CHAPTER 3. MAKING THE MOST OF REGIONAL INTEGRATION

One important message that emerges from the analysis of chapter 2 is that in purely trading terms, an RIA cannot provide benefits that are not open to its member countries through non-discriminatory tariff reductions. In terms of reallocating factors of production (assuming that their quantity and productivity are unaffected by regional integration) or of changing patterns of consumption, there appears to be no economic rationale for forming RIAs. Non-discriminatory tariff reductions would yield a country all the gains of trade creation without the costs of trade diversion, and would, thus, be a superior option. Once we take the transfers between government revenue and producers into account, one member might do better from an RIA than under non-discrimination, but only by driving the other even further below that level. Thus the arguments of Chapter 2 still can not explain why both members would agree to an RIA.

Then why are RIAs so popular? Several explanations are offered in later chapters—e.g. that regionalism helps to reduce the chances of conflict with neighbors, or helps in negotiating agreements to share regional resources—but here we consider whether extending the analysis of purely trade or border factors could explain this popularity. We find that trade-based factors offer little justification for preferring RIAs to free trade, although we do find reasons why if RIAs grow up for other reasons they tend to be between neighbors, and, more controversially why developing countries will generally gain more from RIAs with developed countries than with other developing countries.

The judgements behind these conclusions are necessarily empirical and pragmatic. Although economic theory suggests that RIAs can always be designed to be welfare improving, in practice the necessary information, the incentives and the legal rights to manipulate the instruments of trade policy to guarantee such outcomes are missing. Ultimately it is not theoretical possibilities or empirical generalities that define the appropriateness of an RIA, but practical issues such as which partners are actually chosen and the level of tariffs against the rest of the world. These are the subject of this chapter.

3.1. Free trade is Fine: but with Whom?

One of the rare questions on which trade bloc economists agree is the common sense proposition that an RIA between any set of countries could be designed to leave the rest of the world indifferent, at least one of the members better, and the rest no worse off. (Kemp and Wan, 1976; Krishna and Panagariya, 1997). Unfortunately, however, this result requires setting external tariffs so that members’ total trade with the rest of the world (commodity by commodity) is maintained at pre-union levels, and making lump-sum taxes and transfers between partners. The former are almost impossible to calculate—and might indeed conflict with WTO obligations—while the latter are never feasible in practice.

---

1 Srinivasan (1997) offers one such attempt, but under very restrictive circumstances.
Given these practical constraints on ensuring that any given RIA is welfare improving, and the fact that forming an RIA with one set of partners precludes or at least complicates creating one with others, the question of why an RIA might be attractive can not be divorced from the question of who belongs to it. Thus this section considers both rather generic arguments in favor of discrimination along with rather specific ones about the particular attractions of RIAs with certain types of partner.

### 3.1.1 Why not with everyone?

One argument for preferring an RIA—specifically a customs union— to free trade is that if a trade bloc’s joint economy is large, coordinating its international trade policies will allow it to strike better bargains with trading partners. That is, it can reap terms of trade benefits not available to its members acting independently. This is examined in Chapter 8, where we consider the effects of RIAs on the rest of the world, for such gains are necessarily at the latter’s expense. There is actually very little research on such gains, but Chang and Winters (1999) suggest that they may have been important for Mercosur. In general, however, improving the terms of trade is not commonly believed to be a big motivator of trade policy stances, and besides most RIAs—especially among developing countries—are pretty small.

A second argument concerns market access. An RIA not only affects imports, but also promises the benefits of duty-free access for exports to partner markets. This is an important consideration for a country considering joining an RIA, but the benefits it derives from such market access can only come at the expense of its partner, which, for its part, would be better off if it were to eliminate its tariff. The benefits are the exact counterpart of the tariff revenue losses discussed earlier in chapter 2. Both parties cannot gain simultaneously from such transfers.

Another aspect of the same argument is that countries may feel better able to liberalize their own trade if others are doing so at the same time. If improved access to another market depends on opening one’s own, a new constituency for trade reform may be created among exporters. If countries simultaneously reduce tariffs on each others’ exports, adverse terms of trade changes are less likely than if only one party liberalizes (Bagwell and Staiger, 1997) — and if several liberalize together, the need for any one of them to devalue its currency to maintain its foreign balance is reduced. Also, adjustment might be easier if export sectors are expanding as import-competing sectors contract. These are all reasons for coordinating trade reforms across countries, and RIAs provide one way of doing this. But concerted unilateralism—under which several countries introduce MFN liberalization together —would also fit the bill, as several member countries argued in the early days of APEC.

A third possibility favoring RIAs over MFN free trade arises if member countries wish to industrialize behind protective barriers. Forming a group that provides scope for intra-bloc industrial specialization reduces the cost of protection and can generate welfare gains not open to members by unilateral tariff liberalization (Cooper and Massell, 1965).

---

2 Recall that free trade agreements (FTAs) have free trade internally while customs unions (CUs) both have internal free trade and a common external tariff (trade policy) against non-members. CUs and FTAs are more fully discussed in Section 3.3.
The first wave of RIAs among developing countries in the 1960s and 1970s set such import substitution as a major policy objective, and a number of RIAs (including LAFTA and ASEAN) appear to have been explicitly motivated by gains from pursuing industrialization on a regional basis. But these policies must be based on non-economic governmental objectives—or public goods arguments—given that there are few grounds for concluding that economic gains can be derived from protected industrial development.

If industries can be rendered competitive at higher levels of output when preferential markets are available, it might appear that an RIA could be justified. But it has then to be explained why the industry would need the support of a protected regional market if mere expansion of output (perhaps through scale economies) suffices to make its costs competitive. The issues surrounding the dynamics of industrialization and long-term comparative advantage are addressed further in Chapters 4 and 5.

3.1.2 Choosing Partners – The Natural Partners Fallacy

Ever since customs theory began with Viner (1950), economists have puzzled about who the perfect partners would be. For many years, this proved to be a quite sterile debate, every proposal turning out to be at best a restricted and special case. Recently, however, World Bank research has cast new light on this question, which if it does not close the book on the controversy at least shines a brighter light on it.

Early research focused on whether potential RIA members were complementary in their production and consumption patterns and on whether they were already major trading partners already. The former criterion proved very difficult to pin down, but in the early 1990s the latter appeared to gain considerable currency, along with the notion that countries close enough to each other to reap savings on transport costs were natural partners - Wonnacott and Lutz (1989) and Summers (1991).

Before assessing the validity of the argument that countries that already trade disproportionately with each other are ideal partners for a RIA, we examine whether this was the case for most RIAs. The left side of Table 3.1 shows, at the time of bloc formation, the share of intra-bloc trade in member countries' total trade, for a number of RIAs. The right side of the table shows the share of an individual country's total trade that is carried out with a bloc it is joining. It is clear from the left side of the table that the intra-bloc trade share is much larger in North-North (EEC at 38.6%, CUSFTA at 30.2%) and North-South (NAFTA at 42.1%) RIAs than in South-South RIAs. The largest intra-bloc share in South-South RIAs is in ASEAN (16.7%), followed by MERCOSUR (12.9%), EAC (12.8%) and CEFTA (10.3%). The other RIAs have much smaller intra-bloc trade shares. It is below 1% for three Sub-Saharan Africa RIAs (UDEAC, SADC, CEAO), 1% for the CACM, 2.3% for the Andean Pact, 3.3% for the G-3, 3.2% for SAARC and 4.2% for the GCC.

This evidence leads us to conclude that RIAs are typically formed between neighboring countries but not necessarily between countries that are already major

---

3 Panagariya (1997) lists the most important contributions to this debate.
trading partners. As for the evidence on the right side of Table 3.1, it shows that, except for Costa Rica and the CACM, the countries that signed FTAs with the EU or the EEA were major trading partners with the EU. This can be explained, from the viewpoint of the acceding countries, by the fact that as more countries joined the EU, staying on the outside became ever more costly and the incentive to join increased. This phenomenon of “domino” regionalism is examined in more detail in Chapter 8.

Table 3.1 - Trade Shares

<table>
<thead>
<tr>
<th>Year Formed</th>
<th>Share of Intra-Bloc Trade in Total Trade (%)</th>
<th>Year Joined</th>
<th>Share of Country’s Trade with EU (or CACM) in Country’s Total Trade (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962(^1)</td>
<td>EEC 38.6</td>
<td>1994</td>
<td>Iceland-EU 60.0</td>
</tr>
<tr>
<td>1989</td>
<td>CUSFTA 30.2</td>
<td>1994</td>
<td>Norway-EU 61.3</td>
</tr>
<tr>
<td>1994</td>
<td>NAFTA 42.1</td>
<td>1996</td>
<td>Morocco-EU 74.4</td>
</tr>
<tr>
<td>1969</td>
<td>Andean Pact 2.3</td>
<td>1995</td>
<td>Tunisia-EU 56.8</td>
</tr>
<tr>
<td>1962</td>
<td>CACM 1.0</td>
<td>1994</td>
<td>Hungary-EU 41.5</td>
</tr>
<tr>
<td>1991</td>
<td>MERCOSUR 12.9</td>
<td>1994</td>
<td>Poland-EU 49.4</td>
</tr>
<tr>
<td>1995</td>
<td>G-3 3.3</td>
<td>1995</td>
<td>Czech Rep.-EU 46.7</td>
</tr>
<tr>
<td>1966</td>
<td>UDEAC 0.9</td>
<td>1995</td>
<td>Lithuania-EU 23.0</td>
</tr>
<tr>
<td>1980</td>
<td>SADC 0.1</td>
<td>1995</td>
<td>Latvia-EU 21.7</td>
</tr>
<tr>
<td>1973</td>
<td>CEAO 0.9</td>
<td>1995</td>
<td>Romania-EU 41.8</td>
</tr>
<tr>
<td>1993</td>
<td>CEFTA 10.3</td>
<td>1995</td>
<td>Slovak Rep.-EU 26.4</td>
</tr>
<tr>
<td>1967</td>
<td>ASEAN 16.7</td>
<td>1995</td>
<td>Slovenia Rep.-EU 53.7</td>
</tr>
<tr>
<td>1981</td>
<td>GCC 4.2</td>
<td>1962</td>
<td>Costa Rica-CACM 2.5</td>
</tr>
<tr>
<td>1985</td>
<td>SAARC 3.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The full names of the RIAs are provided in the Annex to Chapter 1.
1. EEC was formed in 1957, but the earliest bilateral trade data obtained is for 1962.

Attractive though it would be to have a shortcut to assessing RIAs, the concept of “natural trading partners” actually has little or no operational content. To be useful, it
must be easily applied and must identify trade that we wish to stimulate artificially by means of preferences. That a trade flow is large already says nothing about the need to stimulate it: the advantages of proximity and conformable commodity composition are already reflected in its size. In fact, some flows are large because of existing distortions and need curtailment rather than boosting.

On transportation costs, the argument is more sophisticated, but ultimately of little more use practically. RIAs can be advantageous if they allow two countries to trade a good with each other that previously one exported to the rest of the world and the other imported at a higher price. This situation could arise either because the rest of the world levies tariffs or export taxes on its trade or because of transport costs - Wonnacott and Wonnacott (1981). By trading together, and so reducing the margin between the buying and selling prices, the RIA improves one or both countries’ terms of trade, but only if at least one of them ceases its trade with the rest of the world. Amjadi and Winters (1999) show that for Mercosur, while high transport costs on trade with the rest of the world offer scope for large terms of trade gains, very few goods seem likely to experience the required change in trade patterns. Similarly, with the advent of NAFTA, Mexico reduced rather than expanded the set of goods (defined at the 6-digit level of the Harmonised System (HS) trade classification) for which the USA and/or Canada were its only suppliers.

Recently Schiff (1999) has revisited the natural trading partners argument and argued that both advocates and detractors of the hypothesis have largely missed the point, because while they have analysed the target country’s relationships with its potential partners and with the rest of the world, they have failed to consider that between the latter two. This is important because it constrains the patterns of trade and prices that could actually rule in the world prior to the formation of the RIA. Once this is taken into account, Schiff confirms that the size of the pre-existing trade flow is no guide to ‘naturalness’ – the relationship is neither necessarily positive (as advocate’s claim) nor necessarily negative (as detractors’ claim). He does, however, argue that, for a given target country, larger partners are better than smaller ones, because they are more likely to satisfy the target country’s import demand without increasing prices, and less likely to cease importing the target’s export goods from the rest of the world and hence to reduce their internal price below the world price plus their tariff on third countries. He also argues that the RIA is likely to be better off if its members are complementary in the sense that, under non-discriminatory trade policies, each would import what each other exports rather than what it imports.

3.1.3 Choosing Partners – Comparative Advantage

The second advance in the area of choosing partners is Venables (2000), who considers the comparative advantage of RIA members relative to each other, and relative to the rest of the world. It turns out that this plays a major role in defining a country’s ideal partners. It also helps to determine whether RIA members’ incomes converge or diverge, a topic to which we turn more fully in chapter 5.
Start by thinking of two developing economies that both have a comparative disadvantage in manufactures relative to the rest of the world, but the disadvantage is less for one of them than the other. Kenya and Uganda can serve as examples. Their comparative disadvantage in manufactures could come from many alternative sources—technological, geographical or institutional differences—but suppose that it is because of different endowments of capital: Kenya has little capital per worker relative to the world average, but Uganda has even less. The initial position is one in which Kenya and Uganda each have some manufacturing, serving local consumers and surviving because of relatively high tariff protection.

Now suppose that these two countries form an RIA. Since Kenya has a comparative advantage in manufacturing relative to Uganda (but not relative to the rest of the world, of course), it will expand manufacturing output to supply the whole RIA (or at least move in that direction), while Uganda’s manufacturing sector will contract. These developments move Kenya’s production structure further away from its true comparative advantage, while moving Uganda’s closer, but surprisingly, Kenya gains from the reallocation, while Uganda may lose (and will certainly do less well than Kenya). The reason is that Uganda suffers trade diversion—some manufactures that were previously imported from the rest of the world are now imported at greater cost from Kenya. But for Kenya, there are gains from being able to supply manufactures in the Ugandan market, protected from competition with the rest of the world.

There is a general argument here, which is that countries with comparative advantage closer to the world average do better in an RIA than do countries with more extreme comparative advantage. Interposing the ‘intermediate’ country between the ‘extreme’ one and the rest of the world distorts the extreme country’s trade, causing it to switch import supplier. But the intermediate country does not experience this switch in supply; its trade with the ‘extreme’ country and with the rest of the world are less close substitutes, and therefore less vulnerable to trade diversion.

A further implication follows. An RIA between two poor countries will tend to cause their income levels to diverge, but an RIA between two rich ones will tend to cause convergence. The logic can be seen from figure 3.1. The vertical line measures each country’s endowment of capital per worker; countries higher up the line have a greater comparative advantage in manufacturing, and also higher initial per capita income. For two countries below the line, we see the extreme country losing, and the intermediate country gaining—Uganda and Kenya, as marked. Similarly, for countries above the line, the extreme country loses and the intermediate one gains (labeled France and Spain respectively on the figure). However, in the former case the extreme country is the poorer of the two, so we have divergence, while in the latter it is the richer, so we have convergence. The same basic forces therefore mean that regional integration between rich countries causes their incomes to converge, whereas integration between poor ones causes divergence.

The two cases analysed above were for a pair of low-income countries and a pair of higher income countries. But what about mixed RIAs—so-called North-South RIAs? As before, the mechanism behind the effects is simply the relocation of industry in
response to differences in factor endowments, and associated differences in factor prices. In general, we would expect North-South RIAs to benefit both members because in each case comparative advantage relative to the partner and to the world push in the same direction. In fact, the changes might be particularly large, and particularly beneficial, for a lower-wage economy in an RIA with an industrialised and high-wage economy. For example, both in NAFTA and in the Europe Agreements between the EU and its eastern neighbours, there are signs of significant shifts in industry, aided, often, by foreign direct investment – see Chapter 4.

Figure 3.1: Convergence and divergence of real incomes in an RIA
3.1.4 Neighborhood RIAs

One of the most striking features of international economic integration is that the vast majority of RIAs are between neighboring countries (hence the term 'regional'). The main exceptions are APEC (if it can be considered a RIA), the Group of 3 (Mexico is not neighboring either Colombia or Venezuela), various plurilateral FTAs (EU-South Africa, EU-Mexico, EU-Chile) and bilateral FTAs (e.g., Canada-Chile, Chile-Mexico and Israel-US). The transportation cost variant of the “natural trading partners” hypothesis might explain some of this, as might regionalism’s possible role in reducing the chances of conflict with neighbors, or helping them negotiate agreements to share regional resources, which we discuss later. Other possible reasons include a history of cooperation with, and better knowledge of, neighboring countries, and a desire to replace past tensions with a institutional framework that promotes cooperation.

One pure trade possibility is that neighbors together constitute a regional market for certain goods that, for reasons of taste or excessive transportation costs, are not tradable with the rest of the world. If an RIA includes all potential suppliers of such a good, it is equivalent to multilateral free trade and hence, for small countries, bound to be welfare improving. This is a more demanding criterion than just observing that members do not import this good from the rest of the world, for that might just reflect existing tariffs. It is also relatively rare: in 1993, for example, the four Mercosur countries imported only 124 types of goods—out of a total of 4,858 types—only from each other (Amjadi and Winters, 1999). These “local” markets were concentrated on agriculture, and accounted for just half of 1% of total imports. Enlarging the criterion to goods for which over 95% of imports come from other members only adds 155 types, or 2.2% of total imports. Poorer countries may well have higher proportions of their trade in such “local” markets, but even so this does not seem likely to be a major possible benefit of regionalism. Moreover, even if trade were restricted to close neighbors, that is not an argument for forming an RIA: a nondiscriminatory liberalization would have the same beneficial effect.

Another important reason for forming a customs union between neighbors is to reduce the transaction costs involved in border formalities. These are often more important hindrances to trade than customs duties, and are far more likely to be avoidable between contiguous than distant countries. Even with neighbors, however, it can be a complex process; the EU took from 1957 until the mid-1990s to get close to having “invisible borders” between even a sub-set of its members.

A third incentive for neighborhood CUs is ensuring that countries get their proper shares of tariff and other fiscal revenues when goods are in transit or are transferred. (That is, the CU contains agreements to transfer tariff revenues from the country of collection to the country of final destination. Such agreements can be contentious, however, if the partners do not trust each other fully.) These administrative considerations — particularly important for land-locked countries — played a major role in the origins of four long-standing African RIAs: UDEAC, CEAO, EAC, and SACU. A customs union can also help avoid the costs of smuggling or tax competition. Senegal has long sought to form a customs union with The Gambia to remove the incentive for its citizens to smuggle imported manufactures in from its neighbor, which then gets the customs revenues from them. Robson (1998) estimates that for many years such
smuggling generated a significant proportion of The Gambia’s revenues from import duties (estimated at 25% in 1980) — at Senegal’s expense.

3.1.5 North-South or South-South RIAs

North-South RIAs mostly involve arrangements between developing countries and the EU or, more recently, North America. Those two trading areas include a high proportion of the world’s most efficient producers of many products, operate behind relatively low tariffs for manufacturers, and are capable of supplying the bulk of the needs of the Southern economies. To see the advantages that this has for RIA formation, consider the polar case of a good for which domestic supplies and imports from partners and non-partners are perfect substitutes for each other.

The Southern partner in a North-South RIA reduces its tariffs on imports from a supplier large enough to satisfy many of its needs at little more than the prevailing international cost. If it imports only from the Northern partner after the bloc is formed, its domestic prices fall to Northern levels and it benefits from increased consumption and reduced production of high-cost domestic substitutes. In effect, in the commodities for which this is true, the South can enjoy gains from a North-South RIA much like those from unilateral liberalization on a MFN basis. In terms of figure 2A.1 it is as if $p^A$ and $p^B$ were very close together, so that losses from diversion (area H) are small.

There are two critical issues here. First how frequently and by how much the Northern partner’s costs exceed international minima. While the EU and US are generally efficient producers and are subject to international competitive pressures, there are still plenty of products for which they are not least-cost producers. And even if their cost disadvantage were only, say 5%, it would apply to a high proportion of imports and so would represent a significant loss of income relative to multilateral free trade. As early as 1977, Roemer (1977) identified the tendency for metropolitan countries to sell far broader ranges of goods in their colonies and former colonies then they were able to do on world markets. This is the direct analogue of the 5% margins.

The second issue is that one needs to be confident that prices would actually fall to the internal levels of the Northern partner. The story above implicitly assumes that the partner becomes the only supplier, but, as we saw in chapter 2 above, if imports do continue from third countries which continue to face the tariff, domestic prices will not fall and there will be no gains at all. The Southern country will just lose tariff revenue. But even if third country imports do cease what makes Northern firms cut their prices to their domestic levels? The answer is competition, and if anything occurs to curtail competition between Northern firms in the Southern market, the price cuts will be curtailed or, indeed wholly absent. Among the things that could curtail competition are collusion (tacit or otherwise) not to compete in these ‘captive’ markets, economies of scale in selling (e.g. consignment size) so that it is not worthwhile for more than one or two firms selling there, and product differentiation. The last increases the chances that the post-tariff prices of Northern goods fall somewhat, but curtails the effects of such price falls on competing domestic suppliers.
The Southern government would, of course, lose tariff revenues: to domestic consumers to the extent that consumer prices fell, and partly to Northern producers if their prices exceeded those at which imports occurred before the RIA. As we saw in chapter 2, the former is just a transfer within the Southern country, but the latter represents a transfer of income to the Northern partner. The most striking example of trade dependence between developing and developed partners is probably NAFTA. In 1991, before NAFTA was a real possibility, 70% of Mexico’s imports came from the US, and for 614 out of the 4,854 trade headings in which imports occurred, the US or Canada were the only sources. By 1996 the share of imports from the US had risen to 78%, but, in line with the general liberalization of the economy accompanying NAFTA, the number of imports brought only from NAFTA partners had decreased to 296. Without information on prices, it is impossible to quantify the effects, but Mexico’s continuing extensive imports from non-preferred suppliers makes it likely that its internal prices are significantly influenced by the remaining tariffs against them. Moreover, because Mexico’s tariffs are typically well above US tariffs, Mexico’s losses of tariff revenue on imports are likely far to exceed its gains on its exports to the US. Thus Mexico’s gains from NAFTA in the simple trade dimension are likely to be considerably less than those possible if tariffs were removed from all import sources.

North-South RIAs are typically FTAs—NAFTA and the Euro-Med agreements—rather than CUs. When they involve customs unions—as with Turkey-EU—the Southern partner is effectively obliged to adopt the lower Northern tariff, implying a substantial measure of trade liberalization. This is a further source of gain.

3.2 How many RIAs?

For an individual country an attractive strategy may be to maintain many RIAs simultaneously. This would combine duty-free access to many markets with zero tariffs on imports from many sources. Putting aside the very real danger that different FTAs have different administrative requirements such as rules of origin (Section 3.3), combining many RIAs could effectively substitute for free trade. Several countries now pursue this strategy in essence. For example, Chile is party to 11 trade agreements, (APEC and LAIA, plus bilateral links with Argentina, Bolivia, Canada, Colombia, Ecuador, MERCOSUR, Mexico, Peru, and Venezuela). Panama is a member of 9; Mexico of 8 (with 5 under negotiation); Bolivia, Costa Rica and Nicaragua of 5; and El Salvador, Guatemala, and Honduras 4 each. The picture is equally complex for Eastern Europe, with the Slovak Republic belonging to 9 RIAs, the Czech Republic and Slovenia belonging to 8, Estonia to 6, and Hungary, Latvia, Poland and Romania to 5. The pattern of RIA memberships in Eastern and Southern Africa is given in figure 3.2, which reveals a pattern of overlapping blocs.
The EU has a wide range of RIA’s – with European and Mediterranean partners – and, counting its non-reciprocal trade agreements, actually offers tariff preferences to all but ten of its trading partners (Winters, 2000).

Modeling results on Chile in a multi-regional CGE suggest strong benefits from such additive regionalism (Harrison, Rutherford and Tarr, 2000). Figure 3.3 estimates the increase in Chile’s economic welfare from signing FTAs with an ever-increasing set of partners.
Chile’s current agreement with Mercosur looks as if it could reduce welfare, at least on the simple trade criteria applied here. An FTA with NAFTA would be more attractive, as it offers better market-access benefits. Indeed, all the positive benefits recorded are due to gaining tariff-free access to the US market for non-grain crops; without this Chile would lose. Adding the two FTAs offers overall gains of 1.48% of GDP. This exceeds the sum from the two separate FTAs by 0.87%, which is broadly a measure of the extent to which liberalizing two sources of imports simultaneously reduces trade diversion. Adding the EU and the rest of Latin America to the cocktail also adds value strongly, both via further reductions in trade diversion and by increasing Chile’s access to export markets. These estimates should not be taken too seriously as absolute values, but the basic insight that, provided it is administratively straightforward, additive regionalism could generate benefits, is worth noting.

One problem with this incremental approach is that Chile’s early partners thought they were negotiating access to a market that was effectively protected by 11% tariffs against everyone else, and were prepared to ‘pay’ for that privilege with their own concessions. They then discovered that Chile was selling essentially the same deal to other partners. In two cases this led to protests, which Chile addressed by accelerating tariff reductions on the complainers’ exports. The solution to partner complaints about unforeseen liberalization was yet more liberalization – a good outcome for Chile – but it is clear that the threat of this kind of problem could undermine the whole approach. Managing multiple RIA’s is clearly not easy. However, if credibility and administrative skills are high, one’s country is an attractive partner, the FTAs cover the bulk of trade, and the cost of trade diversion incurred before a large number of FTAs is in place is not high, it could be a useful strategy.

The benefits of access to Mercosur and to NAFTA markets are basically independent of each other, hence the change in the overall benefits must reflect the combination of the import effects.
Moreover when we come to look at the dynamic effects of RIAs in chapter 4 below, we shall see these might also arise from multiple RIAs. The best example is the hub of ‘hub and spoke’ regionalism. If one country (or group of countries) has RIAs with a number of countries that maintain barriers between each other, then this hub country becomes the preferred location for investment – firms can reach more markets tariff free than they can from any of the other locations – and this will tend to bid up factor prices and raise real income in the hub. The world’s largest hub is the EU, which has separate RIAs with most other European and many Mediterranean countries, most of which do not grant each other free trade.

A quite different situation arises if a country’s multiple RIAs conflict. Customs union members must all set the same external tariffs, yet Colombia, Venezuela, Ecuador, Bolivia and Peru form the Andean pact (a customs union), while Colombia and Venezuela are also in the Group of 3, a free trade area with Mexico, and Bolivia has formed an FTA with Mercosur. Similarly (see Figure 3.3), Namibia, Swaziland and Lesotho belong to COMESA while also belonging to SACU, a customs union with South Africa and Botswana. And Tanzania belongs to SADC while being a member of EAC, a customs union with Kenya and Uganda. These obligations are formally contradictory, so it may be unclear which will prevail in practice, and special conditions and exclusions may have to be formulated. In addition the administrative aspects of the various agreements tend to differ – most notably in terms of rules of origin (see section 3.3 below) – creating a mass of complex rules for traders to cope with, and, frequently, seek to avoid.

### 3.3 Free Trade Areas vs. Customs Unions.

A central issue for countries planning to integrate their trade is whether to choose a free trade agreement (FTA) or a customs union (CU). Of the 162 RIAs notified to the WTO/GATT by August 1998, 143 were FTAs—with zero internal tariffs but no harmonization of external tariffs—and 19 were CUs—with a common external tariff as well. Under most circumstances, CU are more efficient and allow greater market integration, but they require more coordination and tighter constraints on individual member policies and sovereignty.

#### 3.3.1 Trade Deflection and Rules of Origin: More Protection

A major worry for FTA members is trade deflection, the redirection of imports from third countries through the FTA member with the lowest external tariff. If unconstrained, this reduces the effective tariff of every member to that of the lowest plus the transportation cost involved in indirect importing—which is wasted. The usual solution is rules of origin (ROOs)—the apparently reasonable requirement that goods qualifying for tariff-free trade should be produced in a member country, rather than just

---

5 Source: WTO. Of course, many RIAs contain other elements as well; this classification refers only to their policies on trade in goods.
passing through them. In practice, however, ROOs often become instruments of protection.

ROOs can lead to trade diversion when exportables between FTA members are not wholly produced within a partner country, but rely partly on inputs imported from non-member countries. Under rules of origin, exports have to derive a certain proportion of their value from local content or undergo certain production processes within the FTA to obtain duty-free treatment. Trade diversion will result if the ROOs create an incentive for producers in one partner to purchase higher-cost inputs from another, despite there being cheaper inputs from the rest of the world. ROOs can also artificially increase domestic sales if the favored input is domestically produced.

While trade in the NAFTA countries had a high regional content even before the RIA was formed (Cordoba, 1996), its rules of origin have serious protective effects in certain sectors, shifting their trade and investment patterns from lower- to higher-cost sources. Most clothing produced in Mexico gains tariff-free access to the North American market only if its inputs are virtually 100 percent sourced in North America (WTO, 1995). In the automobile industry, the origin requirement of 62.5% local content has induced Japanese automobile manufacturers with plants in Canada to invest to produce components in the US rather than import cheaper ones from Japan. The rules of origin also require the tubes in color televisions to be of North American origin if the televisions are to receive duty-free treatment. Since the inception of NAFTA in 1994, five television tube factories have been planned or established in North America by Japanese or South Korean firms, probably at the expense of expansion in Southeast Asia (Stephenson, 1996). Even tomato catsup is affected – see Box 3.1.

---

**Box 3.1 Rules of Origin are Protectionist**

To Chile’s dismay, the tomato catsup rules changed when CUSFTA evolved into NAFTA. Under CUSFTA, catsup processed out of imported tomato paste qualified for duty-free treatment in internal trade. Under NAFTA rules the tomato paste itself must be produced within a NAFTA member in order for the catsup to qualify for free entry. In 1992 Chile was the leading foreign supplier of tomato paste to the United States and the catsup it went into enjoyed free entry under CUSFTA. Mexico and Chile together accounted for over 80% of US tomato paste imports in roughly equal quantities. Under NAFTA, which meant that catsup made out of Chilean paste could no longer circulate duty-free, Chile’s share dropped to 5%, with Mexico taking the other 75% - Palmenter (1993).

---

Rules of origin effectively allow protection to be “exported” from one member to another even if each country maintains its pre-integration structure of external tariffs. The more restrictive the rule of origin, the greater the scope for trade diversion on intermediate products. At some point, however, the rule becomes so restrictive that producers opt to source inputs from outside the free trade area and forego duty-free access.
ROOs also pose governance problems for developing countries. They take considerable effort to negotiate, and are complex and opaque to operate. The EU’s agreement with Poland has 81 foolscap pages of small print in its ROOs section, and NAFTA some 200 (Krueger, 1997). FTAs require controls on products crossing internal frontiers to ensure compliance with rules of origin and the payment of customs duties for non-complying imports, which can pose significant administration costs. Some years ago these costs were estimated at 3 to 5% of f.o.b. prices for EFTA-EC trade. Documentation and verification—and their costs—must be a continuing part of FTA arrangements for importing countries to avoid tariff revenue losses and keep the protection system effective. Finally, ROOs allow customs authorities—and individual customs officers—a good deal of discretion. The administrative cost of ensuring that this discretion is not abused is considerable; the cost of not doing so is even higher.

3.3.2 Indirect trade deflection: Exporting Protection

Rules of origin may prevent the FTA member with the lowest tariff rate from importing goods and sending them on duty-free to another. But they do not prevent a low-tariff partner from meeting its own requirements for a product from the rest of the world, and then transferring a corresponding amount (or all) of its own production to its partners. This is indirect trade deflection, and its consequences for efficiency depend on the capacity of the low-tariff country to supply all its partner’s needs. If it can meet the latter’s entire import requirements at its own tariff-inclusive price out of its own production, it will render their high tariffs ineffective, and at the same time earn additional rents. In the limiting case that every good fell into this class, indirect trade deflection would render a free trade area equivalent to a customs union that takes the lowest pre-union tariff as the basis for the CET – Robson (1998).

If, on the other hand, the low-tariff country cannot meet the full import needs of its partner, the latter continues to import its marginal requirements from the rest of the world, and the domestic price remains anchored to the world market price plus its unchanged tariff. Then, although the exporter enjoys additional rents—the high-tariff country’s foregone duties—there would be no beneficial effects on resource allocation in the high-tariff partner, and real resources would be wasted as trade was diverted from third country to partner suppliers.

Unfortunately, there has been no empirical investigation evidence of indirect trade deflection within RIAs, but it has been identified working across trading blocs. Trade deflection through the EU from where into the US in textiles and clothing has modified, or even nullified, the effect of increased US non-tariff barriers (Hamilton 1988).

3.3.3. Customs Unions offer lower trading costs and greater integration

---

6 Agreeing on the rules for clothing and autos was a major issue in NAFTA – Krueger (1997)
6 Herin (1986).
Customs unions have common external tariffs, and so do not formally need rules of origin. Thus, provided that they also define and enforce any non-tariff protection measures at the bloc level, CUs can, at one stroke, avoid all the administrative costs and distortions of ROOs. Harmonizing non-tariff barriers is quite a demanding requirement, however: for example, throughout its first 30 years the EU allowed members to maintain their own quotas on certain third country imports (e.g. clothing, footwear, and steel) and use border measures to prevent those goods from crossing internal borders (Winters 1993). The EU-Turkish customs union allows the parties to impose anti-dumping duties on each other, which immediately requires the presence of border formalities and rules of origin to define partner goods. Similarly, if, say, industrial or safety standards differ between members, border controls or their equivalent will be necessary for enforcing them.

In effect, a customs union needs to have not only a common external tariff, but also a trade policy that is common in all respects. Thus they are a good deal more complex to create than FTAs. While CUs offer greater market integration and lower costs, they also require more ongoing coordination. Adopting a CET means reconciling the interests of member states and then establishing continuing political arrangements to deal with subsequent adjustments—such as in rounds of global trade talks, and imposing temporary safeguards, anti-dumping or anti-subsidy duties. Those arrangements may imply substantial sovereignty losses over trade policy instruments and revenue sources. FTAs also need an initial reconciliation of interests, effected through the character and restrictiveness of the rules of origin that are adopted, but once these are settled only relatively light institutional arrangements will be needed.

3.4 External Trade Policy

The net benefits of RIA membership depend directly on the external trade policy stance. There are strong arguments for pursuing a policy of external openness in conjunction with regional integration.

First, trade diversion is both more likely, and more costly should it occur, the higher are external trade barriers. It is more likely, since the relative price differences created by preferential liberalization will be greater with a higher external tariff, inducing trade diversion in more sectors. And it is more costly, since a higher external tariff will provide greater incentives for inefficient sectors to expand. Producers are able to charge high prices (because the tariff protects them from world competition), and capture what was previously tariff revenue on trade between members. Fundamentally, and just as for individual countries, the gains from competition with low cost suppliers – gains to consumers, gains in developing an efficient industrial structure, and competition-induced efficiency gains at the firm level – may be foregone if tariffs inhibit such competition. These are arguments both for low tariffs on average, and for tariff schedules that are relatively uniform, avoiding peaks. Very high rates in particular sectors are almost certain to produce diversion – as in EU agriculture.
A common complaint about such advice is that forming or deepening an RIA causes adjustment costs, and that simultaneous external liberalization magnifies these to an unacceptable level. The problem with this argument is that adjustment costs are only worth paying at all, if the adjustment is in the right direction! High tariffs or tariff peaks might induce costly economic changes that move the country away from economic efficiency. Besides, simulation studies of regional integration involving developing countries and large, high-income nations or blocs (such as the EU) suggest that the adjustment costs associated with RIA implementation are as high as those that would arise if trade liberalization were implemented on a nondiscriminatory basis (Rutherford, Rutstrom and Tarr 1999).

These are strong arguments that governments should accompany their RIAs with a genuinely liberal external policy, but, of course, ‘should’ does not always determine policy. An important question is whether RIAs change the balance of incentives and forces for external liberalism. If so, one could start to answer the question of whether regionalism is a stepping stone or a millstone around the neck of multilateralism. In part this is determined by the interactions between the RIA and outside countries (which we consider in Chapter 9), but it also depends on the RIA’s internal incentives and constraints on trade policy.

As so often, there is no simple empirical regularity to resolve these questions. When the Andean Pact finally agreed a common external tariff (CET) in 1994 (15 years later than originally intended) it replaced average tariffs of 11.1% in Colombia, 10.2% in Ecuador and 11.8% in Venezuela with an average CET of 12.8%. Under NAFTA, Canada reduced tariffs on around 1,500 items imported from non-member countries, while Mexico raised tariffs on 503. Since the EEC was formed in 1957, the average common external tariff on manufactures has fallen from about 13% to about 3% after the Uruguay Round by 2001. Clearly it is necessary to consider the parts of the argument separately and in more detail.

### 3.4.1 Setting external tariffs in an FTA: A race to the bottom?

FTA members decide their own tariff policies—while keeping intra-bloc tariffs at zero. The interactions between members may reduce external tariffs for three reasons.

First, if tariffs on fellow members are constrained to zero, the optimal level on closely competitive goods from third countries will be relatively low to reduce trade diversion. Second, if there is trade deflection, high-tariff countries lose tariff revenue. If they reduce their tariffs to just below the level of their partners, they can recapture the revenue without affecting internal price or resource allocation. A series of such moves in a competition between members will tend to lower external tariffs (Richardson, 1995). Third, if duties on inputs used to make exports to other members cannot be rebated, high import tariffs render final goods exporters uncompetitive. This concern apparently lay behind Canada’s decision to reduce 1,500 tariffs on inputs in 1995, shortly after NAFTA started.

---

7 Echavarria (1998); neither Bolivia nor Peru is adopting the CET.
But there are three counteracting forces that can increase protection. First, rules of origin essentially export protection, allowing producers in one country to benefit from protection in another. Second, the result of fierce FTA competition in the intra-FTA segment of their business, might lead firms to seek protection against third-country imports in another. This view requires the firm to provide different goods or different varieties in the two segments, but that is not unusual. And it might be re-inforced by the third force for protection: as tariffs on partners are fixed rigidly at zero, lobbying resources become more plentiful for lobbying against third country imports.

The balance between these pro- and anti-protection forces will vary from case to case, but there is at least something in the view that FTA could encourage liberalism.

3.4.2 Trade policy institutions in CUs can increase protection

The situation in CUs is quite different. Creating a customs union provides an (unavoidable!) opportunity to review the tariff structures and create new institutions for determining trade policy. National tariffs must be harmonized at some agreed level, and in the process, international obligations—notably those to the WTO—must be respected. The WTO rules are not a particularly good guide to economic policy (Chapter 9) and CUs are best served by using the domestic advantages of a liberal unilateral trade policy as their guidebook. Tariffs should be low, and the number of rates very few. Non-uniformity across products tends to increase administration cost and encourage discretion and corruption, and if the CET is more uniform than the rates it replaces there will be an additional benefit to the RIA.

Unfortunately, the forces created by a CU do not lead in particularly liberal directions, although care in establishing the CU and its institutions can help. By coordinating their trade policies, CU members may be able to increase their negotiating power against the rest of the world. If they are able to negotiate effectively as a bloc, which does not always happen, this will change the nature of world trade talks, probably in the direction of greater protection (Chapter 9).

In the most hegemonic of CUs—SACU—South Africa simply decided on trade policy and compensated the smaller countries for the costs it imposes on them. Similarly, Brazil dominated the determination of Mercosur’s external tariff (Olárreaga and Solóaga, 1998). Large countries are generally less dependent on international trade—and more prone to protection—than smaller ones, and hence tend to support increased protection.

Even in more evenly balanced CUs—the EU—it may pay to allow one member disproportionate power over certain negotiations. If tariffs in the CU and the rest of the world are strategic substitutes (i.e. if, as one bloc increases its tariff, the optimal response of the other is to lower its), letting the more aggressive member “lead” negotiations with the rest of the world on an issue, will result in higher CU tariffs and lower tariffs abroad than if the CU negotiated collectively. In this outcome, the lower tariff in the rest of the world increases CU welfare, possibly by enough to offset the costs of its own tariff. If, on the other hand, tariffs are strategic complements the CU would gain by letting its less aggressive member ‘lead’, resulting in lower tariffs and higher welfare all round (Gatsios and Karp, 1991, 1995). Unfortunately there are no general rules for determining which
tariffs are complements or substitutes (even in simple models), so it is not clear whether delegation within the CU would lead to higher or lower external tariffs.

Moreover, once one recognises that tariff setting is a continuing process (a repeated game, in economic jargon) the situation becomes even more complex. A more aggressive leader will be able to extract a more favorable deal because its threats to retaliate (with the whole of the union’s resources) will be more credible. Whether this leads to lower protection overall, however, depends on whether a more aggressive union can achieve a more liberal outcome by virtue of its readiness to retaliate, or whether it actually needs to use its retaliatory muscle. The smaller the union, the less likely it is to achieve a liberal outcome overall.

By convention, the EU allows countries disproportionate influence over policy in areas in which they claim vital interests, up to (and including) veto power. Given that a country’s interest in a sector is commonly related to that sector’s share in its GDP, it is easy to imagine this feature enhancing the interests of producers. The effect that this has on a CU’s trade policy depends in part on whether a sector’s having a high share of a member’s GDP reflects comparative advantage or past policy distortions. If the former, one might expect relatively liberal stances; if the latter, protection will be strongly defended. One encouraging aspect of this is that, as any trade creation will tend to relocate production in a sector toward relatively more efficient members, over time this could reduce protectionist pressure.

Even with genuinely inter-governmental decision making in a CU, several features can lead to protectionist biases in the process of aggregating individual members’ preferences into a common policy. Policy will generally be made by bureaucrats and ministers representing their own governments. This can lend the process a protectionist bias. The incentives for bureaucrats, who can reap no direct reward from the profits they create, tend toward the protectionist (Messerlin, 1983), and adding layers of inter-governmental decision making tends to swing influence away from voters and toward official preferences for administrative convenience and a quiet life (Scharpf, 1988). If there is no single political location at which the costs and benefits of protection come together—such as the presidency in the USA—the costs, which are usually spread thin and wide, tend to be dominated by the more concentrated and obvious benefits and a bias towards producers and protection introduced. If policies are enshrined in the CU’s constitution—such as in the EU’s Treaty of Rome, which defines the need for and objectives of the Common Agricultural Policy—or if they acquire their own bureaucracies—again, as with EU agricultural policy—interventions are legitimized and defended from within, and reform can become very difficult (Winters, 1997).

Problems can also arise if the costs of policies are allocated across countries in different proportions than their benefits. If a government wishes to minimize the danger of its nationals having to bear the cost of protection (consumer costs and higher factor prices) with no payoff for its own producers, and if it believes that it may not be able to

---

8 Such a sector may prefer high EU protection, so that it can reap high rents on EU sales, but at least it could survive with lower protection. It may also seek liberalization in other import sectors to reciprocate for any liberalization that it seeks in its export markets.
defeat the calls for protection, it will tend to join protectionist coalitions and seek protection for its own industries. The solution is to ensure that some CU-level institution (a hegemony or a central body) takes responsibility for all the costs and benefits of the whole package (Box 3.2).

**Box 3.2 Restaurant Bills, Universalism, and Protection**

Institutions for taking decisions about protection can impart a protectionist bias to the outcome. Two simple but instructive models of institutional failure are the restaurant bill problem and universalism.

*The Restaurant Bill Problem:* Four acquaintances go to a restaurant and decide, before ordering, to split the bill four ways: is it surprising that the bill is higher than when each pays for his own meal?

Similarly, suppose a country’s benefits from a policy on a product are proportional to its share of CU output of that product while its costs proportional to its share of GDP. A production subsidy financed out of general taxation will have this characteristics as, approximately, will a tariff, which benefits producers at the expense of consumers. If each country has a veto, or if consensus is valued very highly, countries sitting down to decide a package of price-increasing policies on several products will press for inclusion of any good for which their share of production exceeds their share of GDP. As each country is likely to have several products for which this true, the easiest package to build—and sell back home to local pressure groups—will increase the prices of nearly all products even if, overall, each country would prefer “no change” to the final outcome.

*Universalism:* Imagine that protecting steel is being discussed and that each of three member states in a CU produces one type of steel. If any one type is protected, the government in whose country it is produced receives all the “benefits” (profits, employment, political convenience etc.), but each member bears some of the costs, which exceed total benefits because protection is inefficient. Net costs are zero if the measure is rejected (Shepsle and Weingast, 1981).

The difficulty facing a government in deciding how to vote lies in taking the position, “we oppose this measure in principle, but if it passes, we want our a share of the spoils.” The worst outcome is that protection passes, but our producers are on the outside. Thus, if governments fear that they might not be able to stop a protectionist measure, they vote for it, but on the condition that it cover their own producers.

Simple arithmetic (Winters 1994) shows this situation is quite likely to arise in small groups of decision makers, and experience suggests the same outcome. Universalism drove the introduction of the USA’s 1929 Smoot-Hawley tariff, with individual Congressmen acquiescing to protection for their fellows’ ‘pet’ industries in return for reciprocal treatment for their own (Schattschneider 1935). In the EU, Britain opposed the inadequate McSharry agricultural reforms in 1992, but, instead of being adamant, used most of her influence to ensure that large farms (of which Britain has relatively many) were not excluded from income support mechanism, as McSharry had
originally proposed. In 1995 the EU Council of Ministers passed 92 of its 94 common trade policy decisions unanimously. Given that trade policy typically redistributes real income from one member to another, this certainly suggests that something like universalism is going on (Bilal 1998).

Institutional responsibility for policies must be clear. In the EU the struggle between the European Commission, representing the center, and the national capitals for control of non-tariff barriers to imports arguably resulted in more active use of those instruments than the parties would otherwise have wished. The Commission had to prove that it could use NTBs to meet national objectives in order to win control of them; while its measures were possibly less restrictive than national ones would have been, they applied to all members rather than just one (Winters, 1993).

3.4.3 Lobbies bias RIAs towards trade diversion

Lobbying is a fact of life for trade policy makers, and creating an RIA gives interest groups new threats to manage and new opportunities to influence policy. When asked their views on the FTAA, the Florida Citrus Mutual said that it would “mean the end of the US industry” and that “citrus products must be exempted from further tariff cuts.” The Rubber and Plastic Footwear Manufacturers Association said that eliminating duties would “cause havoc” and spell the death of the industry. Exempting footwear “would have virtually no impact on any country’s balance of trade figures and would in no sense violate the [WTO requirement that the FTAA cover] substantially all trade”.9 When NAFTA was signed some US industries were granted 15 years to adjust to free trade—including those (such as citrus) that already had the highest protection. In addition, new and more protective ROOs were formulated for textiles/clothing and autos.

In 1998 UEMOA debated the details of its common external tariff within a structure that had been previously agreed to impose 5% tariffs on intermediate goods, 10% on capital goods and 20% on consumer goods. As well as adding some temporary surcharges, one of the notable features of this debate was the conclusion that cement was a consumer good worthy of 20% protection! In 1997, private sector protests led Madagascar to postpone tariff cuts promised under the IOC - Solignac Lecomte (1998, p.5). The Andean Pact allows members to charge tariffs below the agreed CET level—provided there is no local producer of the good concerned (Echavarria, 1998).

One does not need much experience of the world to recognize the fingerprints of the lobbyists on these stories, and, in fact, lobbying is a particular problem when forming an RIA. Not only does it lead to resistance to internal free trade and to liberal external policies, but it does so in ways that emphasize trade diversion.

Governments respond to pressures from domestic interest groups in order to increase popular support and their chances of re-election. Lobbying for a sector or an interest provides a “public good” in the economist’s sense that the policies sought apply to all group members, regardless of how much they contribute to the lobbying effort. The larger the group of beneficiaries, the more difficult it is to prevent people from “free

"riding"—benefiting without contributing; and the smaller the benefit that any individual receives the less likely he is to take the trouble to enter the debate. For both of these reasons consumers find it very hard to organize a lobby, so lobbying is dominated by producers, who typically organize along sectoral lines. This effectively gives profits double weight in the government’s calculations: once as a source of income in the traditional assessment of national economic welfare, and again as the source of lobbying support—which governments value in its own right. The heavy private sector involvement in current regional plans—APEC, TMP, FTAA—shows how seriously firms take the opportunity to influence this element of the business environment. RIAs are undoubtedly better for being business-friendly, but lobbying can bias an RIA too far toward producer objectives.

Trade creation can be a mixed blessing for a negotiating government: it generates surpluses for consumers at home and for exporters in the partner country, but reduces them for one of the main lobbying groups—domestic import-competing producers. Trade diversion, on the other hand, generates no such reduction in profits, and although it generates fewer consumer gains, that may matter less to governments concerned about reelection or keeping the support of business elites. If two governments can exchange trade-diverting concessions, trade diversion becomes attractive politics, even if it is bad economics.

This bias toward trade diversion has two major implications. First, regionalism may be attractive where multilateral liberalization is not—farmers and policy makers in the EU accept internal free trade as managed through the Common Agricultural Policy, but are not ardent supporters of global trade talks on agriculture. Brazil is keen to see Mercosur have internal free trade in information technology products, but was the major holdout on the (nearly) global Information Technology Agreement of 1996. Second, exceptions or delays in achieving free trade tend to be concentrated on sectors with strong trade creation potential, the very sectors that promise economic gains (Viner, 1950, Grossman and Helpman, 1995). If the creation of RIAs is driven primarily by the industrial attractions of trade diversion, regional liberalization can not continue all the way to global free trade, because the last step in that direction would necessarily generate only trade creation.

These dangers are elegantly demonstrated by Krishna (1998) in a simple three-country world in which policy is determined solely by its effects on profits. He shows that, considering two of the countries, the more trade diverting an FTA between them, the stronger its backing, and hence the more likely it is to come about. He then also shows that, following such an RIA, the backing for further (multilateral) liberalization with the third country is reduced. Included in this is the possibility that a multilateral liberalization that was feasible before the FTA would cease to be so afterwards. Very simply, while an exporter might find any liberalization of foreign markets worth lobbying for, having once achieved a regional liberalization, he might find that the (possibly negative) incremental returns to lobbying for global liberalization did not warrant the effort. Thus if the world
attempts to achieve multilateral free trade via regionalism, progress would stop at the intermediate stage: regional stepping stones would leave us stranded in mid-stream. 10

A further factor favoring lobbying for an RIA rather than MFN free trade, is uncertainty. A country’s exporters can be confident that they will benefit from an RIA partner’s tariff concessions if the latter are restricted to RIA members even if they are inefficient. If the market is opened to all comers, on the other hand, these exporters can be much less confident that they will be the beneficiaries. Hence, although global negotiations will open more markets than regional negotiations, for any given market, exporters are likely to lobby harder for regional than for MFN liberalization.

It is sometimes argued that multinational companies are a bulwark against protectionism because they operate on both sides of tariff borders. Multinationals do frequently seek lower barriers to their imported inputs, and this can effectively discipline local pressures for protection. But multinationals also frequently show strong protectionist instincts. In the early 1990s, as the Europe Agreements were under negotiation, West European motor vehicle producers penetrated Eastern European markets in a fairly non-competitive way with one firm setting up per market, their entry encouraged (and was allegedly conditional upon) fierce import restrictions. Similarly, as Brazil raised barriers to non-regional production, world auto producers expanded their capacity in Brazil. Further evidence of the role of pressure groups in defining Mercosur’s CET and in obtaining exceptions both to it and to internal free trade is given in Box 3.3.

---

**Box 3.3 Pressure Groups and Mercosur**

With its highly non-uniform Common External Tariff—with numerous exceptions—Mercosur is an ideal case study of the role of pressure groups in RIAs (Olarreaga and Soloaga 1998). It took several years to negotiate Mercosur’s CET, culminating in the December 1994 Ouro Preto Protocol. Each member was allowed an exceptions list and the initial CET applied only to about 75% of the universe of 9,119 tariff lines (table). Left out were capital goods, computer and telecommunication equipment, and automobiles and sugar. Convergence to the CET should be achieved for most goods by 2006; there is no agreement on convergence dates for sugar and automobiles. The Ouro Preto Protocol also established a list of deviations from intra-Mercosur free trade. These were planned to disappear by 2000.

Olarreaga and Soloaga tried to explain variations and deviations from the CET across 27 industries, using measures of political and interest group activity such as wages, industrial concentration indices, labor/capital ratios, import penetration, intra-industry trade, and trade creation. Their conclusions included:

---

10 Models in which citizens vote directly on their own interests can also generate a similar outcome — Levy (1997) — especially if the partners are similar and relatively capital-abundant compared with the rest of the world (e.g., the EU).
• Sectors where significant trade creation is likely tend to be exempted from internal free trade. Internal free trade is also resisted more successfully by sectors with high employment shares. This reflects government desires to avoid large-scale labor adjustments, and the voting strength and trade union presence of large sectors. High third-country import penetration, which puts competitive pressure on domestic firms, also leads to high protection.

• Political considerations explain 58% of the variation in the CET across industries. The negotiated CET in any sector correlates with the share of capital remuneration in value added in that sector and the share of sector-specific capital in total inputs. If labor is fairly mobile and new firm entry difficult, it is existing capital that receives most of the benefits of protection. They also found a positive correlation with industry concentration: more concentrated industries find it easier to organize lobbying efforts, making them more effective lobbyists.

• The CET in a sector mainly reflects the preferences of the member country that has the greatest production in that sector. Thus, Brazil’s preferences are the main determinants of the structure of the CET, as it represents at least 70% of Mercosur production in each of the 27 sectors considered.

<table>
<thead>
<tr>
<th></th>
<th>Deviations from the CET</th>
<th>Deviations from Internal Free Trade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of tariff lines</td>
<td>% of total tariff lines</td>
</tr>
<tr>
<td>Argentina</td>
<td>1,540</td>
<td>17</td>
</tr>
<tr>
<td>Brazil</td>
<td>1,605</td>
<td>18</td>
</tr>
<tr>
<td>Paraguay</td>
<td>2,101</td>
<td>23</td>
</tr>
<tr>
<td>Uruguay</td>
<td>1,961</td>
<td>22</td>
</tr>
</tbody>
</table>

Note: There are several other deviations from the CET not indicated here, e.g., pre-existing special promotion regimes, and the tax-free areas of Manaus and Tierra del Fuego.

3.4.4 RIAs provide a new environment for Lobbying

As well as providing an opportunity to influence the new tariff structure itself, forming an RIA also changes the environment in which pressure groups seek changes in trade policy. Unfortunately, however, we do not have much idea in which direction this will push.

Some scholars argue that lobbying pressure is diluted by customs unions. It costs more to lobby for a 1% increase in your tariff in a customs union than in a single country:
there is more opposition to overcome (Panagariya and Findlay, 1994, de Melo, Panagariya and Rodrik, 1993) or more representatives to influence (Richardson, 1994). Given the lower individual returns and the greater difficulty of controlling free riding in a larger more diverse group, the sum of the member lobbying activity falls as a result of integration. The lobby from, say, Senegal does not wish to devote resources to lobbying for protection for producers in, say, Togo. Such spillovers will probably reduce average protection, even if this results in some members having higher tariffs on some goods than they would individually.

All these analyses presuppose, however, that fragmented lobbies face a unified CU government, whereas reality might be exactly the opposite. The RIA’s governments might be fragmented while some lobbies may have their power enhanced by integration (Winters 1993, and Bandyopadhyay and Wall, 1999). For example, each member might start with a lobbying game in which industry and agriculture more or less cancel each other out. But if integration lets agriculture lobbies cooperate (because they produce the same things) while the industry lobbies compete (because they produce different things) the CU may end up with agricultural protection. Overall, therefore, while dilution effects will undoubtedly be present, they will not always predominate.

Re-enforcing dilution effects, an RIA that generates trade creation can create a liberalising momentum. Industries that shrink because of lost protection eventually lose lobbying power. By squeezing relatively inefficient sectors and promoting more efficient ones, regionalism might ease future liberalization. Hathaway (1998) attributes the decline in US footwear protection since the 1970s to the sector’s declining ability to lobby effectively as it contracted. Employment fell from 216,000 in 1960 to 58,000 in 1990. The industry association found it ever harder to obtain import restrictions and instead started to focus on helping firms adopt new technologies. In 1990 its president decided to “stop spending one more penny or one more minute … on import restrictions” and the association started to admit importers into its ranks. In 1991 it decided not to oppose tariff cuts on footwear as part of the US Caribbean Basin Initiative. Their reaction to the FTAA, however, see p. ## above, suggests that seeking special protection is a hard habit to kick.

On the other hand, an RIA may well concentrate lobbying activity against outside countries. Once an RIA has been firmly agreed, there is no point in lobbying for protection against partner suppliers. This may well turn attention to third country suppliers and it also means that lobbying resources will be released to make lobbying against them cheaper than previously. The result is likely to be greater pressure on external tariffs (Findlay and Panagariya, 199#).

It is certainly the case that lobbying activity has increased hugely in Europe since the advent of the EU, and that it has become much more heavily focused on Brussels. The number of lobbying organizations in Brussels grew from 300 in 1970 to some 3,000 in 1990 (Anderson and Eliassen 1991); expenditure on lobbying was approximately $150 million in 1990, and increasingly rapidly. By 1998, there were 13,000 professional lobbyist in Brussels—approaching one for every European Commission staff member (The Economist 14 August 1998)! Magee and Lee (1997) offer one of the few formal
attempts to quantify their effects. They analyze external protection in France and Italy as the EC widened and deepened over 1968-83. The major effect was the supposedly exogenous OECD-wide liberalization trend, which broadly halved average tariffs from 15% to 7.5%. In addition, however, they identified a 1.7 percentage point increase in the average due to increased pressure from industries suffering from trade creation, and where mergers had increased industry concentration. This was partly offset by a 1.1 percentage point decrease due to the dilution of lobbying efforts as the political arena enlarged from national to EEC bounds and some industries grew rapidly as a result of increased market size and efficiency gains.

It is not possible to say a priori whether, once it is formed, a CU will be more or less vulnerable to lobbying. As noted above, much will depend on the institutions it adopts.

3.4.5 RIAs and protection—a summing up

Unfortunately there is too little empirical evidence to balance the arguments in this and the previous section against each other accurately: it remains an empirical matter whether RIAs will increase or decrease protection. Foroutan (1998) suggests that RIAs have not increased their protection levels recently, but the evidence is not particularly strong (these results are discussed in Chapter 9 (Box 9.5), when we consider external pressures on RIA trade restrictions). Whatever the net effect, however, there are likely to be benefits from establishing good trade policy institutions at the CU level.

Members should avoid excessively bureaucratic decision-making methods, ensure that a single institution takes responsibility for the whole package of measures, and avoid writing protectionist policies into an RIA’s basic documents. Even where members want an explicitly inter-governmental approach and eschew a central authority for the CU, they should establish a central informational/analytical body responsible for determining and widely disseminating the collective CU interest in trade policy. Its composition and constitution, however, should avoid equating benefits solely with production or high intra-bloc trade shares with the collective interest.

3.5 Integration and taxes

3.5.1 Tax Revenue

Many developing countries are heavily dependent on trade taxes as a source of revenue, with some African countries raising as much as one half of government revenues from them.11 Membership of an RIA erodes these revenues – directly, as tariffs on intra-RIA trade are reduced, and indirectly, when trade diversion occurs, as importers switch away from external imports subject to tariffs. It can also arise if a CU sets tariffs at levels below a country’s pre-RIA levels. As we saw above, the loss of government revenue lies at the heart of the trade diversion argument, with revenue being transferred

---

11 Trade taxes include export and import taxes, trade monopoly profits remitted to governments, and taxes and official profits on foreign exchange transactions.
to partner producers and representing a direct loss to our country. However, if the government has difficulty in mobilising alternative revenue sources and is spending its revenue efficiently, then even losses that get transferred to domestic consumers – such as when tariffs on non-members fall - could be costly. Thus forming an RIA may need to be accompanied by steps to improve the tax take from domestic excise, sales, or value added taxes.\(^\text{12}\)

But how much revenue has typically been lost by RIA formation? In practice, we often find that countries which are less dependent on trade taxes lose larger amounts of revenue. This paradox arises because intra-RIA trade volumes are typically high in RIAs where dependency on trade taxes has been quite low (e.g. the EU), while countries with higher trade tax dependency have tended to form RIAs with countries with which they have relatively little trade.

However, there are exceptions to this. Cambodia derived 56 percent of its total tax revenues from customs duties prior to its entry into the ASEAN free trade area, with two thirds of these levied on imports from ASEAN countries (Fukase and Martin 1999). Entry into ASEAN provided a powerful stimulus for the introduction of a value added tax in early 1999. By the time UEMOA is fully operational it is predicted roughly to halve Senegal’s tariff take, from CFA 100 billion to between CFA 39 billion and CFA 63 billion. Of this, the bulk is due to reductions in external protection, with intra-bloc trade preferences accounting for no more than CFA 5 billion.\(^\text{13}\) In SADC also, where some countries are quite heavily dependent on trade with South Africa, substantial amounts of revenue are involved. Table 3.2 gives estimates of the revenue cost of going to free internal trade, and we see that this will approximately halve customs revenue in Zambia and Zimbabwe, losing the governments 5.6 percent and 9.8 percent of government revenue respectively. These are very substantial revenue losses, and point to the need to ensure that alternative tax systems are in place before removing sources of trade tax revenue.

### Table 3.2 Customs Revenue Collected as a Percent of Total Government Revenue in 1996 and the Implications of a Free Trade Area for SADC Members*

<table>
<thead>
<tr>
<th>Member Country</th>
<th>Customs Duty as % of Total Tax Revenue</th>
<th>Estimated Change in Customs Duty</th>
<th>% Custom Duty</th>
<th>% Total Tax Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malawi</td>
<td>14.3</td>
<td>-36.7</td>
<td>-5.3</td>
<td></td>
</tr>
<tr>
<td>Mauritius</td>
<td>29.8</td>
<td>-18.2</td>
<td>-5.4</td>
<td></td>
</tr>
</tbody>
</table>

\(^{12}\) It is important to remember that not every m.f.n tariff cut reduces tariff collections. The initial duty may be above the revenue-maximizing rate, especially taking account of exemptions and evasion. Thus, revenue replacement is not always as daunting as it first appears.

\(^{13}\) These estimates, from a CGE model, are made as if UEMOA were fully operational in 1996. They vary according to whether or not quantities of trade are allowed to change in response to the RIA, and if so, according to what assumptions. Ng and Winters (1998).
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>South Africa</td>
<td>3.6</td>
<td>4.9</td>
<td>0.2</td>
</tr>
<tr>
<td>Tanzania</td>
<td>24.0</td>
<td>-8.3</td>
<td>-2.0</td>
</tr>
<tr>
<td>Zambia</td>
<td>12.3</td>
<td>-45.3</td>
<td>-5.6</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>18.4</td>
<td>-53.3</td>
<td>-9.8</td>
</tr>
</tbody>
</table>

*Source: Staff calculations, IMF*

* There are discrepancies between the duty revenue reported by customs departments and that reported in budget numbers. For example Malawi reported FY96 duty revenues of 1,505.2 and 2,028.7 million kwacha against the 615 million reported by customs. For consistency, we have used the numbers reported by customs.

Note: The FTA assumes Free Trade on Intra-SADC trade. The projections assume that each country’s average tariff rates against SADC members are zero.

It is important to enquire what form revenue replacement takes, not only because inefficient taxation reduces economic welfare, but also because the net benefits of signing an RIA can depend critically on how distortionary taxes are. Chile’s FTA with Mercosur is predicted to require raising VAT rates by half—from an average of 7.9 % to around 12%, to maintain revenue-neutrality (Harrison, Rutherford, and Tarr 1997). But, in fact Chile’s VAT is fairly distortionary. If Chile improved collection and eliminated ad hoc exemptions to create a fully uniform VAT, it could halve the nominal average rate and increase real incomes by 0.3%, while at the same time making tariff revenue replacement cheaper and trade reform more attractive. Rather different results apply to the Egypt-EU FTA (Konan and Maskus, 2000). If lost revenue is replaced by current discriminatory taxation on capital the FTA would gain Egypt about 0.3% of GDP. However if the tax system is cleaned up, total gains of 1.4% are recorded and the marginal effect of the FTA is to reduce welfare by 0.2%. That is, an FTA may be better than current policy but once the gains from tax reform are reaped, the RIA is actually harmful.

For poor economies, import tariffs represent one of the cheapest and easiest ways of collecting government revenue, especially if they have only a few well-defined outlets for international trade. The observation above that in most cases RIAs between such countries pose few threats to revenue, raises another intriguing possibility. Maybe small countries can obtain a given tariff revenue at lower economic cost by combining into an RIA than by acting independently. The trade-off is straightforward to state if not to quantify. An RIA between small countries may offer benefits in terms of larger markets and more competition than non-distortionary tariff policy, but at the same time will require higher taxes on imports from third countries. If the former gains outweigh the latter losses (plus, of course, the trade diversion to which small-country RIAs are condemned), the RIA would help revenue-constrained countries.

There are several caveats to this argument, however. First, if an RIA is better than non-distortionary tariff policy and yet still entails an increase in tariffs against the rest of the world, there are clearly significant gains to be had by developing other sources of revenue, and this should receive high priority. Second, revenue tariffs, whether

---

14 Relative to the pre-RIA position, the partners import less from third countries (trade diversion) and need to make up the revenue foregone on intra-bloc trade.
discriminatory or not, should have different patterns from protective tariffs. If varied across commodities, they will be high on those with inelastic demand for imports, and have no correlation with whether or not there is local production. Probably better from a practical perspective, would be to have a single tariff rate, uniform across commodities. Uniform tariffs eliminate distortions between different imports, significantly reduce administration costs (which were, after all, the reason for using tariffs to raise revenue in the first place) and substantially reduce the scope for lobbying and/or corruption. Third, when there is no domestic production, a tariff is equivalent to an excise tax (which is levied on all sources of a good including domestic sources). Then it is better to call the measure an excise tax to make clear that, should domestic production grow up behind the tariff wall, it will be taxed at the same rate as imports. This will prevent the emergence of inefficient domestic activity in response to the side-effect protection associated with a revenue tariff. Fourth, as we have argued extensively so far, even without revenue considerations there is no presumption that an RIA between small countries will improve economic welfare; indeed, the balance of the argument is to the contrary.

3.5.2 Fiscal Compensation

A problem for many trade blocs among developing countries is that some members account for disproportionately shares of the bloc’s import-competing production and, hence, of the tariff-free trade within the bloc. This problem was evident in CACM, (where El Salvador and Guatemala accounted for the bulk of industrial production), CEAO (Côte d’Ivoire), the East African Community (Kenya), and SACU (the Republic of South Africa). The other members, whose losses of tariff revenues were translated into higher prices for industrial imports from the more industrialized countries (i.e. who suffered trade diversion), felt that they needed compensation.

These concerns are magnified if trade policies—as well as trade—are unbalanced. An extreme example of this is the Lebanon-EU FTA that is under negotiation. In 1995, Lebanon’s imports from the EU were LL3,547 billion with an average tariff rate of about 15%; its exports were LL143 billion with an average tariff of, perhaps, 4%. Potential revenue losses are about LL532 billion—about one-sixth of government revenue—but the gains on exports only about LL5 billion. Arguments for similar discrepancies have been made for Mexico in NAFTA (Bhagwati and Panagariya, 1997) and South Africa-EU (Teljeur, 1998).

In some RIAs compensation is explicitly ruled out in the negotiation process—e.g. NAFTA. Where it is not, the only practical way to estimate the warranted compensation in an existing RIA is to compare actual revenue with what would be collected if members imposed the external tariff on all their imports including those from other members. This is essentially how the SACU agreement handles compensation for Botswana, Lesotho, Namibia and Swaziland. But such estimates are essentially accounting exercises that make no allowance for the fact that the quantities and pre-tax prices of tradable goods—and welfare—change as a result of an RIA. Rather, compensation based on foregone revenues represents a political settlement to allow for

---

15 Data from Fulcihan (1997), Moukabel (1997) and Martin (1996).
the most obvious source of disparity in the costs and benefits of RIAs using a transparent and analytically plausible formula. But no simple formula can provide more than a crude approximation for offsetting the disparities in costs and benefits in a broader appraisal of an RIA. Nor, in the face of economic change, can any such formulae maintain the balance of advantage initially agreed by political negotiation.

One such broader element that governments sometimes claim in assessing an RIA is that protection helps to maintain value added and employment in manufacturing, and thus that imports from partners are economically costly even when they displace higher cost domestic output. This was among the arguments that led to SACU including an enhancement factor in its revenue distribution formula, which initially increased compensation some 40% above estimated revenue losses. The plausibility of such development benefits and the role of RIAs in achieving or frustrating them are among the subjects of the following chapter.

Two major problems beset such compensation arrangements, and make it very difficult to devise mutually acceptable compensation formulas. First, while the transfer benefits accrue to the private sector (which may be partly foreign owned), compensation has to be provided out of public revenues by authorities who often face practical constraints on revenue raising. Second, the costs that are compensated are direct, immediate and obvious, while any wider benefits from integration tend to be diffuse, uncertain, and accrue only in the longer term. The failure of compensation has often been accompanied by failure of the RIA itself—as happened to the East African Community and the CEAO—or the withdrawal of members—Chad from UDEAC, or Honduras from CACM. In the light of this experience, unless prospective partners are structurally balanced, actually and potentially, the formation of an RIA is highly risky without the simultaneous adoption of low external tariffs to minimize the transfers implied in duty-free intra-RIA trade. Low external tariffs also, of course, increase economic efficiency.

3.5.3 Tax Competition

Except for the very largest of countries, RIAs constrain members’ fiscal discretion and sovereignty. This is most obvious for customs unions, where the CET transparently removes tariffs from the policy domain of any single country. But there will also be constraints on a wider set of instruments in both CUs and FTAs: minimizing transaction costs, maximizing the tax base and avoiding smuggling, will all create pressure for a degree of indirect tax harmonization.

We have already seen how these factors have stimulated the formation of some RIAs, as countries have sought to manage their neighbors’ tax policies and eliminate smuggling by eliminating fiscal borders. Here we consider the opposite side of the coin, in which constraints are unwelcome side-effects of RIAs signed for other reasons.

16 A further argument was that SACU lacks any institutional arrangement for controlling administrative interventions by South Africa that hinder the industrial development of the smaller member countries, or, in general, for adjudicating disputes relating to the operation of the customs union.
If goods flow unimpeded across an RIA’s internal borders, governments may be tempted to lower indirect tax rates to capture revenue from cross-border shopping. This tax competition could lead to a widespread reduction in indirect tax rates. Such an outcome could be welcomed in that it brings about a downward convergence of effective tax rates, and leads to efficiency gains by reducing consumption distortions. The alternative view, however, is that disregarding such fiscal spillovers leads to a welfare loss for the RIA as a whole because tax receipts and total expenditure fall below optimal levels. This view is frequently argued by the larger continental members of the EU, and Box 3.4 illustrates the same problem in classical times.

Tax harmonization has been a major issue within the EU over many years, generating fierce controversy and even, at times, threatening to derail the Single Market Program. Recognizing that some harmonization is indispensable to the operation of a single market, members have agreed to adopt a minimum standard VAT rate of 15%, with no more than two reduced rates. No agreement was reached on precisely which goods and services fall into which tax bracket, however, or on the use of higher-than-standard rates. Efforts to harmonize excise taxes—on alcohol or gasoline—have formally been even less successful, although tax competition driven by demands from users and/or producers has induced a measure of de facto harmonization.

Fiscal discretion has also long been an issue in SACU, as the smaller members have virtually no role in determining tariffs or other indirect taxes in the customs union area. But even where they formally retain the power to set rates in a closely integrated customs area like SACU, the pressures to harmonize VAT rates at or close to South African levels are so strong that the BLNS countries still have little control over their indirect tax regimes.

Harmonizing taxes elsewhere than the EU and SACU—for instance, in UEMOA and UDEAC—has not been successful. For developing countries, the moral is not to underestimate the political difficulties of tax harmonization—and to recognize that trying to eliminate mild differences across borders may be more trouble than it is worth. In the US and Canada, for example, differences in state and provincial sales taxes of the order of 5 percentage points are an irritant, rather than a major problem.

**Box 3.4 Taxes: A Race to the Bottom**

Taxes can be a battleground for countries—and for traders. In the 4th century BC, Rhodes was a key commercial power in the Eastern Mediterranean, controlling the sea lanes and with a vibrant port. Rhodes charged a 2 percent tax on the value of cargo carried on all ships entering its harbor, including transit cargo. To escape the tax, Roman traders lobbied for the creation of a free port in Delos. Once established, trade rapidly shifted away from Rhodes, which lost most of its harbor tax revenues. But this tax competition shortly proved very costly to shippers: Rhodes had used part of its tax proceeds to police the sea lanes and prevent piracy. Without the revenue, these activities declined, piracy increased significantly, and trade became more costly (Adams, 1993, pp. 83-84).
References


83


