International trade; regional economic integration.

Prepared for the ‘International encyclopedia of Social and Behavioral Sciences’ (article 3.4 34), by Anthony J. Venables, London School of Economics.

Abstract:

Regional economic integration occurs when countries come together to form free trade areas or customs unions, offering members preferential trade access to each others’ markets. The article reviews the economic effects of such agreements on member countries and on the world trading system. Effects on member countries include the benefits and costs of trade creation and trade diversion, as well as gains from increased scale and competition. ‘Deeper’ integration can be pursued by going beyond abolition of import tariffs and quotas, to further measures to remove market segmentation and promote integration. Effects on the world trading system are not clear-cut. There is little evidence that regionalism has retarded multilateral liberalization, but neither is there support for the view that continuing expansion of regional agreements will obviate the need for multilateral liberalization efforts.

Correspondence address:

Anthony J. Venables
Department of Economics
London School of Economics
Houghton Street
London WC2A 2AE
UK

Email: a.j.venables@lse.ac.uk
3.4 34: Regional Economic Integration

Regional integration agreements (RIAs) are groupings of countries formed with the objective of reducing barriers to trade between members. They have long historical antecedents, sometimes being a stage to nation building (most famously in the Zollverein of nineteenth century Germany), and sometimes as part of colonial trading arrangements. In the post-war period, developments were led by the European Union (EU, originally the European Economic Communities, founded in 1958), and in the 1960s and 1970s there was a number of rather inward looking (and largely unsuccessful) RIAs between developing countries. Since the mid-1980s there has been a dramatic increase in regional integration activity. Of the 194 RIAs notified to the General Agreement on Trade and Tariffs/ World Trade Organization (GATT/WTO) at the beginning of 1999, 87 were notifications since 1990. Now almost all countries are members of at least one RIA, and more than one third of world trade takes place within such agreements. The new developments include the expansion and deepening of the EU; the construction of new and more open RIAs between developing countries; and the advent of RIAs in which both high income and developing countries are equal partners, lead by the North American Free Trade Area (NAFTA) which, in 1994, extended the Canadian-USA free trade agreement to Mexico.

RIAs come in many shapes and sizes. They vary in income levels, in openness to trade, and in the share of trade that takes place within the RIA, (60% for the fifteen members of the EU, but just 10% for the eight members of the West African Economic and Monetary Union). Structures vary, from the loose agreements to facilitate trade of the African Cross-Border Initiative and Asia Pacific Economic Cooperation (APEC) forum, through to the deep integration of the EU, involving the construction of shared executive, judicial and legislative institutions. The focus of this article, in common with most of the economics literature, is on RIAs as mechanisms for trade liberalization, not addressing the integration of factor markets, monetary systems, or political institutions. Even with this focus, RIAs differ widely in structure. ‘Free trade areas’ remove internal tariff and non-tariff barriers, but permit members to retain independent external tariff policies. ‘Customs unions’ go further, in fully harmonizing external trade policy. And it
is increasingly recognized that effective integration of markets requires more than reducing tariffs and quotas. Many other barriers – such as differing national product standards and a host of minor border frictions -- impede trade and support market segmentation, and some RIAs are now pursuing ‘deep integration’ policies to eliminate these barriers.

The material of this article is organized under four main headings. Section 1 outlines the main effects of RIAs on member countries, section 2 discusses some of the wider policy choices faced in RIA formation, and section 3 looks at the implications of RIAs for the world trading system as a whole. Section 4 concludes.

1. Economic effects on member countries.

1.1 Trade creation and diversion.

The modern analysis of RIAs dates from Viner (1950), who drew the distinction between the trade creating and trade diverting effects of RIA formation. The classical source of gains from trade is that global free trade allows consumers and firms to purchase from the cheapest source of supply, hence ensuring that production is located according to comparative advantage. In contrast, trade barriers discriminate against foreign supply, inducing domestic import competing producers to expand even though they have higher costs than do imports. This in turn starves domestic export sectors of resources and causes them to be smaller than they otherwise would be. Since a RIA liberalizes trade, reducing at least some of the barriers, doesn’t it follow that it too will generate gains from trade? Viner’s contribution was to show that the answer is; not necessarily. The gains from trade argument applies if all trade barriers are reduced, but need not apply to a partial – and discriminatory – reduction in barriers, as in a RIA. This is because discrimination between sources of supply is not eliminated, it is just shifted. If partner country production displaces higher cost domestic production then there will be gains – trade creation. But it is possible that partner country production may displace lower cost imports from the rest of the world, and this is welfare reducing trade diversion.

The analysis of trade creation and trade diversion constitutes one of the first
formal analyses of the more general problem of ‘second-best welfare economics’. Given that distortions remain in place in some activities in the economy, it is not necessarily the case that removing just some of the distortions (e.g. eliminating trade barriers on partner countries and leaving them in place on external countries) is welfare improving. In the literature on regional integration the response to the fundamental ambiguity created by the second-best took three main forms.

First, authors established circumstances under which there is no interaction between formation of the RIA and external trade flows, so no possibility of trade diversion. Meade (1955) pointed out that if trade barriers with non-members take the form of fixed quantitative restrictions, then a RIA must raise the total welfare of member countries since there is no possibility that imports from the rest of the world are displaced. Ohyama (1972) and Kemp and Wan (1976) showed how, when external trade barriers take the form of tariffs, it is possible to adjust these to hold external trade volumes constant, so preventing trade diversion from occurring.

Second, researchers identified conditions, in terms of changes in endogenous variables, for welfare gain. For example, welfare increases if the initial-tariff weighted change in trade volume is positive (Meade 1955). If internal tariffs are close to zero, then reducing them to zero raises welfare if it increases tariff revenues earned on external trade (Ethier and Horn 1984).

The third approach is to identify features of economies (in terms of their underlying exogenous characteristics) under which they are more or less likely to gain or lose from RIA membership. Lipsey (1957) argued that joining with countries that are already one’s largest trading partners is unlikely to lead to diversion, since the fact that the countries were originally the largest trading partners suggests that they are the lowest cost source of supply. Similar reasoning, including transport costs in the costs of supply, leads to the ‘natural trading bloc’ argument (Wonnacott and Lutz 1989, Summers 1991). Venables (2000) shows that those members of an RIA with comparative advantage most different from the world average are most likely to lose from trade diversion, as their trade is diverted to partner countries with comparative costs between theirs and the world average.
Empirical work on trade creation and trade diversion has taken two main forms; econometric studies of changes in trade flows, and computer simulation studies of the full general equilibrium effects of RIA membership.

Econometric studies seek to quantify the changes in trade flows attributable to membership of a RIA, and thereby identify trade creation and diversion. A variety of different econometric models have been developed, the most common being based on the gravity model which estimates bilateral trade between countries as a function of their GDPs, populations, the distance between them, and physical factors such as sharing a land border, and being landlocked or an island. Dummy variables capture whether or not countries are in a particular RIA, their estimated effect indicating whether countries in a RIA trade more or less than would otherwise be expected. Using this technique, Bayoumi and Eichengreen (1997) find that the formation of the EEC reduced the annual growth of member trade with other industrial countries by 1.7 percentage points, with the major attenuation occurring over 1959-61, just as trade preferences were phased in. Soloaga and Winters (1998) look at a wide range of RIAs, producing a mixed picture with little evidence of widespread trade diversion. Overall, there appears to be weak evidence that external trade is smaller than it otherwise might have been in at least some of the blocs that have been researched, but the picture is sufficiently mixed that it is not possible to conclude that trade diversion has been a major problem. Furthermore, it cannot be inferred that trade diversion has been economically damaging without information on relative costs and tariff structures, variables that are not revealed in this sort of aggregate exercise.

The second empirical approach is based on computable equilibrium modeling. This involves construction of a full computer model of the economies under study and use of the model to simulate the effects of the policy changes associated with the RIA. Such a model typically contains a great deal of microeconomic detail, so can be used to predict changes in production in each sector, and changes in factor prices and real incomes. In models that assume a perfectly competitive environment, the combined effects of trade diversion and trade creation typically give very small welfare gains -- just a fraction of 1% of GDP (see Baldwin and Venables 1997 for a survey). The strength of
these models is that they have sufficient micro-economic structure for the effects of a policy change to be traced out in detail, and its real income effects to be calculated. They are also often used for prediction – to estimate the likely effects of a policy change before it is implemented. But they have the major weakness that they are not usually fitted to data as carefully, or subject to the same statistical testing, as econometric models. The cost of the micro-economic detail is a complexity that makes rigorous econometric estimation impossible.

Although the focus of the trade creation and diversion literature has been on the changes in trade flows induced by regional integration, two consequent effects are important. The first is that changes in trade flows may change world prices, possibly improving the terms of trade of member countries, although this gain arises at the expense of outside countries. For example, if trade diversion occurs then RIA imports from outside countries are reduced, and any reduction in import prices that this causes is a terms of trade gain. Empirical work on this issue by Winters and Chang (2000) shows that Brazil’s membership in Mercosur has been accompanied by a significant decline in the relative prices of imports from non-member countries.

The second is that changes in tariffs and trade volumes will lead to loss of government tariff revenue. This can occur directly (as intra-RIA tariffs are cut) and as a consequence of trade diversion (as imports are diverted away from external, tariff inclusive, sources of supply). Its cost depends on the social cost of raising funds by alternative means, and can be severe in some developing countries. For example, in the South African Development Community, where some countries are quite heavily dependent on trade with South Africa, substantial amounts of revenue are involved, amounting to perhaps 5.6% and 9.8% of government revenue for Zambia and Zimbabwe respectively. Cambodia derived 56% of its total tax revenues from customs duties prior to its entry into the Association of South East Asian Nations (ASEAN), and Fukase and Martin (1999) argue that entry into ASEAN provided a powerful stimulus for the introduction of a value added tax.

1.2 Scale and competition effects:
A second mechanism through which member countries are affected by RIA membership derives from the fact that countries may be too small to support, separately, activities that are subject to large economies of scale. Regional cooperation offers a route to overcome the disadvantages of smallness, by pooling resources or combining markets. These scale benefits can arise in public projects (see World Bank 2000) and also at the level of the private firm, where they typically interact with imperfectly competitive market structures. These considerations are absent from the trade creation and trade diversion approach outlined above, which is based on the perfect competition and constant returns to scale paradigm of traditional trade theory. It was only in the 1970s and 1980s that formal analysis of the interaction between trade, economies of scale and imperfect competition began with the ‘new trade theory’, and these techniques have now been extensively applied to regional integration.

The basic argument is that there is a trade-off between the extent to which firms can achieve economies of scale, and the intensity of competition in the market. For a given size market, larger firms means fewer firms and hence more monopolistic outcomes. If regional integration combines markets, then it shifts this trade-off, potentially allowing firms to be bigger and markets to be more competitive (Smith and Venables 1988). For example, there might be an initial situation in which two economies each have two firms in a particular industry, and these firms exploit their ‘duopoly’ power, setting prices well above marginal cost. After formation of the RIA this becomes four firms in one combined RIA market. This increases the intensity of competition, and possibly induces merger (or bankruptcy), perhaps leaving only the three most efficient firms. The net effect is increased competition, increased firm scale, and lower costs. ‘Triopoly’ competition is likely to be more intense than the original duopolies, and surviving firms are larger and more efficient, so can better exploit economies of scale. A further source of gains comes from possible reductions in internal inefficiencies that firms are induced to make. If the RIA increases the intensity of competition, it may induce firms to eliminate internal inefficiencies (X-inefficiency) and raise productivity. Since competition raises the probability of bankruptcy and hence layoffs, it also generates stronger incentives for workers to improve productivity, and increases labor turnover across firms within sectors.
Although these are significant potential sources of gain, they have been difficult to achieve in many RIAs. This is addressed further in section 2.1.

1.3 Winners and losers:

A continuing concern is with the distribution of the costs and benefits of regional integration between member countries. Do central regions gain at the expense of peripheral ones, and do poor countries tend to catch up or get left behind? The evidence is, broadly, that RIAs composed of developed countries tend to show convergence (for example the narrowing of per capita income differentials observed in the EU, see Ben-David 1993). However, the picture for RIAs composed of developing countries is more mixed, with some examples of divergent performance (World Bank 2000).

The analytical literature on these questions is quite sparse, but provides several clues why this might occur. First, as mentioned above, trade diversion is more likely for countries with ‘extreme’ comparative advantage, suggesting that in a RIA amongst developing countries it might be the lowest income countries that experience diversion. For example, their imports of manufactures might be diverted from non-member countries to a partner that has a comparative advantage in manufactures within the RIA, but not relative to the world at large. Second, industries might tend to cluster in locations that have relatively good market access, or that are well supplied with business services or provision of other intermediate goods. This is more likely to occur in developing countries than in developed ones, partly because of their sparser provision of business infrastructure, and partly because the small size of their manufacturing sectors means that clustering is less likely to run into congestion and other sources of diminishing returns. The clustering may lead to wages being bid up in one member country at the expense of others.

2. Policy making

2.1 The depth of integration

The simplest form of regional integration is the elimination of tariffs (or quotas) between member countries. Beyond this there is a wide range of policy options open to
countries considering integration, many of which turn on the ‘depth’ of integration sought by member countries – ranging from modest trade liberalization, through full economic integration, to the formation of shared institutions.

The distinction between a free trade area and a customs union, in which a common external tariff is set, has already been made. The latter involves greater sharing of sovereignty and requires establishing procedures for revenue sharing, but in return can yield much greater market integration. In a free trade area where countries set different external tariffs the free internal circulation of goods is impossible; border formalities have to be maintained to ensure that external imports do not all enter through the member with the lowest external tariff, for re-export to other member countries. Since these imports include intermediate goods that are further processed in member countries, in practice this involves enforcing complicated ‘rules of origin’ governing trade flows within the RIA. (The EU’s agreement with Poland has 81 pages of small print in its rules of origin section, and NAFTA some 200, Krueger 1997). These rules are not necessary in a customs union, enabling simplification – or elimination – of internal border formalities.

It is increasingly recognized that tariffs and quotas alone may be just a small part of the overall barriers to trade created by an international border. Rules of origin create frictions, and so too does contingent protection (such as anti-dumping rules, known to have a ‘trade chilling’ effect even when not used), duplicative customs procedures, differing national product standards, and simple border red tape. The cost of border formalities on intra-EU trade in the early 1990s has been estimated to more than one percent of the gross value of internally trade, despite implementation of procedures to cut these costs; in many RIAs the costs of border formalities are much larger. Furthermore, relatively minor border frictions may constitute large real trade barriers because it can be in the interests of firms to try and limit effective cross-border competition in order to maintain collusive market sharing arrangements. This will prevent the ‘scale and competition’ effects discussed in section 1.2 from being achieved.

Recognition of the importance of these barriers – and of the failure of more than 25 years tariff-free trade to truly integrate markets – was a motivation behind the ‘deep
integration’ of the EU’s 1992 Single Market Program, (Flam 1992). The program involved adoption of almost 300 measures falling into the following five main types. (i) Simplification and in some cases abolition of intra-EC border controls, involving, eg, replacing border paper-work by an EU wide system of administering value added tax on cross-border transactions. (ii) Adoption of the ‘mutual recognition’ principle for product standards, under which a product that can be legally sold in any EU country can be legally sold in all, thereby removing the need for expensive re-testing and re-certification of products. (iii) Progress towards deregulation of the transport sectors of EU countries, including measures to reduce restrictions on truckers from one country accepting loads in another. (iv) Opening of public procurement in EU countries to effective competition from suppliers in all EU countries. (v) Deregulation of service sector activities, including opening financial services to competition and giving service providers and professionals the right of establishment in other EU countries. Estimates of the gains from these measures range up to as much as 5% of EU GPD, although these are based largely on computable equilibrium studies rather than an actual survey of outcomes (see Baldwin and Venables 1997).

2.2 Policy reform and commitment

Trade policy reforms – and other sorts of reform – are often hampered by the expectation that they may be reversed. Adjusting to reform typically involves investments, but these investments will not be made unless investors are confident that the reform will persist. These problems are mitigated if a country has a ‘commitment mechanism’ guaranteeing that the reform will be durable, and membership of a RIA can, in some circumstances, provide such a mechanism (Fernandez and Portes 1998).

The commitment mechanism operates most obviously for trade policy – membership requires that tariffs with member countries be cut, and reneging on agreed internal liberalization is likely to bring swift retaliation by partner countries. However, it has been argued that RIAs are valuable as commitment mechanisms for a much wider range of measures. Although NAFTA was ostensibly about trade policy, an important part of its motivation was the desire on the part of both the Mexican and US governments to lock in
the broad range of economic reforms that the Mexican government had undertaken in the preceding years. The EU Articles of Agreement with eastern European accession candidates are explicit in promoting ‘full integration into the community of democratic nations’. And the intervention of other Mercosur countries is credited with having averted a military coup in Paraguay in 1996 (Survey on Mercosur, *The Economist*, October 12, 1996). Paradoxically, it is even suggested that the value of a RIA as a commitment mechanism is greatest in areas other than trade policy, because there is already a way committing to tariff reductions – the tariff bindings of the GATT/WTO.

3. **Regional integration agreements and the world trading system.**

The effect of regionalism on the world trading system as a whole has been the subject of extensive recent debate. Is the growth of regionalism part of a process towards global free trade, or is it a substitute for it, damaging to multilateral negotiations and likely to lead to a situation of protected trading blocs? It is noteworthy that RIAs go against the fundamental principle of the GATT/WTO, the ‘most favored nation’ principle under which trade policy must be non-discriminatory. They are permitted only via Article XXIV of the GATT, which requires that they cover ‘substantially all trade’ of member countries and lead to external tariffs that ‘shall not on the whole be higher or more restrictive’.

There are essentially three issues. First, does a given structure of RIAs create a force for more or less liberal external trade policy? Second, is there an inherent dynamic of RIA formation which will lead to ever larger RIAs – perhaps even a RIA of the whole world, giving global free trade? And third, is there evidence that the presence of RIAs has assisted or retarded progress in the multi-lateral trade system?

An insight into the effect of the number of RIAs on incentives to set external tariffs was provided by Krugman (1993) who noted that if tariffs are set to improve member countries’ terms of trade, then they will be lowest -- and consequently world income greatest -- in two opposite circumstances. One is when there is a single world trade bloc containing all countries, i.e. global free trade. The other is when trade policy is set by many small independent jurisdictions, each so small as to have no market power.
and no reason to deviate from free trade. Between these extremes each trading bloc has 
an incentive to use external tariffs to try and improve its terms of trade, suggesting that a 
situation of relatively few large trading blocs might be the worst possible outcome.

This line of argument has been criticized from a number of different angles. From the theory side, it should be recognized that tariff setting is a repeated game, so the 
incentives to cooperate, or to deviate from cooperation, need to be taken into account 
(Bond and Syropoulos 1996, Winters 1998). From the institutional side, it has been 
pointed out that article XXIV of the GATT expressly forbids RIAs from raising external 
tariffs. And from the empirical side, no evidence has been found to suggest that RIAs are 
in fact prone to set higher external tariffs than are separate countries (Foroutan 1998).

What of the dynamics of RIA formation? Does adding a member country to a 
RIA increase the incentives for further countries to join, and will existing members be 
willing to permit unrestricted entry? A number of researchers have argued that the 
incentives to join increase as RIAs become larger – a phenomenon termed ‘domino 
regionalism’ by Baldwin (1995). This may be because of perception of growing benefits 
of membership, or because of increasing costs of being outside. These costs can arise as 
outsiders terms of trade may decline (the other side of the members improving terms of 
trade referred to in section 1.1). Perhaps more importantly, countries fear that firms may 
relocate, in search of the benefits of a larger market (evidence of foreign direct 
investment flows in Europe at the time of the Single Market Program suggests that this 
occurred, Baldwin, Forslid and Haaland, 1996). Another source of loss from non-
membership of RIAs is the risk of being isolated if a trade war occurs (Whalley 1996).

Although these arguments suggest a growing demand for membership of RIAs, it 
is less clear that this will be matched by a willingness to accommodate new members, at 
least within existing RIAs. The most important trade-off here is between the benefits of 
‘deeper’ integration, and the difficulty of achieving this with larger memberships. Thus, 
while the EU has continued to enlarge, worries about the difficulty of accommodating 
new entrants impedes progress. An important contribution to the ideas on regionalism 
was made when the Asia Pacific Economic Cooperation (APEC) forum announced the 
principle of ‘open regionalism’. Although this term has been given a number of quite
different meanings, the key idea is that of open access, whereby the RIA announces that any country willing to abide by its rules may join. However, at time of writing, the idea remains to be put to the test, since APEC itself has yet to develop as a RIA. Overall, then, it seems that regional integration does create its own dynamic, attracting further countries to want to join. But there is no reason to believe that all such requests should be successful, or that the end of the process is more likely to be a single world free trade area than it is to be a number of competing trade blocs.

Finally, are there any reasons to believe that the presence of RIAs has facilitated or impeded progress in multilateral trade negotiations? The argument for impedance hinges mainly on finite government capacity. Investment of time and effort in regionalism reduces the capacity of governments to invest in multilateral negotiation. On the other side, it is argued that regionalism can help, by reducing the number of separate negotiators (since members of a customs union negotiate as a single body) -- an argument that begs the question of how members formulate their common position. Also, it is suggested that regionalism has proved valuable by providing ‘laboratories’ for trying new approaches to issues of trade reform, particularly on more difficult issues. As for the track record of RIAs and multilateral negotiations, views are again divided, although there is a view that fear of fragmentation into regional blocs provided a spur for successful completion of the Uruguay round of trade negotiations (WTO 1995).


Joining a RIA is a major ‘one-off’ event for a country. It affects all aspects of the economy – goods prices, the structure of production, and income flows to different economic agents. Generalizing about the implications of such occasional and large changes is inevitably difficult. Nevertheless, research has established the circumstances under which gains or losses are more or less likely, and drawn out the implications of this for the design of agreements (World Bank 2000). Future research needs to pay greater attention to strategies for developing countries, and the role of RIAs in assisting them to participate effectively in the world trading system.
References


Krugman, P. 1993 ‘Regionalism versus multilateralism; analytical notes’ in New dimensions in regional integration, eds J. de Melo and A. Panagariya, CUP, Cambridge.


Anthony J. Venables, London School of Economics