemonic powers is profitable and hence possible, a hegemonic power will gain a greater payoff by bargaining *sequentially* with them, using bilateral and plurilateral PTA approaches, picking the countries that are the most vulnerable and then moving on to the next one and so on.[58]

The second is that this insight has now been appreciated by several lobbies (for example, the intellectual property protection lobby, the environmental and labor standard groups), which are piggybacking on to trade liberalization and trade institutions to secure their maximalist objectives and which see that the PTA approach (which may be seen as an "incentive" strategy), combined with the occasional use of aggressive unilateralism à la punitive Section 301 actions (which may be seen as a "punishment" strategy), is more likely to procure their objectives at the WTO and multilaterally than if pursued directly there through MTN alone.

The third is that the two processes, the MTN and the PTA paths, are to be traveled in tandem since the ultimate goal is indeed to arrive at multilateral, universal obligations in the areas desired by these lobbies by the nonhegemonic powers.

If this "model" provides insight into the political process driving the legitimation of the PTA time-path, then no hegemonic power is likely to abandon the PTA path simply because the WTO exists and is jump-started. A "selfish hegemon," looking after its own narrowly defined interests, reflecting its own lobbying-derived needs, will indeed want to "walk on both legs." But the multilateral outcome, so affected and determined, need not then be considered to have been affected in the socially optimal direction unless one makes the assumption, made effortlessly by hegemonic spokesmen in their policy pronouncements, that "what is good for the hegemon (and its lobbies) is good for the world trading system." Indeed, when we see that the intellectual property protections that were built into the WTO are almost certainly excessive according to the analytical and empirical argumentation of many of the best international economists today, it is hard to regard the ability of the hegemon to induce such outcomes with the aid of PTAs (and aggressive unilateralism) as creating a "benign" effect of the PTAs on the MTN path.

Implications for Current Policy

The case for PTAs, whether on static or on dynamic grounds, appears far less compelling and attractive than many politicians and policy-makers now believe. In fact, it is likely that most of them, misled by the inevitable confusion between free trade and free trade areas that some economists have wittingly or unwittingly encouraged, are not even aware that the scholarly scene is rife today with serious opposition to PTAs.[59]

The Politics of PTAs

The current preoccupation with PTAs reflects overriding political factors. Recall our earlier discussion of the sequential-bargaining advantage to hegemonic powers. Or consider the fact that the leaders of the smaller, nonhegemonic powers get to play a more prominent role, with better photo-opportunities, with smaller summits, especially when a hegemonic power such as the United States features its own president, than would ever be the case at the WTO. Or consider that where the PTAs are regional, as is MERCOSUR (among Argentina, Brazil, Paraguay, and Uruguay), the discriminatory trade agreement can be depicted politically as an act of foreign policy statesmanship. Or consider simply the operation of Gresham's Law: PTAs by some encourage PTAs by others, especially when they are being continually misportrayed by other politicians and countries as statesmanlike moves to free trade. And, of course, there are always the amateur geopoliticians and geoeconomists. Like little boys playing Nintendo games on their computer screens, they think of playing the game of "trade blocs" to indulge their pet prejudices against Europe or Japan. Some want to make the APEC into a PTA to play off against a "protectionist" Europe, while others think of TAFTA as a weapon to play off against the "unfairly trading" Japan.[60]

The "Spaghetti Bowl" Phenomenon

Our view, for reasons explored fully in this chapter, is that the spread of PTAs is desirable only when two justifications obtain: you are building a Common Market with full-scale integration of factor markets and even political harmonization; or the multilateral MFN, MTN process is not working. As we argued earlier, neither rationale is operative today.

In fact, the proliferation of PTAs today poses the danger, indeed the certainty, that a veritable "spaghetti bowl" phenomenon, as Bhagwati

has called it, will emerge where trade barriers, including duties, will vary depending on origin, and complex and protection-accommodating rules of origin will find their way into practice.[61] And this, too, at a time when multinationals are getting truly global, and the identification of "local content" and hence origin of traded goods and services is becoming increasingly meaningless and hence subject to inevitable arbitrariness. PTAs are just one, and indeed a gigantic, step backward from this reality: the need today is to intensify the commitment to the basic tenet of nondiscrimination that the architects of GATT correctly saw as a principal virtue, not to undermine it.

PTAs with and among Hegemons

We would therefore suggest that Article XXIV-sanctioned PTAs that involve hegemonic powers should be actively discouraged. They involve NAFTA extension southward or overseas, EU free-trade-area agreements with non-EU countries, APEC's transformation into a PTA, and TAFTA.

Such a self-denial would appear anti-free-trade, given the current state of confused thinking and the political capital invested by many in the cause of the PTAs. But it would be speaking to a far more compelling, and truer, version of free trade. It would also require true statesmanship on the part of the leaders of the hegemonic powers, as against the political advantages of opting for what is an inefficient and indeed harmful option.

PTAs among the Nonhegemons

Our view of PTAs among the nonhegemons, principally developing countries, is just a trifle less critical.

To begin with, what MERCOSUR does, for example, has only a fraction of the significance that the United States and the European Union have individually. The trade policy choices of the nonhegemons have comparatively more consequences for themselves than for the world. This contrast is sufficient to regard what they are doing with a less fiercely critical eye than that directed at the hegemons.

Remember again that the impact on their own welfare of PTAs is not necessarily benign. Especially, when these countries get into a PTA with hegemonic powers (for example, Mexico joining the United States in NAFTA), the outcomes for them may be welfare worsening (in the static sense) because of the tariff-revenue-redistribution effect, among other reasons. Failure to understand the differential economics of PTAs, as con-

trasted with that of free trade, underlies many of the favorable assessments often advanced in behalf of the developing countries that seek to join PTAs with the hegemonic powers.[62] A similar caveat would be relevant to PTAs among the nonhegemons themselves.

We may still consider these PTAs, such as MERCOSUR, with some favor, although nondiscriminatory free trade is the best option. After all, the acceptance of Article XXIV discipline (imperfect as it is) is an improvement over protectionism or over the utterly chaotic and arbitrary ECDC (economic cooperation among developing countries) at the GATT under which these countries were free from such discipline and could indulge in any level and kind of preferences among themselves.

Conclusion

At present, the spread of hegemonic PTAs has been halted. The Osaka meeting of APEC in November 1995 witnessed the Asian members of APEC reaffirming their desire to stick to MFN and hence implicitly to reject the PTA approach even though the U.S. position on the issue apparently remained problematic and ambiguous (with several pro-PTA proponents in the administration). Equally, at Madrid, the idea of TAFTA has been deflected away from an Article XXIV agreement to the New Trans-Atlantic Agenda that merely seeks, and in a presumably nondiscriminatory fashion, the lowering of trade and investment barriers in the area. For the time being, the extension of NAFTA to the South has also been halted, for reasons that may not hold for long beyond the presidential election in 1996.

All this yields enough time to take a closer look at the dangerous drift to PTAs that has been aided by the unfortunate conversion of the United States to the thesis that any trade liberalization is as good as any other. Perhaps, as often happens in economic policy, what presently looks like a politically irreversible trend will yield to economic wisdom. We will see.

Appendix 2A1: Varieties of PTAs within the World Trade Organization

There are three categories of PTAs within the World Trade Organization framework. First, under Article XXIV, countries can form Free Trade Areas or Customs Unions. Of the one hundred thirty-four arrangements notified to the GATT/WTO as of June 1, 1995, 108 fell under this category (see table 2A.1). Second, developing countries can form PTAs under the

Table 2A.1
134 Regional Trading Arrangements Notified to the GATT/WTO, 1949–1995

Official title	Usual reference	Date of entry into force	GATT cover
Interim Agreement for a Customs Union between the Union of South Africa and Southern Rhodesia	South Africa–South Rhodesia Customs Union	Apr. 1, 1949	Article XXIV
Free-Trade Treaty between the Republics of Nicaragua and El Salvador	El Salvador–Nicaragua Free Trade Area	Aug. 21, 1951	Article XXIV
Rome Treaty (European Economic Communities and European Atomic Energy Community)	EEC and EURATOM	Jan. 1, 1958	Article XXIV
Multilateral Central American Free Trade and Economic Integration Treaty (participation of Nicaragua)	Central American Free Trade Area	June 2, 1959	Article XXIV
Stockholm Convention (European Free Trade Association)	EFTA	May 3, 1960	Article XXIV
The Montevideo Treaty (Latin American Free Trade Area)	LAFTA	June 2, 1961	Article XXIV
Association of Finland with the European Free Trade Association	EFTA–Finland Association (FINEFTA)	June 26, 1961	Article XXIV
General Treaty for Central American Economic Integration (participation of Nicaragua)	Central American Common Market	Oct. 12, 1961	Article XXIV
The Borneo Free Trade Area	Borneo Free Trade Area	Jan. 1, 1962	Article XXIV
Trade Agreement between the Republic of Ghana and the Republic of Upper Volta	Ghana–Upper Volta Trade Agreement	May 9, 1962	Article XXIV
Regulation of Economic and Customs Relations between the Member States of the Equatorial Customs Union and the Federal Republic of Cameroon	Equatorial Customs Union—Cameroon Association	July 1, 1962	Article XXIV
Agreement setting up an association between the European Economic Community and Greece	EEC–Greece Association Agreement	Nov. 1, 1962	Article XXIV
African Common Market	African Common Market	June 1, 1963	Article XXIV
Convention of Association between the European Economic Community and the African and Malagasy States Associated with that community	Yaoundé I	Jan. 1, 1964	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Agreement for Economic Unity among Arab League States	Arab Common Market	Apr. 30, 1964	Article XXIV
Association between the EEC and certain non-European Countries and Territories maintaining special relations with France and the Netherlands, "PTOM-I"	EEC–PTOM I	June 1, 1964	Article XXIV
Agreement creating an association between the European Economic Community and Turkey; "The Ankara Agreement"	EEC–Turkey Association Agreement of 1963	Dec. 1, 1964	Article XXIV
New Zealand/Australia Free Trade Agreement	Australia–New Zealand Free Trade Agreement	Jan. 1, 1966	Article XXIV
United Kingdom/Ireland Free Trade Area Agreement	Ireland–United Kingdom Free Trade Area	July 1, 1966	Article XXIV
Agreement Establishing the Caribbean Free Trade Association	CARIFTA	May 1, 1968	Article XXIV
Agreement establishing an association between the European Economic Community and the Kingdom of Morocco	EEC–Morocco Association Agreement of 1969	Sept. 1, 1969	Article XXIV
Agreement establishing an association between the European Economic Community and the Republic of Tunisia	EEC–Tunisia Association Agreement of 1969	Sept. 1, 1969	Article XXIV
European Free Trade Association; Accession of Iceland	EFTA/FINEFTA–Iceland Accession	Mar. 1, 1970	Article XXIV
Agreement between the European Economic Community and the State of Israel	EEC–Israel Agreement of 1970	Oct. 1, 1970	Article XXIV
Agreement between the European Economic Community and Spain	EEC–Spain Agreement of 1970	Oct. 1, 1970	Article XXIV
Agreement establishing an Association between the European Economic Community and the United Republic of Tanzania, the Republic of Uganda, and the Republic of Kenya	Arusha II Agreement	Jan. 1, 1971	Article XXIV
Association between the EEC and Certain Non-European Countries and Territories	EEC–PTOM II	Jan. 1, 1971	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
EEC; African and Malagasy states and overseas countries and territories agreements	Yaoundé II	Jan. 1, 1971	Article XXIV
Agreement Establishing an Association between Malta and the European Economic Community	EEC–Malta Association Agreement	Apr. 1, 1971	Article XXIV
Agreements between Austria and the European Communities	EC–Austria Agreements of 1972	Oct. 1, 1972	Article XXIV
Treaty concerning the accession of the Kingdom of Denmark, Ireland, the Kingdom of Norway, and the United Kingdom of Great Britain and Northern Ireland	EC–Accession of Denmark, Ireland and United Kingdom	Jan. 1, 1973	Article XXIV
Agreements between the European Communities and Portugal	EC–Portugal Agreements of 1972	Jan. 1, 1973	Article XXIV
Agreements between the European Communities and Sweden	EC–Sweden Agreements	Jan. 1, 1973	Article XXIV
Agreement between the European Economic Community and the Swiss Confederation	EC–Switzerland/ Liechtenstein Agreements	Jan. 1, 1973	Article XXIV
EEC; Turkey additional protocol to the Association Agreement	EEC–Turkey Additional Protocol to the Association Agreement	Jan. 1, 1973	Article XXIV
Agreement between the European Economic Community and the Republic of Iceland	EC–Iceland Agreements	Apr. 1, 1973	Article XXIV
Agreement between the European Economic Community and Cyprus	EEC–Cyprus Association Agreement	June 1, 1973	Article XXIV
Agreement between the European Economic Community and the Kingdom of Norway	EC–Norway Agreements	July 1, 1973	Article XXIV
Treaty establishing the Caribbean Community	CARICOM	Aug. 1, 1973	Article XXIV
Agreement between the European Economic Community and the Arab Republic of Egypt	EEC–Egypt Agreement of 1972	Nov. 1, 1973	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Agreement between the European Economic Community and the Lebanese Republic	EEC–Lebanon Agreement of 1972	Nov. 1, 1973	Article XXIV
Agreements between the European Communities and Finland	EC–Finland Agreements	Jan. 1, 1974	Article XXIV
Supplementary protocol to the Association Agreement between the European Economic Community and Turkey consequent on the accession of new member states to the Community	EC–Turkey Association Agreement of 1973	Jan. 1, 1974	Article XXIV
Agreement between the Republic of Finland and the People's Republic of Bulgaria on the reciprocal removal of obstacles to trade	Bulgaria–Finland Agreement	Jan. 1, 1975	Article XXIV
Agreement between the Republic of Finland and the Czechoslovak Socialist Republic on the reciprocal removal of obstacles to trade	Finland–Czechoslovakia Agreement	Jan. 1, 1975	Article XXIV
Agreement between the Republic of Finland and the Hungarian People's Republic on the reciprocal removal of obstacles to trade	Finland–Hungary Agreement	Jan. 1, 1975	Article XXIV
Additional protocol to the agreement establishing an association between the European Economic Community and Greece consequent on the accession of new member states to the Community	EEC–Greece Additional Protocol	July 1, 1975	Article XXIV
Agreement between the European Economic Community and the State of Israel	EEC–Israel Agreement of 1975	July 1, 1975	Article XXIV
Agreement between the Republic of Finland and the German Democratic Republic on the removal of obstacles to trade on the basis of Reciprocity concerning advantages and obligations	Finland–German Democratic Republic Agreement	July 1, 1975	Article XXIV
ACP; EEC First Convention of Lomé	First Convention of Lomé	Apr. 1, 1976	Article XXIV
Interim agreement between the European Economic Community and the Peoples Democratic Republic of Algeria	EC–Algeria Agreements of 1976	July 1, 1976	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Interim agreement between the European Economic Community and the Kingdom of Morocco	EC–Morocco Agreements	July 1, 1976	Article XXIV
Interim agreement between the European Economic Community and the Republic of Tunisia	EC–Tunisia Agreements of 1976	July 1, 1976	Article XXIV
Interim agreement between the European Economic Community and the Portuguese Republic	EEC–Portugal Interim Agreement	Nov. 1, 1976	Article XXIV
Australia–Papua New Guinea Trade and Commercial Relations Agreement (PATCRA)	Australia–Papua New Guinea Agreement (PATCRA)	Feb. 1, 1977	Article XXIV
Interim cooperation agreement between the European Communities and the Arab Republic of Egypt	EEC–Egypt Interim Agreement of 1977	July 1, 1977	Article XXIV
Agreement between the European Economic Community and Jordan	EEC–Jordan Interim Agreement of 1977	July 1, 1977	Article XXIV
Agreement between the European Economic Community and Lebanon	EEC–Lebanon Interim Agreement of 1977	July 1, 1977	Article XXIV
Agreement between the European Economic Community and Syria	EEC–Syria Interim Agreement of 1977	July 1, 1977	Article XXIV
Agreement between the Republic of Finland and the Polish People's Republic on the reciprocal removal of obstacles to trade	Finland–Poland Agreement	Apr. 1, 1978	Article XXIV
EFTA–Spain Agreement	EFTA–Spain Agreement	May 1, 1980	Article XXIV
Interim agreement between the European Economic Community and the Socialist Federal Republic of Yugoslavia on trade and trade cooperation	EEC–Yugoslavia Interim Agreement	July 1, 1980	Article XXIV
EEC–Greece Accession Agreement	EEC–Greece Accession Agreement	Jan. 1, 1981	Article XXIV
ACP; EEC Second Convention of Lomé	Second Convention of Lomé	Jan. 1, 1981	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Australia–New Zealand Closer Economic Relations Trade Agreement (ANZCERTA)	Australia–New Zealand (ANZCERTA)	Jan. 1, 1983	Article XXIV
Agreement on the establishment of a free trade area between the Government of the United States of America and the Government of Israel	Israel–United States Free Trade Area Agreement	Aug. 19, 1985	Article XXIV
Accession of Portugal and Spain to the European Communities	EEC–Portugal and Spain Accessions	Jan. 1, 1986	Article XXIV
ACP; EEC Third Convention of Lomé	Third Convention of Lomé	Mar. 1, 1986	Article XXIV
Canada–United States Free Trade Agreement	Canada–U.S. Free Trade Agreement	Jan. 1, 1989	Article XXIV
Agreement between the European Community, of the one part, and the Government of Denmark and the Home Government of the Faroe Islands, of the other part	EC–Denmark and Faroe Islands Agreement	Jan. 1, 1992	Article XXIV
Interim agreement on trade and trade-related matters between the European Economic Community and the ECSC, of the one part, and the Czech and Slovak Federal Republic (CSFR), of the other part	EC–Czech and Slovak Federal Republic Interim Agreement of 1991	Mar. 1, 1992	Article XXIV
Interim agreement on trade and trade-related matters between the European Economic Community and the European Coal and Steel Community, of the one part, and Hungary, of the other part	EC–Hungary Interim Agreement of 1991	Mar. 1, 1992	Article XXIV
Interim agreement on trade and trade-related matters between the European Economic Community and the European Coal and Steel Community, of the one part, and Poland, of the other part	EC–Poland Interim Agreement of 1991	Mar. 1, 1992	Article XXIV
Agreement between the EFTA states and Turkey	EFTA–Turkey Agreement	Apr. 1, 1992	Article XXIV
Free trade agreement between the Kingdom of Norway and the Republic of Estonia	Estonia–Norway Free Trade Agreement	June 15, 1992	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Free trade agreement between the Kingdom of Norway and the Republic of Latvia	Latvia–Norway Free Trade Agreement	June 16, 1992	Article XXIV
Free trade agreement between the Kingdom of Norway and the Republic of Lithuania	Lithuania–Norway Free Trade Agreement	June 16, 1992	Article XXIV
Agreement between the EFTA states and the Czech and Slovak Federal Republic	Czech and Slovak Federal Republic–EFTA Agreement	July 1, 1992	Article XXIV
Free trade agreement between the Kingdom of Sweden and the Republic of Estonia	Estonia–Sweden Free Trade Agreement	July 1, 1992	Article XXIV
Free trade agreement between the Kingdom of Sweden and the Republic of Latvia	Latvia–Sweden Free Trade Agreement	July 1, 1992	Article XXIV
Free trade agreement between the Kingdom of Sweden and the Republic of Lithuania	Lithuania–Sweden Free Trade Agreement	July 1, 1992	Article XXIV
Estonia–Finland protocol regarding temporary arrangements on trade and economic cooperation	Estonia–Finland Agreement	Dec. 1, 1992	Article XXIV
Czech Republic and Slovak Republic Customs Union	Czech Republic and Slovak Republic Customs Union	Jan. 1, 1993	Article XXIV
EFTA–Israel Free Trade Agreement	EFTA–Israel Free Trade Agreement	Jan. 1, 1993	Article XXIV
Central European Free Trade Agreement concluded by the Czech Republic, the Republic of Hungary, the Republic of Poland and the Slovak Republic	CEFTA	Mar. 1, 1993	Article XXIV
Free Trade Agreement between the Swiss Confederation and the Republic of Estonia	Estonia–Switzerland Free Trade Agreement	Apr. 1, 1993	Article XXIV
Free Trade Agreement between the Swiss Confederation and the Republic of Latvia	Latvia–Switzerland Free Trade Agreement	Apr. 1, 1993	Article XXIV
Free Trade Agreement between the Swiss Confederation and the Republic of Lithuania	Lithuania–Switzerland Free Trade Agreement	Apr. 1, 1993	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Interim agreement on trade and trade-related matters between the European Economic Community and the European Coal and Steel Community, of the one part, and Romania, of the other part	EEC–Romania Interim Agreement	May 1, 1993	Article XXIV
Agreement between the EFTA states and Romania	EFTA–Romania Agreement	May 1, 1993	Article XXIV
EFTA–Bulgaria Free Trade Agreement	EFTA–Bulgaria Free Trade Agreement	July 1, 1993	Article XXIV
Finland–Latvia protocol regarding temporary arrangements on trade and economic cooperation	Finland–Latvia Protocol	July 1, 1993	Article XXIV
Finland–Lithuania Protocol regarding Temporary arrangements on trade and economic cooperation	Finland–Lithuania Protocol	July 1, 1993	Article XXIV
Cooperation agreement between the European Economic Community and the Republic of Slovenia	EEC–Slovenia Cooperation Agreement	July 19, 1993	Article XXIV
Agreement between the EFTA states and the Republic of Hungary	EFTA–Hungary Agreement	Oct. 1, 1993	Article XXIV
Agreement between the EFTA states and the Republic of Poland	EFTA–Poland Agreement	Nov. 15, 1993	Article XXIV
Interim agreement on trade and trade-related matters between the European Economic Community and the ECSC, of the one part, and the Republic of Bulgaria, of the other part	EEC–Bulgaria Interim Agreement	Dec. 31, 1993	Article XXIV
Free trade agreement between the Czech Republic and the Republic of Slovenia	Czech Republic–Slovenia Free Trade Agreement	Jan. 1, 1994	Article XXIV
North American Free Trade Agreement	NAFTA	Jan. 1, 1994	Article XXIV
Free trade agreement between the Slovak Republic and the Republic of Slovenia	Slovak Republic–Slovenia Free Trade Agreement	Jan. 1, 1994	Article XXIV
Austria, Finland, Sweden–EU accession agreement	Austria, Finland, Sweden–EU Accession Agreement	Jan. 1, 1995	Article XXIV

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Czech Republic–Romania Free Trade Agreement	Czech Republic–Romania Free Trade Agreement	Jan. 1, 1995	Article XXIV
Agreement on free trade and trade-related matters between the European Community, the European Atomic Energy Community and the ECSC, of the one part, and the Republic of Estonia, of the other part	EC–Estonia Agreement	Jan. 1, 1995	Article XXIV
Agreement on free trade and trade-related matters between the European Community, the European Atomic Energy Community and the ECSC, of the one part, and the Republic of Latvia, of the other part	EC–Latvia Agreement	Jan. 1, 1995	Article XXIV
Agreement on free trade and trade-related matters between the European Community, the European Atomic Energy Community and the ECSC, of the one part, and the Republic of Lithuania, of the other part	EC–Lithuania Agreement	Jan. 1, 1995	Article XXIV
Free trade agreement between the Republic of Hungary and the Republic of Slovenia	Hungary–Slovenia Free Trade Agreement	Jan. 1, 1995	Article XXIV
Slovak Republic–Romania Free Trade Agreement	Slovak Republic–Romania Free Trade Agreement	Jan. 1, 1995	Article XXIV
EFTA–Slovenia Free Trade Agreement	EFTA–Slovenia Free Trade Agreement	June 1, 1995	Article XXIV
The Unified Economic Agreement among the countries of the Gulf Cooperation Council	Gulf Cooperation Council		Enabling Clause
Additional protocol on preferential tariffs among members of the organization for economic cooperation (ECO)	Preferential Tariffs among ECO-members		Enabling Clause
South Asian Association for Regional Cooperation; Preferential Trade Arrangement (SAPTA)	SAPTA		Enabling Clause

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Protocol relating to trade negotiations among developing countries	Protocol relating to Trade Negotiations among Developing Countries	Feb. 11, 1973	Enabling Clause
First agreement on trade negotiations among developing member countries of the Economic and Social Commission for Asia and the Pacific	Bangkok Agreement	June 17, 1976	Enabling Clause
Association of South-East Asian Nations ASEAN declaration	ASEAN Preferential Trading Arrangements	Aug. 31, 1977	Enabling Clause
South Pacific Regional Trade and Economic Cooperation Agreement	SPARTECA	Jan. 1, 1981	Enabling Clause
Second Treaty of Montevideo	Latin American Integration Association, "LAIA"	Mar. 18, 1981	Enabling Clause
Cartagena Agreement	Andean Group	May 25, 1988	Enabling Clause
Global System of Trade Preferences among developing countries (GSTP)	GSTP	Apr. 19, 1989	Enabling Clause
Trade agreement between the government of the Kingdom of Thailand and the government of the Lao People's Democratic Republic	Lao–Thailand Trade Agreement	June 20, 1991	Enabling Clause
Treaty of Asunción, Treaty Establishing a Common Market between the Argentine Republic, the Federal Republic of Brazil, the Republic of Paraguay, and the Eastern Republic of Uruguay	MERCOSUR	Nov. 29, 1991	Enabling Clause
Common effective preferential tariff scheme for the ASEAN Free Trade Area	Preferential Tariff Scheme for the ASEAN Free Trade Area	Jan. 28, 1992	Enabling Clause
Common Market for Eastern and Southern Africa	COMESA	Dec. 8, 1994	Enabling Clause
Bolivia-Mexico Free Trade Treaty	Bolivia-Mexico Free Trade Treaty	Jan. 1, 1995	Enabling Clause
Mexico-Costa Rica Free Trade Area	Mexico–Costa Rica Free Trade Area	Jan. 1, 1995	Enabling Clause

Table 2A.1 (continued)

Official title	Usual reference	Date of entry into force	GATT cover
Colombia, Mexico, and Venezuela Free Trade Agreement	Treaty of the Group of Three (G3)	Jan. 1, 1995	Enabling Clause
Australian treatment of products of Papua New Guinea	Australian Treatment of Products of Papua New Guinea		Waiver–Art. I:1
French trading arrangements with Morocco	France–Morocco Trading Arrangements		Waiver–Art. I:1
	Australia–Federation of Rhodesia and Nyasaland Agreement	July 1, 1955	Waiver–Art. I:1
United States Caribbean Basin Economic Recovery Act	U.S.–Caribbean CBERA	Jan. 1, 1984	Waiver–Art. I:1
Canadian tariff treatment for commonwealth Caribbean countries	CARIBCAN	May 12, 1986	Waiver–Art. I:1
ACP—EEC Fourth Convention of Lomé	Fourth Convention of Lomé	Sept. 1, 1991	Waiver–Art. I:1
Andean Trade Preference Act	U.S.–Andean Trade Preference Act	Dec. 4, 1991	Waiver–Art. I:1
Trade agreement between the governments of the Federation of Rhodesia and Nyasaland and the Union of South Africa	Federation of Rhodesia/Nyasaland–South Africa Agreement of 1955	July 1, 1955	Waiver–Art. I:2
	Federation of Rhodesia/Nyasaland–South Africa Agreement of 1960	July 1, 1960	Waiver–Art. I:2

Enabling Clause. Here a full FTA or CU as defined in Article XXIV is not required, and partial preferences are allowed. Seventeen arrangements fall under this category. Finally, within the Generalized System of Preferences (GSP), a waiver from the MFN Article I may be given for preferences granted by developed countries to developing countries. Nine agreements fall under this category.

Notes

We thank Jeffrey Frankel, Philip Levy, T. N. Srinivasan, Robert Staiger, and participants in the CIE-AEI Conference in June 1995, as well as Pravin Krishna, for many helpful conversations and comments on an earlier draft. We have also benefited from suggestions made at seminars at Harvard and Osaka universities, the University of Maryland, and the Stockholm School of Economics. Special appreciation is expressed to Maria Pillinini of the Development Division of the World Trade Organization for providing the list of PTAs at the end of this chapter in appendix table 1A-1.

1. The focus of our chapter will be on Article XXIV-sanctioned PTAs, rather than on every kind of preferential arrangement among any subset of World Trade Organization (WTO) members. PTAs, often grouped together into a single category, actually fall into three different WTO categories: Article XXIV arrangements involving FTAs and CUs, Enabling Clause arrangements limited to developing countries and permitting partial preferences, and Generalized System of Preferences (GSP) arrangements permitted via a grant of an exception to Article I. Appendix table 1A-1 provides a complete list of PTAs reported to WTO to date according to the WTO category within which they fall.

2. The reasons why these did not succeed are discussed in Bhagwati (1991).

3. These different approaches, and the later approaches to the static theory of preferential trading areas by Kemp and Wan (1976) and Brecher and Bhagwati (1981), have been distinguished and discussed in the graduate textbook by Bhagwati and Srinivasan (1983). The Cooper-Massell-Johnson-Bhagwati argument has also been formalized recently, using the Kemp-Wan approach and combining it with the theory of noneconomic objectives, by Krishna and Bhagwati (1994).

4. Our analytical synthesis draws on Bhagwati, Krishna, and Panagariya (1996) and also on our paper for the 1996 American Economic Association meetings in San Francisco, Bhagwati and Panagariya (1996).

5. In Viner's analysis, reproduced in figure 2.1, with constant costs everywhere, the concepts translate immediately into a shift of imports from the outside to the partner country as trade diversion and a shift from the home country production to imports from the partner country as trade creation. This translation does not hold fully in figure 2.3, for example.

6. In a generous introductory footnote to his article entitled "Emerging Regional Arrangements: Building Blocks or Stumbling Blocks?" Lawrence (1991) writes, "I owe this phrase to Jagdish Bhagwati." Bhagwati (1991, 77) refers to the expansion of membership as a test of PTAs serving as "building blocks" for worldwide freeing of trade: this concept is illustrated in figure 2.9, reproduced from Bhagwati (1993a). Evidently, if going down the PTA path can trigger multilateral negotiations and their successful conclusion, that too can be a way in which PTAs may serve as building blocks, as discussed here.

7. The "natural trading partners" hypothesis comes therefore in two forms. In the first form, the emphasis is on a large initial volume of trade that may result, inter alia, from geographical proximity. In the second form, the emphasis is on transport costs that are assumed to be low between countries within the same region. We have been firmly informed by Paul Wonnacott that the term "natural trading partner" originated in Wonnacott and Lutz (1989). Many authors have attributed the term instead to Krugman (1991a), who, along with Summers, should nevertheless be credited with popularizing it.

8. We assume a specific rather than an ad valorem tariff for geometric simplicity when supply curves are rising. Nothing in the analysis hangs on it.

9. Because imports expand, some of the inefficient domestic production is replaced by imports from B. A also gains from an increase in the consumers' surplus in excess of the tariff revenue.

10. B gains nothing and C loses nothing, given the constant-cost assumptions on their supply curves in trade.

11. The measure used is the conventional Hicksian equivalent variation: keeping the initial nondiscriminatory tariff, how much income can A withdraw to yield the same welfare loss as the FTA imposes?

12. Many of the points in this and the following section have been made earlier in Panagariya (1995a, 1995b). The tariff-revenue-transfer effect central to our analysis is normally present in all models characterized by flexible terms of trade. Thus, see the three-good, three-country general-equilibrium analyses of Berglas (1979) and Riezman (1979), which are neatly summarized within a unified framework by Lloyd (1982). Both Berglas and Riezman find, as we do, that when intra-union terms of trade are flexible, a large volume of imports from the partner country is inversely related to the welfare effect of a preferential liberalization. Neither of these authors makes many of the points we make or looks at the problem as we do, however.

13. There is no trade creation in the example as the FTA leaves the domestic price and therefore total imports into A unchanged.

14. We discuss the natural trading partners hypothesis in the alternative context of transport costs later in the chapter.

15. In a similar vein, Krugman (1991a, 21) notes, "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."

16. Ceteris paribus, the less A trades with the outside country, the less tariff revenue it collects and the less is its gain. Thus, in the spirit of our previous discussion, a high proportion of trade with the partner implies smaller gains from preferential liberalization.

17. In addition, a fraction of the large imports from the United States could well be a result of preferential policies rather than competitiveness.

18. It is a common practice in the computable general-equilibrium (CGE) models to differentiate goods by the country of origin and yet impose the small-country assumption. To a general equilibrium theorist, this is not correct. If a country is the sole producer of its exports, it necessarily has market power.

19. Rules of origin can and do, of course, restrict trade in other ways. For a recent analysis of how rules of origin can lead to welfare-worsening outcomes, see Krueger (1993, 1995).

20. See Grossman and Helpman (1995) in this context.

21. Note that the horizontal difference between E_BE_B and S_BS_B declines as price rises. This is because the demand in B must fall with a rise in the price.

22. This simple point seems to have escaped a number of CGE-modelers of NAFTA who distinguish products by the country of origin and continue to impose the small-country assumption.

23. The same would also hold true if we were to use a monopolistic-competition or oligopoly model.

24. In a small, open economy with tariffs as the only distortion, the change in welfare (real income) from an infinitesimally small change in any set of tariffs equals the change in tariff revenue evaluated at initial tariff rates (Eaton and Panagariya 1979). For an infinitesimally small change in the tariff on B, the vertically shaded area in figure 2.5 is the increase and the horizontally shaded area the decrease in tariff revenue measured at the original tariff rates.

25. Observe that the world price of each product is unity. Therefore, the base of the rectangle represents both the quantity and value of imports at world prices.

26. To make this point another way, start with a zero tariff on good b and a positive tariff on c. The introduction of a small tariff on b will not lead to an efficiency loss in the b market but will generate an efficiency cost in the c market.

27. The effects shown in figure 2.5 do not arise in the partial equilibrium model of figures 2.2 and 2.3. Because these effects require the presence of at least two importables, they do not arise even in a two-good general equilibrium model.

28. Recall that in figures 2.1 and 2.2, the *internal* terms of trade are variable. Country A's terms of trade with respect to country B deteriorate by the full amount of the tariff reduction. But because of the small-union assumption, the external terms of trade do not change there.

29. One hundred percent revenue seeking means that the entire revenue is available for those who wish to seek it. Perfectly competitive revenue seeking leads to a dollar's worth of resource loss for a dollar of revenue sought. The two assumptions together imply that the resources used up in revenue seeking equal the tariff revenues in equilibrium. For rent seeking, see Krueger (1974) and for revenue seeking, see Bhagwati and Srinivasan (1980).

30. A detailed, general equilibrium analysis of this issue is provided in Panagariya (1996a, 1996b).

31. We assumed earlier that each country is the sole producer of its export good. This assumption necessarily makes the terms of trade variable. In the conventional analysis, as also in the present discussion, the outside country is assumed to produce all goods and is large. The terms of trade are then determined in the outside country, and the only effects that arise are those depicted in figure 2.5. In arriving at the conclusions discussed in this paragraph, Lipsey also assumed that preferences are Cobb-Douglas. For further details, see Panagariya (1996a, 1996b).

32. As quoted in footnote 12 of Bhagwati (1993a), according to Lipsey, "the larger are purchases of domestic commodities and the smaller are purchases from the outside world, the more likely is it that the union will bring gain." If the liberalizing country's preferences are of the CES variety, the compensated crossprice elasticity of its demand for the partner's good with respect to the price of its own good reduces to the product of the expenditure share of its own good and the elasticity of substitution. A similar statement applies to the compensated crossprice elasticity of the country's demand for the partner's good with respect to the

price of the outside country's good. Thus, under CES preferences, our condition in the text reduces to Lipsey's. As noted in the previous footnote, Lipsey himself had relied on Cobb-Douglas preferences to derive the conclusion quoted at the beginning of this footnote.

33. This would not be generally true even if we were to take the measure just for one individual country with every other country instead of pooling all possible pairs together.

34. Thus, intra-African exports were only 12.8 percent of total African exports in 1990.

35. There are countless other examples. Bhagwati (1993a) cites India-Pakistan versus India-United Kingdom and India-USSR as an example.

36. Although Frankel and Wei find that a common border increases trade volumes, Dhar and Panagariya (1994), who estimate the gravity equation on a country-by-country basis for twenty-two countries, find the common-border effect to be negative in six cases. This conflict of results underlies the serious reservations we have about the use of these gravity models to infer "trade diversion," and so on: the coefficients vary considerably depending on the dataset, and sometimes the signs do as well.

37. We note, however, that the recent critique of gravity models by Jacques J. Polak (1996) casts serious doubt on even this conclusion. Polak estimates a gravity equation for total imports as a function of income, population, and a location index measuring how favorably a country is located for purposes of international trade. He finds that, for 1960 trade data, the location index yields a statistically significant effect, as in Frankel-Wei regressions. But for the 1990 sample used by Frankel and Wei, the effect is statistically insignificant.

38. Of course, even if the relationship was "natural," it does not justify preferences as argued already by us.

39. Bhagwati, in several writings, for example, Bhagwati and Dehejia (1994) and Bhagwati (1996a), has argued that comparative advantage has become "kaleidoscopic," that is, thin and volatile, as technical know-how has converged, multinationals have become global, interest rates have become closer across nations, and access to different capital markets has become more open. More and more industries are thus footloose.

40. In this paragraph, we abstract from the demand effects. The inclusion of demand effects will modify the discussion but not the fundamental point.

41. And if costs indeed fall below those of Taiwan, there is no extra gain from the PTA since in that case Canada would have replaced Taiwan as the supplier even under a non-discriminatory tariff.

42. This makes the analysis complicated because the countries now wield market power, and unilateral free trade is no longer optimal.

43. The point can also be made under "iceberg" type transport costs that are frequently employed in international trade literature. In this formulation, a constant fraction of the good melts away in transit.

44. If the initial tariff is above the optimum tariff, given the elasticity assumption, a small preferential reduction in the tariff on B is welfare improving. For a complete removal of the tariff on B to be welfare improving, the initial tariff must be substantially higher than the optimum tariff.

45. De Melo, Panagariya, and Rodrik (1993) note a similar possibility when the country faces a revenue constraint.

46. We discuss these noneconomic factors later in the chapter. Our analysis, which has focused mainly on the effects on the member countries, has not addressed adequately the issue of the effects on nonmembers. However, there is a revival of interest in that issue as well. See, in particular, Srinivasan (1995) and Winters (1995a, 1995b).

47. The common-water definition, of course, excludes land-locked countries such as Nepal and countries with shores only on land-locked seas such as the Caspian. These, however, add up to only a small fraction of world trade. See Bhagwati (1996b) for more on common-waters FTAs.

48. Also see Frankel, Stein, and Wei (1995a, 1995b). Interestingly, Haberler (1943) appears to have been the first to use the term *continental blocs.*

49. As matters stand currently, however, APEC and TAFTA are extremely unlikely to become Article XXIV-sanctioned PTAs, despite the U.S. obsession with PTAs, whereas the extension of NAFTA to the South looks like a long-term process.

50. Similarly, the MTN path may facilitate or obstruct the expansion of PTA membership, so that the interaction between the two paths may be mutual.

51. Christopher Bliss (1994) has tried to give the argument some structure. More recently, T. N. Srinivasan (1995) has done so in the context of examining the question of the impact of PTAs on nonmember welfare.

52. See Deardorff and Stern (1994).

53. Bhagwati (1993a, 40–44) also discussed skeptically the claims that PTA formation is quicker, more efficient, and more certain than MTN.

54. For an early analysis of the political factors underlying the formation of PTAs, see the work by the political scientist Edward Mansfield (1992) cited and discussed in Bhagwati (1993a). Other political scientists, such as Miles Kahler and Joseph Grieco, have written in this area recently

55. In comparing incentives for export-oriented firms, for lobbying for a PTA (for example, NAFTA) as against MTN (for example, the Uruguay Round), a dollar's worth of lobbying would go a longer way in the former case because any preferential opening of the Mexican market would be better for the U.S. exporter than such an opening on an MFN basis that yields the benefits equally to U.S. rivals in Japan, the EU, and elsewhere. This argument applies only to the extent that the MTN process simultaneously does not open other markets to the U.S. exporter on a reciprocal basis.

56. Bhagwati (1993a) cites Irwin's (1993) study of trade liberalization in the nineteenth century, which shows that the Anglo-French Treaty may well have served this purpose. Richard Baldwin's (1993) subsequent formalization of this basic idea in what he calls the "domino" theory of PTA expansion is discussed below.

57. In this review, we do not include the important contributions to the political economy-theoretic analysis of PTAs that do not directly address either of the two dynamic time-path questions at issue in the text. For example, Grossman and Helpman (1995) have modeled the formation of PTAs, demonstrating the critical role played by the possibility of trade diversion in the outcome, a conclusion also arrived at independently by Pravin Krishna (1993) in a different model. Similarly, Panagariya and Findlay (1996) have formalized the endogeneity argument that reduced protection between members in a PTA can lead to increased protection against nonmembers. Using a political process consisting of lobbying by owners of specific factors, they also investigate the external tariffs that emerge under an FTA and a CU.

For answers to a similar set of questions, but under the assumption of a welfare-maximizing government, see our discussion of Bagwell and Staiger (1993).

58. As noted in Bhagwati (1993a, 1994), this is exactly what the United States achieved, in terms of intellectual property protection and even concessions on environmental and labor standards enforcement, by getting then president Carlos Salinas into a one-on-one bargaining situation in NAFTA. And now Chile is poised to accept these obligations as the price of getting into NAFTA. On the other hand, as the virtually unanimous developing country objections to labor standards demands at the WTO show, neither Mexico nor Chile would have agreed to these demands in the purely WTO context.

59. This was stated to be the case for Washington, D.C., by a well-placed trade economist in the Clinton administration, at a recent conference on the subject of PTAs. The first author, at a Stockholm conference on WTO issues in 1996, organized by the Swedish trade minister, Mats Hellström, found a similar unawareness among some of the trade ministers and bureaucrats present, even as the response of the attending economists to his critical remarks about the current obsession with PTAs was enthusiastic.

60. And then there are also those who think that the APEC, turned into a PTA that excludes the extension of trade barrier reductions to Europe, will prompt Europe to its own tariff cuts in a benign outcome. This viewpoint, ascribed in the media to C. F. Bergsten, is premised on his view that the Seattle APEC summit pushed the Europeans into settling the Uruguay Round. The latter view is unpersuasive since, in the end, it was the U.S. administration that decided to accept the advice of many to close the Round with whatever it could get and to proceed to build on that in future negotiations. For a critique of similar, special pleading to justify NAFTA, see Bhagwati (1995, 11–12).

61. For a detailed statement of this critique, see Bhagwati (1995).

62. Unfortunately, this comment also applies to many of the numerical models, including the computable general-equilibrium models, estimating the gains from PTAs, as discussed by Panagariya in a forthcoming essay. And then there are the more elementary conceptual errors that afflict the numerical estimates of gains in *employment* from NAFTA. These errors were widely repeated by the media at the time.

References

Bagwell, Kyle, and Robert Staiger. "Multilateral Cooperation During the Formation of Free Trade Areas." NBER Working Paper no. 4364, 1993.

Baldwin, Richard. "A Domino Theory of Regionalism." CEPR Working Paper no. 857, November 1993.

Berglas, Eitan. "Preferential Trading: The n Commodity Case." *Journal of Political Economy*, vol. 87, 1979, pp. 315–31.

Bhagwati, Jagdish. "Trade Liberalization Among LDCs, Trade Theory and GATT Rules." In J. N. Wolf, ed., *Value, Capital, and Growth: Papers in Honour of J. R. Hicks*. Oxford: Oxford University Press, 1968.

———. "The Theory of Political Economy, Economic Policy, and Foreign Investment." In M. Scott and D. Lal, eds., *Public Policy and Economic Development*, Essays in Honor of I.M.D. Little, pp. 217–30. Oxford: Clarendon Press, 1990.

———. *The World Trading System at Risk.* Princeton, N.J.: Princeton University and Harvester Wheatsheaf, 1991.

———. "Regionalism and Multilateralism: An Overview." In Jaime de Melo and Arvind Panagariya, eds., *New Dimensions in Regional Integration.* Cambridge: Cambridge University Press, 1993a.

———. "Beyond NAFTA: Clinton's Trading Choices." *Foreign Policy,* Summer 1993b, pp. 155–62.

———. "Threats to the World Trading System: Income Distribution and the Selfish Hegemon." *Journal of International Affairs,* vol. 48, Spring 1994, pp. 279–85.

———. "U.S. Trade Policy: The Infatuation with Free Trade Areas." In Jagdish Bhagwati and Anne O. Krueger, *The Dangerous Drift to Preferential Trade Agreements.* Washington, D.C.: AEI Press, 1995.

———. "Trade and Wages: A Malign Relationship?" In Susan Collins, ed., *The American Worker: Exports, Imports and Jobs.* Washington, D.C.: Brookings Institution, forthcoming (1996a).

———. "Watering of Trade." *Journal of International Economics,* forthcoming (1996b).

Bhagwati, Jagdish, and Vivek Dehejia. "Trade and Wages: Is Marx Striking Again?" In Jagdish Bhagwati and M. Kosters, eds., *Trade and Wages: Leveling Wages Down?* Washington, D.C.: AEI Press, 1994.

Bhagwati, Jagdish, Pravin Krishna, and Arvind Panagariya. "Introduction." In Jagdish Bhagwati and Pravin Krishna, eds., *Contributions to the Theory of Preferential Trading Areas.* Cambridge, Mass.: MIT Press, forthcoming.

Bhagwati, Jagdish, and Arvind Panagariya. "The Theory of Preferential Trade Agreements: Historical Evolution and Current Trends." *American Economic Review,* vol. 86, pp. 82–87.

Bhagwati, Jagdish, and T. N. Srinivasan. "Revenue Seeking: A Generalization of the Theory of Tariffs." *Journal of Political Economy,* vol. 88, December 1980, pp. 1069–87.

———. *Lectures in Trade Theory.* Cambridge, Mass.: MIT Press, 1983. Bliss, Christopher. *Economic Theory and Policy for Trading Blocs.* Manchester and New York: Manchester University Press, 1994.

Brecher, Richard, and Jagdish Bhagwati. "Foreign Ownership and the Theory of Trade and Welfare." *Journal of Political Economy,* vol. 89, June 1981, pp. 497–511.

Cooper, C. A., and B. F. Massell. "A New Look at Customs Union Theory." *Economic Journal,* vol. 75, 1965a, pp. 742–47.

———. "Towards a General Theory of Customs Unions for Developing Countries." *Journal of Political Economy,* vol. 73, 1965b, pp. 461–76.

Deardorff, Alan V., and Robert M. Stern. "Multilateral Trade Negotiations and Preferential Trading Arrangements." In Alan V. Deardorff and Robert M. Stern, eds., *Analytical and Negotiating Issues in the Global Trading System.* Ann Arbor: University of Michigan Press, 1994.

de Melo, Jaime, and Arvind Panagariya, ed. *New Dimensions in Regional Integration.* Cambridge, Great Britain: Cambridge University Press, 1993.

de Melo, Jaime, Arvind Panagariya, and Dani Rodrik. "The New Regionalism: A Country Perspective." In Jaime de Melo and Arvind Panagariya, eds., *New Dimensions in Regional Integration*, chap. 6. Cambridge, Great Britain: Cambridge University Press, 1993.

Dhar, Sumana, and Arvind Panagariya. "Is East Asia Less Open than North America and the European Economic Community? No." Policy Research Working Paper no. 1370. Washington, D.C.: World Bank, 1994.

Eaton, Jonathan, and Arvind Panagariya. "Gains from Trade under Variable Returns to Scale, Commodity Taxation, Tariffs and Factor Market Distortions." *Journal of International Economics*, vol. 9, 1979, pp. 481–501.

Findlay, Ronald, and Stanislaw Wellisz. "Endogenous Tariffs, the Political Economy of Trade Restrictions and Welfare." In Jagdish Bhagwati, ed., *Import Competition and Response*. Chicago and London: University of Chicago Press, 1982.

Frankel, Jeffrey, and Sheng-Jin Wei. "The New Regionalism and Asia: Impact and Options." Paper presented at the Asian Development Bank Conference on Emerging Global Trading Environment and Developing Asia, May 29–30, 1995.

Frankel, Jeffrey, E. Stein, and Shing-Jin Wei. "Trading Blocs and the Americas: The Natural, the Unnatural and the Supernatural." *Journal of Development Economics*, vol. 47, June 1995a, pp. 61–96.

———. "Continental Trading Blocs: Are They Natural or Super-Natural?" NBER Working Paper no. 4588, 1995b.

Grossman, Gene, and Elhanan Helpman. "The Politics of Free Trade Agreements." *American Economic Review*, September 1995, pp. 667–90.

Haberler, Gottfried. "The Political Economy of Regional or Continental Blocs." In Seymour. E. Harris, ed., *Postwar Economic Problems*. New York. 1943.

Irwin, Douglas. "Multilateral and Bilateral Trade Policies in the World Trading System: An Historical Perspective." In Jaime de Melo and Arvind Panagariya, eds., *New Dimensions in Regional Integration*. Cambridge, Great Britain: Cambridge University Press, 1993.

Johnson, Harry. *Money, Trade and Economic Growth*. Cambridge, Mass.: Harvard University Press, 1962.

———. "An Economic Theory of Protectionism, Tariff Bargaining, and the Formation of Customs Unions." *Journal of Political Economy*, vol. 73, June 1965, pp. 256–83.

Kemp, Murray C., and Henry Wan. "An Elementary Proposition Concerning the Formation of Customs Unions." *Journal of International Economics*, vol. 6, February 1976, pp. 95–98.

Krishna, Pravin. "Regionalism and Multilateralism: A Political Economy Approach." Economics Department, Columbia University, mimeo, December 1993. Presented to the NBER Universities Research Conference on International Trade and Regulations, Cambridge, Mass., 1993.

Krishna, Pravin, and Jagdish Bhagwati. "Necessarily Welfare-Enhancing Customs Unions with Industrialization Constraints: A Proof of the Cooper-Massell-Johnson-Bhagwati Conjecture." Columbia University Working Papers, April 1994.

Krueger, Anne O. "The Political-Economy of the Rent Seeking Society." *American Economic Review*, vol. 69, 1974, pp. 291–303.

———. "Rules of Origin as Protectionist Devices." NBER Working Paper no. 4352. April 1993. [Forthcoming] in J. Melvin, J. Moore, and R. Riezman, eds. *International Trade Theory: Essays in Honour of John Chipman*. United Kingdom: Routledge, 1996.

———. "Free Trade Agreements versus Customs Unions." NBER Working Paper no. 5084, April 1995.

Krugman, Paul. "The Move to Free Trade Zones." In Symposium sponsored by the Federal Reserve Bank of Kansas City, *Policy Implications of Trade and Currency Zones*, 1991a.

———. "Is Bilateralism Bad?" In E. Helpman and A. Razin, eds., *International Trade and Trade Policy*. Cambridge, Mass.: MIT Press, 1991b.

———. "Regionalism versus Multilateralism: Analytical Notes." In Jaime de Melo and Arvind Panagariya, eds., *New Dimensions in Regional Integration*. Cambridge, Great Britain: Cambridge University Press, 1993.

Lawrence, Robert Z. "Emerging Regional Arrangements: Building Blocks or Stumbling Blocks?" In Richard O'Brien, ed., *Finance and the International Economy*, Amex Bank Prize Essays. Oxford: Oxford University Press for the Amex Bank Review, 1991.

Levy, Philip. "A Political Economic Analysis of Free Trade Agreements." *American Economic Review*, forthcoming.

Lipsey, Richard. "The Theory of Customs Unions: Trade Diversion and Welfare." *Economica*, vol. 24, 1957, pp. 40–46.

———. *The Theory of Customs Unions: A General Equilibrium Analysis*. University of London, Ph.D. Thesis, 1958.

———. "The Theory of Customs Unions: A General Survey." *Economic Journal*, vol. 70, 1960, pp. 498–513.

Lloyd, Peter J. "3 × 3 Theory of Customs Unions." *Journal of International Economics*, vol. 12, 1982, pp. 41–63.

Ludema, Rodney. "On the Value of Preferential Trade Agreements in Multilateral Negotiations." Mimeo. 1993.

Mansfield, Edward. "The Concentration of Capabilities and International Trade." *International Organization*.

Mayer, Wolfgang. "Endogenous Tariff Formation." *American Economic Review*, vol. 74, December 1984, pp. 970–85.

Meade, James E. *The Theory of Customs Unions*. Amsterdam: North-Holland, 1955.

Mundell, Robert A. "Tariff Preferences and the Terms of Trade." *Manchester School of Economic and Social Studies*, vol. 32, 1964, pp. 1–13.

Panagariya, Arvind. "Should East Asia Go Regional? No, No and Maybe." WPS 1209. Washington, D.C.: World Bank, 1993.

———. "East Asia and the New Regionalism." *World Economy*, vol. 17, November 1994, pp. 817–39.

———. "Rethinking the New Regionalism." Paper presented at the Trade Expansion Program Conference of the United Nations Development Programme and World Bank, January 1995a.

———. "The Free Trade Area of the Americas: Good for Latin America?" Working Paper no. 12, *World Economy*, forthcoming. Center for International Economics, University of Maryland, 1995b.

———. "Preferential Trading and the Myth of Natural Trading Partners." Working Paper no. 200, Center for Japan-U.S. Business and Economic Studies, Stern School of Business, New York University, New York, 1996a.

———. "The Meade Model of Preferential Trading: History, Analytics and Policy Implications." Mimeo, University of Maryland, Department of Economics, 1996b.

Panagariya, Arvind, and Ronald Findlay. "A Political Economy Analysis of Free Trade Areas and customs Unions." In Robert Feenstra, Gene Grossman, and Douglas Irwin, eds., *The Political Economy of Trade Reform*, Essays in Honor of Jagdish Bhagwati. Cambridge, Mass.: MIT Press, 1996.

Polak, Jacques J. "Is APEC a Natural Regional Trading Bloc? A critique of the 'Gravity Model' of International Trade." Mimeo (revised version to appear in the *World Economy*).

Richardson, M. "Why a Free Trade Area? The Tariff Also Rises." *Economics & Politics*, vol. 6, March 1994, pp. 79–95.

Riezman, Raymond. "A 3 × 3 Model of Customs Unions." *Journal of International Economics*, vol. 9, 1979, pp. 341–54.

Srinivasan, T. N. "Discussion." In Jaime de Melo and Arvind Panagariya, eds., *New Dimensions in Regional Integration*. Cambridge, Great Britain: Cambridge University Press, 1993.

———. "Common External Tariffs of a Customs Union: The Case of Identical Cobb-Douglas Tastes." Mimeo, Yale University, 1995.

Summers, Lawrence. "Regionalism and the World Trading System." Symposium Sponsored by the Federal Reserve Bank of Kansas City, *Policy Implications of Trade and Currency Zones*, 1991.

Viner, Jacob. *The Customs Union Issue*. New York: Carnegie Endowment for International Peace, 1950.

Wei, Sheng-Jin, and Jeffrey Frankel. "Open Regionalism in a World of Continental Trade Blocs." Revised version of NBER Working Paper no. 5272, November 1995.

Winters, L. A. "Regionalism and the Rest of the World: The Irrelevance of the Kemp-Wan Theorem." Mimeo. Washington, D.C.: World Bank, 1995a.

———. "European Integration and Economic Welfare in the Rest of the World." Mimeo. Washington, D.C.: World Bank, 1995b.

Wolf, Martin. "Comments." In Jeffrey J. Schott, ed., *Free Trade Areas and U.S. Trade Policy*, pp. 89–95. Washington, D.C.: Institute for International Economics, 1989.

Wonnacott, Paul, and Mark Lutz. "Is There a Case for Free Trade Areas?" In Schott, Jeffrey J., ed., *Free Trade Areas and U.S. Trade Policy*, pp. 59–84. Washington, D.C.: Institute for International Economics, 1989.