Regionalism: Old and New, Theory and Practice

Mary E. Burfisher (Economic Research Service, USDA), Sherman Robinson (IFPRI) and Karen Thierfelder (U.S. Naval Academy)
Regionalism: Old and New, Theory and Practice

Mary E. Burfisher
Sherman Robinson
Karen Thierfelder¹

May 2003

Paper prepared for presentation to
The International Agricultural Trade Research Consortium (IATRC) Conference
Capri, Italy
June 2003

¹ Mary E. Burfisher is a Senior Economist at the Economic Research Service, USDA. Sherman Robinson is an Institute Fellow at the International Food Policy Research Institute. Karen Thierfelder is a Professor at the U.S. Naval Academy. The authors would like to thank Mary Bohman, Praveen Dixit, Moataz El Said, Mark Gehlhar, and Steven Zahniser for helpful comments on an earlier draft.
Abstract

In this paper, we review the debate on “new regionalism,” focusing on the tools used to evaluate regional trade agreements (RTAs). We find that much analysis uses tools from old trade theory in the Viner-Meade tradition, focusing on trade creation, trade diversion, and terms-of-trade effects. These tools are adequate for the analysis of the effects of removing commodity trade barriers (“shallow” integration), but the comfortable Viner-Meade framework misses many of the impacts associated with new regionalism, which typically involves “deep integration,” often between developing and developed countries. A framework for analyzing new regionalism should include dynamic changes such as trade-productivity links and endogenous growth theory, international factor mobility, the role of imperfect competition, rent seeking behavior, and political-economy considerations such as potential conflicts between regionalism and multilateralism. Agriculture poses problems for new regionalism because of high tariffs, the use of domestic subsidies and entrenched special interest groups, but the role of trade liberalization on its productivity is often overlooked. For developing countries, a crucial issue is whether and how regionalism can be part of a successful development strategy. While “new trade theory” is concerned with a number of the issues relevant to new regionalism, and is providing new tools, the work is eclectic and is far from providing a unified framework for empirical analysis of new regionalism. Both theoretical and empirical research is needed to improve the reach and scope of new trade theory applied to issues of new regionalism.
Introduction

The world economy after World War II has become much more integrated. Eight successive rounds of negotiations under the General Agreement on Tariffs and Trade (GATT) have resulted in significant global trade liberalization and there has been an accelerating trend toward regional integration in every part of the world. Most of the early attempts at regional trade agreements (RTAs) in the 1950’s and 1960’s, many of them among developing countries, met with little success.\(^1\) This “first wave” of regionalism has been eclipsed by the exponential growth in the number of RTAs formed over the past 10 years (figure 1). As of May 2003, 184 RTAs were in force. Almost every WTO member has now joined at least one RTA and some have entered 20 or more.\(^2\) The most dramatic policy-driven exercise in regional integration has been the establishment of the European Common Market in 1958 and its evolution into the European Union (EU).

\[\text{Fig. 1- RTAs in force by year of entry into force}\]

\[\begin{array}{c}
\text{1950} & 0 \\
\text{1955} & 20 \\
\text{1960} & 40 \\
\text{1965} & 60 \\
\text{1970} & 80 \\
\text{1975} & 100 \\
\text{1980} & 120 \\
\text{1985} & 140 \\
\text{1990} & 160 \\
\text{1995} & 180 \\
\text{2000} & 200 \\
\end{array}\]

Source: World Trade Organization.

In the U.S., Special Trade Representative Zoellick has described the U.S. pursuit of regionalism as a strategy to achieve short-term economic goals, help break the logjam in the multilateral negotiations, and achieve longer term, strategic objectives that can be fostered by trade liberalization.\(^3\) The EU has pursued regionalism aggressively as a means of encouraging investment and competition, and to

\(^1\) We will use the term “regional trade agreement” to include preferential trade agreements between countries, including those between countries not geographically contiguous or even nearby.

\(^2\) Facts about RTAs are available and regularly updated by the World Trade Organization (WTO) at its website: http://www.wto.org/english/tratop_e/region_e/region_e.htm. See also World Trade Organization (2002).

\(^3\) The U.S. has also established criteria for deciding which partners to engage in free trade agreements (FTAs). These include the size and importance of the economy to the U.S., the country’s willingness to negotiate a comprehensive agreement that includes topics such as intellectual property protections, and whether the RTA will help advance WTO or FTAA (Free Trade Agreement of the Americas) negotiations (Inside U.S. Trade, January 10, 2003).
reinforce a multipolarity in the international system (Lamy, 2001 and 2002). Even Japan, Korea, and China are now engaged in regionalism—with their first agreements signed at the end of 2002.4

These trends have led to a number of questions and research challenges for trade economists:

- What are the empirical characteristics of these new RTAs that distinguish them from earlier RTAs?
- Does the existing body of “old” trade theory, based largely on the theory of comparative advantage, provide an adequate framework for analyzing new regionalism?
- Does recent work on “new” trade theory provide a better framework for analyzing new regionalism?
- What are the major knowledge gaps, both empirical and theoretical, that need to be addressed for better analysis of new regionalism?

The objective of this paper is to review the major elements of the economics debate on new regionalism. First, we provide a brief overview of the characteristics of old and new regionalism. We then describe recent trends in the types of RTAs being formed, focusing on delineating the elements of deep integration and the links between developed and developing countries that represent the main distinctions between new and old regionalism. We also discuss the practical challenges to regional integration, particularly the inclusion of sensitive sectors like agriculture. We then review the use of old trade theory to evaluate new regionalism and empirical work in this tradition. Next, we consider studies using new trade theory to analyze new regionalism. We conclude that progress in analyzing current and potential RTAs must move beyond the limited framework of old trade theory. Finally, we identify some important knowledge gaps, both theoretical and empirical, that should be the focus of future work.

Old and New Regionalism

Historically, one can distinguish differing degrees of “integration” among countries, along a continuum from “shallow” to “deep.” Shallow integration involves only reducing or eliminating barriers to trade in commodities.5 Deep integration involves additional elements of harmonizing national policies, and allowing or encouraging internal factor mobility. The first 20 – 30 years after World War II can be seen as a period characterized by shallow integration, both globally and, where tried, regionally (“old regionalism”). With the Uruguay Round of GATT negotiations, and perhaps earlier in some areas, there were strong elements of deeper integration, going beyond commodity trade and standard border protection policies. Many RTAs introduced in the past 10 – 15 years have involved elements of deeper integration, and many of them have

4 Japan signed an agreement with Singapore in November 2002, and is now negotiating agreements with Mexico, South Korea, the Philippines and Thailand. China signed its first agreement with ASEAN (Association of South East Asian Nations), while Korea’s first agreement was with Chile.
5 The notion of “commodities” is sometimes expanded to include non-factor services.
linked developing and developed countries—the twin characteristics of “new regionalism.”

The theoretical analysis of the impact of increased integration and trade liberalization has followed the historical trends. International trade theory has a long tradition of analysis of the impact of shallow integration—liberalization of barriers to commodity trade. This work program has resulted in an elegant and coherent body of general equilibrium trade theory that underlies most policy analysis of both global trade liberalization and regional integration, with broad consensus in the economics profession about the desirability of achieving free trade globally. Under global free trade, countries would reallocate factors of production to achieve structures of trade, production, and employment of primary factors consistent with their comparative advantage, with welfare gains arising from increased efficiency. There is a large body of theoretical and empirical work in this tradition, working within what is commonly called the Heckscher-Ohlin-Samuelson (HOS) theoretical framework.

Theoretical analysis of shallow regional trade agreements—old regionalism—is more complex, because it is inherently an exercise in “second best” analysis. Some distortions are eliminated (tariffs on trade within the RTA), while others remain (e.g., other within-RTA domestic tax/subsidy policies and tariffs on non-RTA trade). The core theoretical analysis of shallow RTAs is the theory of customs unions, with seminal contributions by Viner (1950), Meade (1955), and Kemp and Wan (1976). In this framework, which adheres closely to the standard general equilibrium trade theory in the HOS framework, the welfare impact of an RTA is determined by a few crucial variables: changes in commodity trade in the countries within the RTA (“trade creation” effects), changes in trade between the RTA and the rest of the world (“trade diversion” effects), and changes in international prices facing the countries (“terms-of-trade” effects). In general, trade creation and terms-of-trade gains are welfare enhancing, and trade diversion and terms-of-trade losses are potentially damaging. Given the second-best environment, it is impossible to draw strong general conclusions about the desirability of forming an RTA from the perspective either of members or of the rest of the world, from theory alone. There are tradeoffs, and empirical work is required to make any welfare judgments in particular cases.

While there is a general view that there are elements, in addition to resource reallocation, that are an important part of the story in new regionalism, there is still widespread use of old trade theory to analyze RTAs, focusing only on trade creation, trade diversion, and terms-of-trade changes. The HOS and Viner-Meade frameworks are well established, representing a kind of conventional wisdom and coherent theoretical structure that is comfortable to use, even though it is widely understood that this framework misses much of the action in the new regionalism.\footnote{For a recent survey of theoretical work largely in the Viner-Meade framework, see Panagaryia (2000).}

\footnote{In two recent examples, Galal and Lawrence (2003) analyze the potential Morocco-U.S. and Egypt-U.S. free trade areas, and Leith and Whalley (2003) consider a potential South Africa-U.S. free trade area. Both studies discuss issues of deep integration and note that their effects may be large, but focus their empirical analysis on issues of trade creation and trade diversion, acknowledging the difficulty of measuring the effects of deep integration. Both papers work within the framework of old trade theory, with little reference to any theoretical work in new trade theory.}
Along with the emergence of deeper integration both globally and regionally, and the reliance on RTAs by developing countries, there is now emerging a body of “new trade theory” that has sought to incorporate the impact of forces that go beyond efficiency gains from reallocating resources according to comparative advantage. This work has been stimulated in part by the observation that, while efficiency gains from various regional schemes are significant, they are small in relation to national product and appear to be much too small to explain the rapid economic growth that has accompanied trade expansion in many countries. This body of work is much more eclectic and less coherent than work in the HOS and Viner-Meade frameworks, although there are certainly many examples of “elegant” models in new trade theory. There are partial and general equilibrium models incorporating a variety of new elements, including, for example, rent seeking, political economy, game theory, industrial organization (especially imperfect competition), geography, open-economy macroeconomics, and new growth theory. An important strand of this research agenda analyzes the links between international trade and total factor productivity, which provides additional sources of growth and welfare gains from expanded trade. There is an active literature seeking to understand the links between productivity and trade, especially in an environment with various elements of deep integration, and to measure their quantitative importance.

**Trends in “New Regionalism”**

**Diversity of RTAs: From Shallow to Deep Integration**

“Regional trade agreement” is a general term that refers to a whole spectrum of levels of economic integration. The lowest level of integration is represented by trade preferences, or partial scope agreements, which liberalize trade in specific commodities or sectors. This type of agreement, with its selective liberalization, does not conform to GATT/WTO rules on RTAs, which under Article 24 require that preferential arrangements for trade in goods meet two criteria:

- “substantially all trade” must be included, and
- the “general incidence of duties and other regulations of commerce” must not on the whole be higher or more restrictive against third parties than before the formation of the RTA.

The reasoning behind the requirement for substantial inclusion is to prevent members from taking a mercantilist approach to liberalization, and including only those sectors in which they anticipate export growth. Such agreements would likely be crafted

---

8 RTAs are also covered by GATT Article 5, which governs the preferential liberalization of services, and by the 1979 Enabling Clause, which governs trade in goods among developing countries. Of the 184 RTAs notified to the WTO as of May 2003, 25 RTAs were notified under Article 5 and 19 were notified under the Enabling Clause.
to carve export growth out of non-members’ market shares, and hence be essentially protectionist and trade-diverting in character. \(^9\)

The most common type of RTA is a free trade area in which members liberalize internal trade but retain their independent external tariffs. Seventy percent of the RTAs that have been notified to the WTO, are free trade agreements. Examples of free trade agreements include NAFTA, and U.S. agreements with Israel, Jordan, Singapore, and Chile. Since free trade agreements allow members to retain different tariffs against the rest of the world, they must include detailed rules of origin (ROOs). ROOs prevent goods that enter the member country with the lower external tariff from being transshipped duty free to members with higher tariffs. ROOs require that some proportion of products traded within the free trade area be of domestic content. ROOs can become complex because they can specify domestic content thresholds on a commodity basis and can in themselves become a focus of market access negotiations.

The GATT/WTO does not place any discipline on the rules of origin used in free trade areas. These are being increasingly recognized as an insidious form of trade protection. By increasing the domestic content requirement, ROOs can increase demand for local inputs, and divert trade from lower-cost, non-member suppliers. Krueger (1995) has argued that special interest pressures on the content requirements in ROOs gives them the potential to be used as non-tariff barriers on imported intermediates, causing them to become an important but hidden source of trade diversion in RTAs. \(^10\)

Customs unions liberalize internal trade and its members adopt common external tariffs against the rest of the world, eliminating the need for ROOs. About 8 percent of the RTAs currently in force are customs unions, including MERCOSUR, the Andean Pact, and the Central American Common Market (CACM). This is the type of RTA at which the second criteria of GATT article 24 is aimed. The prohibition against RTA’s raising their common external tariff is, like the first criteria, an attempt to minimize trade diversion. Low external tariffs reduce the margin of preference offered to pact members, and therefore the price incentives that lead to trade diversion. Kemp and Wan (1976) showed that it is possible to eliminate trade diversion entirely if a customs union adopts a sufficiently low set of common external tariffs at the same time that they liberalize internal trade. \(^11\)

---

\(^9\) Bhagwati (1990) argues that there were other reasons for the GATT adoption of a standard of 100 percent preferences. This approach creates a single-nation characteristic and a quasi-national status to the RTA that can legitimate the exception to the MFN principal. Such a difficult standard may also have been expected to act as a deterrent to an eruption of bilateral arrangements like those that created protectionism in the inter-war period. Finally, this standard may have been used to define regionalism as a supplement to, rather than a substitute for, multilateralism by establishing as a standard the ultimate goal of the GATT – free trade.

\(^10\) Analysis of trade data seems to support the negative views on ROOs. In a review of textile trade in NAFTA, Burfisher et al. (2001) and James and Umemoto (1999) both found strong evidence that NAFTA ROOs led to trade diversion. In the EU, Brenton and Manchin (2003) found a low level of utilization of EU trade preferences, which they attributed to ROOs.

\(^11\) MERCOSUR is an example of an RTA that simultaneously lowered its external tariffs when internal trade barriers were removed. Analyses of MERCOSUR related to agriculture show that the RTA therefore created trade for both members and nonmembers (Gelhar (1998), Zahniser et al. (2002)). Yeats (1998) found that MERCOSUR is net trade-diverting. However, his analysis is based on a partial-equilibrium
In a common market, members move beyond a customs union, and beyond shallow integration or commodity trade reforms, to allow the free movement of labor and capital within the union. The European Economic Community (EEC) by the early 1990’s had achieved a common market. With the decision to become the European Union, in which members adopted compatible fiscal and monetary policies, and (many) a common currency (the Euro), the Europeans are achieving full economic or deep integration, or an economic union.

New regionalism can be characterized as involving many of the elements found in the deepest level of integration, or the achievement of full economic (and monetary) union, and may include (in rough order of increasing depth):

- facilitating financial and foreign direct investment flows (real and financial capital mobility) by establishing investment protocols and protections;
- liberalizing movement of labor within the RTA;
- harmonizing domestic tax and subsidy policies, especially those that affect production and trade incentives;
- harmonizing macro policies, including fiscal and monetary policy, to achieve a stable macroeconomic environment within the RTA, including coordinated exchange rate policy;
- establishing institutions to manage and facilitate integration (e.g., regional development funds, institutions to set standards, dispute resolution mechanisms);
- improvements of communications and transportation infrastructure to facilitate increased trade and factor mobility;
- harmonizing legal regulation of product and factor markets (e.g., anti-trust law, commercial law, labor relations, financial institutions); and
- monetary union—establishment of a common currency and completely integrated monetary and exchange rate policy.

New Regionalism and Developed-Developing Country Linkages

New regionalism is also characterized by the linkages that recent RTAs have created between developing countries and one or more large, developed country partners (table 1). Such RTAs have been important for the United States, Canada and the EU, but not for Japan, Australia or New Zealand. Of the four RTAs in which the United States participates (currently in force and notified to the WTO), 3 are with developing countries: NAFTA (Mexico), Chile and Jordan. The EU has used RTAs as a key part of its strategy for economic development assistance. Thirty-five of its 50 RTAs are with developing country partners, some of which include Eastern European partners who have now graduated to EU membership. Especially when considering the impact of an RTA on developing countries, an important issue in new regionalism, elements of deep integration with developed countries are considered to be crucial in achieving potential links between

---

study of individual sectors, excluding agriculture, and so cannot yield conclusions on the aggregate impact of the RTA, which requires an economywide analysis. Using a CGE framework, Robinson et al. (1998) found MERCOSUR to be net trade-creating and welfare enhancing.
Increased trade and improvements in productivity. The nature of such links, and questions about how different elements of deep integration facilitate or hinder them, are a central part of much new trade theory.

Non-reciprocal preferences are agreements between developed and developing countries that typically allow for low or duty-free access for developing-country exports. Although not considered to be RTAs because they are not mutual, they deserve mention for two reasons. First, they are the most geographically comprehensive type of tariff preference in world trade. They include the Generalized System of Preferences (GSP) offered by developed countries to almost all developing countries. In addition, there are special non-reciprocal preference programs such as the U.S. Caribbean Basin Economic Recovery Act (CBERA) and Andean preferences; the EU Africa, Caribbean, and Pacific (ACP) Partnership, the Cotonou Agreement, and the “Everything But Arms” initiative. A disadvantage of these schemes compared to RTAs is that they generally require renewal and can expire, which happened in the U.S. during the recent debate in Congress on granting the president Trade Promotion Authority.

Non-reciprocal preferences are also relevant to the regionalism debate because their value to developing countries is reduced by regional integration, which can erode margins of preference within the union. This outcome has spurred debate about the costs to some developing countries of advancing from existing non-reciprocal preferences to regional trade agreements with developed countries, as in the proposed FTAA, for example. On one hand, developing countries will lose any benefits accrued in its favor due to the trade-diverting effects of the non-reciprocal preferences. On the other hand,

<table>
<thead>
<tr>
<th>Country or region</th>
<th>Number of regional trade agreements</th>
<th>Total in force</th>
<th>With developing country partners</th>
<th>With developing country partners, formed since 1989</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td></td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>European Union</td>
<td></td>
<td>50</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Japan</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Australia and New Zealand</td>
<td></td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>European Free Trade Area</td>
<td></td>
<td>24</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>Agreements between developing country partners</td>
<td></td>
<td>76</td>
<td>76</td>
<td>68</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>19</td>
<td>Na</td>
<td>Na</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>184</td>
<td>139</td>
<td>123</td>
</tr>
</tbody>
</table>

Note: NAFTA is included for both the U.S. and Canada.
such preferences are increasingly viewed as having offered few real benefits to developing countries. They often exclude the products of greatest export interest, such as sugar and dairy products. For example, Brown (1989) finds that exclusions of important export sectors of developing countries from preferences diminish the potential gains from preferences. Furthermore, they may actually slow economic development by reinforcing traditional or low-productivity sectors. Panagariya (2002) argues that EU preferences essentially transfer rents and undermine reforms that could promote faster growth. Lamy (2002) argues that non-reciprocal preferences lead to clustering of economic activity in traditional activities and slows industrialization. Because of this, the EU now plans to move away from this model, toward reciprocal preferences (RTAs) with developing countries.

**Sensitive Sectors and RTAs: Problems Posed by Agriculture**

Despite the GATT/WTO requirement that RTAs liberalize substantially all trade, full trade liberalization remains the exception rather than the rule. One reason this has been possible is the ambiguity of Article 24 criteria, which makes it difficult to determine whether RTAs are in compliance with the GATT/WTO. Whether the criteria refer to tariff lines or the value of current or potential trade, and what the threshold is for “substantial” have not been determined.\(^{12}\)

While the legal definitions remain unresolved, in practice, recent trade agreements have more extensive product coverage than previous agreements. Of the agreements in force in 1998, a recent WTO examination found that 43 percent had 100 percent coverage of industrial goods compared to 11 percent in the 1990 period, although few have 100 percent coverage. Agriculture continues to pose the greatest obstacle to comprehensive RTA coverage. Generally, RTAs’ inclusion of agriculture has been more limited than of industry, although it is has been addressed more comprehensively in recent RTAs than in earlier agreements.

---

\(^{12}\) Two measures of trade coverage are now being debated at the WTO to clarify the definition of substantially all trade. One is qualitative coverage, which would require that every major sector be included in trade liberalization. The alternative is a quantitative benchmark that would require that a certain percentage of trade be covered. The measurement of this benchmark might be the share of total trade, the share of bilateral trade, or a share of tariff lines.
Agriculture poses two challenges for regional agreements. First, tariffs in agriculture remain very high, despite the progress made in the Uruguay Round. These high tariffs signal the presence of effective special interest groups that have been able to forestall trade reform at both the global and the regional levels. Both the U.S. and the EU already allow preferential access to most of their agricultural trade partners: for both, about two-thirds of their agricultural imports come from countries benefiting from their trade preferences (figures 2 and 3). However, the value of the preferences is diminished by their exclusion of many agricultural commodities and related food products linked to their domestic support programs.\(^\text{13}\) The EU’s RTAs (with non-CAP members) exclude those agricultural commodities that receive domestic support under the CAP (Hasha, 2001). Similarly, U.S. GSP and other developing country preferences exclude commodities linked to U.S. domestic support programs, including processed food products such as confections with sugar and dairy content (Burfisher, Robinson and Thierfelder, 2002).

Second, some agricultural policies are essentially global in character, and are not amenable to regional reform. Domestic production subsidies, for example, have a global, rather than regional impact. In the FTAA, the U.S. has therefore argued that their negotiation in a regional forum would reduce needed U.S. leverage on their negotiation in the multilateral discussions. Likewise, the reform of export subsidies needs a global solution. While an RTA can discipline the use of export subsidies within a region, subsidies by third countries into the region would be difficult to monitor and the creation of regional compensation mechanisms to protect regional exporters would be a challenge.

One way that RTAs can resolve such incompatibilities in members’ domestic support or export subsidies is to harmonize or to adopt common policies, as the EU has done internally in its Common Agricultural Policy (CAP). Another is to provide long transition periods for certain commodities. In one of the only comparative empirical analyses of tariff phase-out periods, Estevadeordal (2002) studied the phase-out paths of trade agreements in the Western Hemisphere. He found that over 50 percent of products become duty-free in the first year of the RTAs’ implementation, and only about 5 percent

\(^{13}\) Crawford and Laird provide a general overview of the treatment of agriculture in RTAs notified to the WTO. Sheffield (1998) reviews the treatment of agriculture in major RTAs and finds that most include agriculture in trade liberalization. Estevadeordal reviews the treatment of agriculture in Western Hemisphere RTAs, finding that most will eliminate internal barriers for almost all products.
of trade is exempted. However, the phase-out periods differ significantly on a sectoral basis, with agriculture generally having the most gradual liberalization path.

WTO members intend to address some of the deficiencies of the GATT/WTO disciplines on RTAs in the Doha Development Agenda. The Doha Declaration’s Article 29 states that “We also agree to negotiations aimed at clarifying and improving disciplines and procedures applying to regional trade agreements.” When these issues are resolved, there will be stronger global trade rules that reinforce the benefits and constrain the protectionist character of RTAs. Ultimately, the most effective discipline on RTAs will be the continued pace of trade liberalization in the WTO. Further reductions in MFN tariffs will reduction of margins of preference available through RTAs, and stronger enforcement and dispute resolution mechanisms can help ensure fair trade despite the essentially uncompetitive nature of preferential tariffs in world markets.

Old Trade Theory and New Regionalism

Theoretical Arguments: Trade Creation, Trade Diversion, and Terms of Trade

There is a large literature using old trade theory to analyze the impact of RTAs in the Viner-Meade framework. In Annex 1, we review the key points of this framework, particularly how models in this tradition consider trade creation, trade diversion, and terms-of-trade effects. While some work in this tradition attempts to extend the Viner-Meade framework to capture some elements of new regionalism, all of it is characterized by its focus on trade creation, trade diversion, and prices, and by its neglect of potential trade-productivity links and other essential elements of new trade theory.

Bhagwati and Panagariya (1996), (also Panagariya 1996, 1999) oppose RTAs because of their trade diverting effects. Instead, they advocate multilateral reform. Their theoretical arguments rely on a special case in which an RTA is specified as purely trade-diverting. In this case, which they call the “small union,” the home country has no domestic production of the traded good (hence, no possibility of trade creation), and the RTA partner cannot supply all of the imports demanded. Some imports continue to come from the non-RTA partner, which is a large country and the price setter. By construction, there can be only a welfare loss as consumers in the importing country do not benefit from a lower import price and lose the tariff revenue on imports from the RTA partner. The larger the initial trade share with the RTA partner, the greater the welfare loss associated with the tariff revenue transfer. They also refer to this as a terms-of-trade loss because the importing country moves up the export supply curve of its RTA partner, worsening its terms of trade with the RTA partner (see discussion in the annex).

14 Article 29 in addition states that “(T)he negotiations shall take into account the developmental aspects of regional trade agreements.”
15 Old trade theory tools were appropriate to analyze old regionalism. For example, the RTAs in Latin America in the 1960s failed because trade diversion effects dominated, as theory predicted.
16 Others dismiss this case because it is not realistic. For example, de Melo et al. (1993) note that the case of pure trade diversion while unambiguously welfare-worsening, is too extreme a model to characterize actual RTAs. For a recent discussion of the theory with models that allow both trade creation and trade diversion, see Winters (1996) and DeRosa (1998).
One of the implications of their terms-of-trade argument is that RTAs such as NAFTA will be particularly harmful to undiversified partners such as Mexico. For example, Panagariya (1997) does a back-of-the-envelope calculation, estimating welfare losses as high as $3.26 billion for Mexico from NAFTA, looking only at tariff losses. In contrast, Brown (1993) argues from the same initial conditions for Mexico—high trade shares with the United States and high tariffs—but finds that little trade can be diverted from the rest of the world due to NAFTA. The difference between the two approaches is the underlying assumption about the analytical model. Brown presumes both trade creation and trade diversion are possible, while Panagariya’s model allows only for trade diversion.

The general conclusion of the theoretical literature in the Viner-Meade framework is that whether an RTA is welfare-increasing is essentially an empirical question that must be settled by examination of data specific to each RTA under consideration. The various theoretical models point to potentially important effects and causal channels, but no general conclusions can be drawn from theory alone.

**Empirical Evidence: Trade Creation, Trade Diversion, and Terms of Trade**

Although both trade volume and terms-of-trade effects determine the welfare impacts of RTAs in the Viner-Meade framework, empirically, it has been easier, ex post, to analyze trade creation and diversion, which can be measured with data on trade volumes, than to analyze terms-of-trade effects, which require price data. In a recent, comprehensive review of trade flows, Crawford and Laird (2001) analyze trade data from six regional trade agreements (APEC, MECOSUR, NAFTA, the EU, the Association of South East Asian Nations—ASEAN, and the Andean Community). They find that, between 1990 and 1999, trade within these pacts increased an average 7.1 percent annually, compared to the average annual growth in each pact’s imports from non-members of 6 percent annually. In the case of the EU, imports from members and nonmembers grew at the same rate. The high overall growth in the RTAs’ imports may reflect the effects of liberalization on their economic growth and demand, and suggests that the RTAs have been net trade-creating for both members and nonmembers.

The challenge, however, is to identify the impact that RTA policies have had in accounting for these trade flows. Gravity models are becoming an important tool for such analysis because they can be used to econometrically estimate the relationship between trade and a policy shock. But these models are also criticized because of their weak theoretical foundation, and because the estimated coefficients for the trade-agreement variables may capture the influence of unrelated developments that are contemporaneous to these accords. The growing body of research utilizing these models is not yielding a

---

17 Similarly, Panagariya (1996) projects losses for Latin America in a Free Trade Agreement of the Americas (FTAA). Since Latin American countries have higher tariffs than the USA, they will lose tariff revenue when they enter an FTAA. This is, of course, a theoretical argument and its empirical validity depends on the size of the effects. In the case of NAFTA, there is no empirical evidence that these tariff-diversion effects have been empirically important. See Burfisher et al. (2001).

18 See Anderson and van Wincoop (2003) for a recent discussion of the theoretical weakness of gravity models.
consensus view on RTAs’ trade impacts. Frankel et al. (1996) analyze trade among 63 countries between 1965 and 1992, finding intra-bloc trade biases that are statistically significant, with mixed impacts of trade creation and trade diversion with nonmembers. Soloaga and Winters (1999) find no evidence of increased intra-bloc trade during 1980-96, although in Latin America, they found RTAs had a positive impact on the imports of bloc members. In agriculture, Zahniser et al. (2002) estimate a modified gravity model that controls for the importing country’s long-term bilateral trading relationships and allows the impacts of an RTA to vary from one participant to another. They estimate the impacts of Western Hemisphere RTAs on U.S. agricultural exports; in contrast, the other gravity models discussed here analyzed total trade. They find that NAFTA had a positive and significant impact on Mexico’s trade. Mexico’s unilateral reforms accounted for 39 percent of U.S. agricultural exports to Mexico during 1989-93, while the reforms and NAFTA together account for 59 percent of this trade during 1994-99.

Multi-country computable general equilibrium (CGE) models include the price analysis necessary to evaluate terms-of-trade effects. In these models, commodities are differentiated by country of origin. The effects of an RTA on the demand for non-member goods depends on the elasticity of substitution between member and non-member country goods. CGE models allow controlled simulations of the effects of trade reforms on endogenous prices and quantities relative to some benchmark base year. Burfisher and Jones (1998) survey the results of CGE-based case studies related to agriculture, including analyses of NAFTA, Western Hemisphere integration, EU expansion and APEC. These studies report changes in the U.S. agricultural terms of trade, which improve in the short run for all RTAs except for the EU. The inclusion of both terms-of-trade effects and trade changes allow CGE models to generate welfare outcomes. The RTAs surveyed by Burfisher and Jones were all welfare-increasing for members and the world, but not for the U.S. when it is a non-member.

Schiff and Winters (2003) critique the use of CGE models to draw ex post conclusions because the models are used for counterfactual simulations, not forecasts, and because of their typically ad hoc estimates of behavioral and trade/productivity parameters. The counterargument is that: (1) sensitivity analysis indicates that the broad conclusions are robust to reasonable variation in parameter estimates, and (2) that CGE models provide the most appropriate tool for examining the impact of trade liberalization on world prices, trade, and welfare in the Viner-Meade framework. These empirical models have become a work-horse of policy analyses because of their capacity to capture bilateral trade flows, input-output relationships, factor market effects, price and quantity changes, and welfare impacts—all within a framework that has a consistent foundation in microeconomics and trade theory.

Schiff and Winters (2003) also argue that CGE models overstate the terms-of-trade benefits to RTA members because the models use the assumption that products are differentiated by country of origin, giving each country some degree of market power. This characterization is incorrect because it focuses only on the terms-of-trade gains members experience at the expense of non-members. One needs to consider the changes in both intra-union terms of trade and terms of trade with the non-union countries, which are both captured in CGE models.
The applied literature on the estimated welfare impacts of regionalism is large and growing, and for the most part supports a consensus view that RTAs have been net trade-creating and world welfare-improving. Baldwin and Venables’ (1995) review of the empirical literature found generally positive impacts on the living standards of RTA members and negligible impacts on nonmembers. Robinson and Thierfelder (2002) review the CGE-based literature and conclude that there are robust conclusions from the many existing studies of RTAs: (1) they increase welfare of participants, (2) aggregate trade creation is much larger than trade diversion, (3) positive welfare effects are even larger if features of new trade theory are considered, and (4) there are additional welfare gains from expanding membership.

**RTAs and Global Liberalization: Complements or Conflicts**

Bhagwati’s (1990) characterization of regional trade agreements as “building blocs or stumbling blocs” for multilateralism is part of an ongoing debate on whether new regionalism helps or hurts prospects for continuing global liberalization. In addition to RTAs’ static trade and welfare impacts, their effects on the multilateral negotiations then taking place in the Uruguay Round became an important criteria for evaluating the net benefits of an RTA. In Bhagwati and Panagariya (1996), building blocs were defined as RTAs that either prompt an acceleration in multilateral negotiations (by going further and faster than multilateralism, and creating successful experience with reform), or which add new members until the bloc converges on global free trade. Stumbling blocs are the opposite: they create or entrench trade diversion and protectionism, and are closed to expansion.

Much of the literature on RTAs as building or stumbling blocs is theoretical and stays in the Viner-Meade framework. One way an RTA can act as a stumbling bloc is by creating incentives for its members to raise their external tariffs against nonmembers, thereby increasing the overall level of protection in world markets. Although this is not allowed by GATT/WTO Article 24, it is at least conceivable that developing countries whose applied rates are well below bound rates could exercise this option since WTO rules allow countries to increase their tariffs up to their bound rates. If RTAs lock in bilateral rates, then any policy responses to protectionist pressures must fall on outside countries. Panagariya (1996) argued that this dynamic was evidenced during Mexico’s peso crisis, although most observers of NAFTA have concluded that the opposite occurred. Since about 90 percent of Mexico’s trade is with the U.S., NAFTA effectively prevented Mexico from a pursuing a protectionist response to the foreign exchange crisis, and helped Mexican policy makers to pursue a successful macro stabilization program. NAFTA, by locking in Mexico’s policy reforms of the 1980’s, was

---

19 See, for example, Bergsten (1997), who is an eloquent proponent of the notion of open regionalism, and Bond, Syropoulos, and Winters (2001), who consider issues of deep regional integration and multilateral liberalization.
20 Panagariya pointed to Mexico’s response to its peso crisis, when it increased tariffs on non-NAFTA countries on 502 products by 35 percent, about a 75 percent increase. Bhagwati (1993) has made a similar point, arguing that countries entering a FTA will tend to increase their use of non-tariff measures such as anti-dumping and countervailing duty actions against third countries as weaker industries struggle to survive regional free trade.
widely considered to have been a major factor in maintaining investor confidence and initiating a macro recovery.\textsuperscript{21}

Krugman (1991) explains the incentive small countries have to form an RTA that lets them exert market power, as a group, and benefit from a higher external tariff. This analysis uses the old trade theory tools from the Viner-Meade framework, particularly the terms-of-trade effects of an RTA on non-members. Krugman finds that the welfare-minimizing number of blocs is three, when tariffs are set non-cooperatively.\textsuperscript{22} Because the formation of a trade bloc creates trade among its members, the bloc now has more market power than that wielded by its individual members and therefore a greater incentive to levy higher tariffs and manipulate its terms of trade.\textsuperscript{23} Krugman describes a U-shape to world welfare according to the number of trade blocs. When there are many small blocs, they will set low external tariffs due to their limited market power. As the number of blocs diminishes, they will seek their optimal tariff and set their tariffs high, tending to increase trade diversion relative to trade creation. When there is only one bloc, the world has achieved free trade on a multilateral basis, and trade diversion is eliminated.

While terms-of-trade issues have an important place in theoretical models, there is general agreement that large countries or blocs in fact do not appear to be motivated by the exercise of market power, nor are they observed to be raising tariffs. More likely, they are motivated by social welfare goals or interest-group satisfaction as described by Bhagwati and Panagariya’s stumbling bloc theory. Krugman describes this as “GATT-think” in which countries are mainly acting to maximize exports, minimize imports, and trade off some import growth for export gains. Domestic producer interests dominate—perhaps due to Olsen-type lobbying and collective action.

In our view, the continuing reliance on old trade theory in the Viner-Meade framework to analyze the impact of new regionalism is not very useful and distracts attention from important issues related to these new RTAs. First, while old trade theory concludes that FTAs “may” reduce welfare, the empirical work in the narrow Viner-Meade framework indicates that RTAs are generally good for their members, that they are not seriously detrimental to nonmembers, and that global liberalization is always better. With few exceptions, there is no convincing empirical evidence that trade diversion dominates in the RTAs considered. These results support the “open regionalism” view of RTAs, in which their continued expansion supports an evolution toward global liberalization. A second robust conclusion from empirical work in this tradition, however, is that the potential benefits of trade liberalization in general, and RTAs in particular, are rather small as shares, say, of national product—neoclassical efficiency gains tend to yield triangles, not rectangles. Certainly, the dramatic increases

\textsuperscript{21} See, for example, De Long, De Long, and Robinson (1999).

\textsuperscript{22} Krugman (1993) later argued that the current multilateral system is one in which tariffs are set cooperatively. Countries are therefore likely to believe that they have more to lose from the collapse of a cooperative global system than they can gain from non-cooperative trade policies that exploit their market power.

\textsuperscript{23} Krugman (1993) later argued that trade diversion is likely to dominate trade creation even without an increase in a trade bloc’s external tariff, an argument that rests on the restrictive assumption of a very high elasticity of substitution between regional and extra-regional goods.
in world trade in the past 50 years, and the apparent large benefits arising from that increased trade, do not seem to be captured in standard neoclassical trade models. More is needed and it is time to move beyond the static HOS and Viner-Meade frameworks.

**New Trade Theory and New Regionalism**

While old trade theory focuses on commodity trade and prices, new trade theory considers a variety of other effects of trade and mechanisms other than more efficient sectoral allocation of factors of production. New trade theory considers trade-productivity links (i.e. “new growth theory”), imperfect competition, and rent-seeking behavior, especially in considering the issue of regionalism versus multilateralism. Characteristics of new regionalism suggest that their welfare impacts cannot be fully explained using old trade theory. Features of new regionalism that have been prominent in recent literature include:

- technology and knowledge transfers, and technology diffusion, especially from developed countries to developing countries, that increase productivity,
- dynamic comparative advantage and “learning by doing” efficiency gains through increased demand from expanded trade,
- elimination of wasteful rent seeking activities through trade liberalization,
- pro-competitive gains from increasing import competition in an environment of imperfect competition, allowing exploitation of potential economies of scale in production,
- increased geographical dispersion of production through trade that supports (1) exploitation of different factor proportions for parts of the production process (Ricardian efficiency gains) and/or (2) local economies of scale through finer specialization and division of labor in production (Smithian efficiency gains),
- increased foreign direct investment that carries with it advanced technologies and hence increases in productivity,
- “challenge-response” increases in efficiency through increased competition due to expanded involvement in world markets, and
- Schumpeterian innovation and “creative destruction” induced by increased competition arising from expanded trade.

We review three of these features below: trade-productivity links, imperfect competition, and rent-seeking behavior.

**Trade-Productivity Links and Development Strategies**

New regionalism provides a broad context to explain the motivations of developing countries to engage in regionalism As Ethier (1998a and 1998b) notes, recent RTAs typically involve more than commodity trade liberalization. His “stylized facts” of new regionalism are:
recent RTAs typically feature one or more developing countries linking up with a developed country—for example, the union of Mexico with the U.S. and Canada in NAFTA;

• membership in an RTA often follows significant unilateral liberalization by developing countries, including both trade and macro policy reforms—for example, the unilateral reforms by the Central European countries and Mexico that preceded their respective unions with Western Europe and NAFTA;

• RTAs seldom address only trade barriers—in fact, the degree of trade liberalization may be modest—and they invariably incorporate elements of deep economic integration; and

• developing countries make bigger trade concessions in RTAs; often because the developed countries have low tariffs to begin with.

Ethier’s focus on the role of new regionalism in achieving deep integration between developed and developing countries highlights their role as part of an appropriate “development strategy” for the poorer countries. There is a fairly broad, but not universal, consensus that expanded trade is an important part of a successful development strategy. There is also a consensus that trade liberalization is not sufficient to achieve improved performance, including rapid growth and elimination of poverty.

Developing countries hope that their internal reforms will attract foreign direct investment (FDI) from developed countries (“investment creation”), which carries with it the prospect for the transfer of global technology and increased productivity. The payoff of the RTA through FDI and the resulting unskilled wage gains must be great enough to overcome resistance from special interest groups. The potential FDI may even be large enough to induce countries to attempt reform that otherwise would not do so—a “reform creation” effect of an RTA.

Within the framework of new regionalism, developing countries form an RTA with a developed country to compete with nonmembers for the developed partner’s FDI. Ethier argues that the effect of this “investment diversion” is likely to increase the resolve of outsiders to reform themselves, acting as a positive force behind reform creation, in his view the major benefit of new regionalism.

The role of FDI and productivity growth in Ethier’s framework incorporates the endogenous growth theory that has become embedded in recent empirical work on RTAs. Typically, trade is assumed to have a role in stimulating productivity growth through channels that include technology differences among countries, knowledge spillovers, the transmission of ideas, and market expansion that lead to increasing returns to scale and/or Smithian economies of “fine specialization” (as opposed to Ricardian differences in factor proportions). In CGE analyses, the operational links between trade liberalization and total factor productivity growth are frequently based on the stylized trade-productivity externalities described by de Melo and Robinson (1992). There is an export externality link between export growth and an increase in TFP within the sector. On the

---

24 See Schiff and Winters (2003) for a more detailed discussion of whether and how RTAs support economic development. They have an extensive discussion of the characteristics of new regionalism, especially as they relate to developing countries.
import side, imports of intermediate and capital goods are linked with sectoral TFP. Finally, an increase in aggregate exports leads to economy-wide increases in the efficiency of capital inputs.

Although this modeling approach typically incorporates ad hoc assumptions about the parameters that describe trade-productivity links, there is a growing body of empirical literature that seeks to measure links between trade volumes and productivity. Coe, Helpman, and Hoffmaister (1997) estimated trade-productivity links for 77 developing countries, finding sizable spillover benefits of research and development in developing countries through exports of machinery and equipment to developed countries. They estimated that a one-percent increase in the import share of machinery and equipment to GDP results in a .3 percent increase in TFP. Frankel and Romer (1999) analyzed a 98-country sample, controlling for capital inputs per worker and schooling. They found that a one-percentage point increase in the trade share of GDP increased the contribution of productivity to output by about two-percentage points.

There remain many skeptics of the role of trade or openness per se in stimulating growth. Rodriguez and Rodrik (1999), for example, argue that the positive links between openness and income growth are greatly overstated and that the empirical work is suspect given the mixed quality of the data and problems related to measurement and empirical methodology. Furthermore, most of the trade externalities are based on macro relationships between measures of openness and measures of income or productivity growth. Instead of openness and trade expansion, Rodrik et al. (2002) argue for the primacy of institutions in explaining economic growth. They find that the effect of trade on income, after controlling for institutions and geography, is almost always insignificant, although it is positively related to effective institutions.

Trade-productivity externalities have been a defining element of the new trade theory and regionalism, incorporating elements of new growth theory, and related empirical literature. Open questions remain for theoretical analysis of new regionalism. In addition to the need for stronger evidence concerning trade-productivity linkages in general, there is the potential for exploring and testing the applicability of new growth theory to agriculture. Analysis of the role of institutions, and how they determine the manner links are exploited, could make an important contribution to the literature on regionalism. The role of trade rules and institutions is another area that has yet to be fully incorporated into analysis of regionalism. In particular, the question remains open as to whether there remains room in the multilateral system for countries to exploit productivity links as we find out more about them, or perhaps whether future Asian tigers may lose critical opportunities because of current constraints on policies such as export subsidies.

---

26 There is a parallel literature on micro studies of the impact of openness on productivity, using firm-level data. See, for example, the survey by Evenson and Westphal (1995). There is also some work at the sectoral level that argues for links between sectoral TFP growth and trade performance (e.g., Nishimizu and Robinson, 1984).
27 Of course, the emergence of good “institutions” might well be related to openness and is a potential source of productivity growth—TFP potentially includes much more than technology.
Trade Productivity Links and Agriculture

Trade-productivity externalities are typically linked to manufacturing, although economy-wide capital productivity gains can benefit agriculture as well. In addition, agricultural exports can serve as a source of foreign exchange to finance increased imports of productivity-enhancing intermediate inputs and capital goods. In general, though, there is a gap in this literature regarding the potential for trade expansion to lead to structural and/or technological change in agriculture, which could well be very important for the poorest developing countries.

Some recent literature at the industry level suggests this potential. Haley (2003) has described structural change in the North American hog and poultry sectors following NAFTA. She argues that open borders are among the factors that have contributed to the evolution of the U.S. and Canadian hog industries into a single, integrated industry in which each country specializes in stages of hog production that best meet its economic resources and regulatory framework. These changes have helped the industry to reduce production costs and increase product quality. USDA (2002) analyzed the effects of North American integration on the retail and wholesale food supply chains. The increased volume of trade and the development of strategic alliances among North American firms are spurring a “virtuous cycle” of continued innovation and integration that is lowering costs, improving quality and achieving year-round supplies of produce. Reardon and Berdegeu (2002) summarize a series of studies on the rise of supermarkets in Latin America. They conclude that trade liberalization has served as a supply side factor contributing to the transformation of the food marketing chain. Liberalization has made it easier and cheaper to import food and non-food products (such as refrigerators), which leads to what they call economies of scope (the possibility for large stores to hold more inventory than small stores). In addition, drastic change in FDI regulations, some of them due to RTAs, has contributed to a surge of FDI into the retail food sector and the rapid growth in the presence of multinational grocery chains

Imperfect Competition

Imperfect competition, game theory, and product differentiation are additional components of new trade theory that have been included in analysis of new regionalism. CGE models incorporating imperfect competition and economies of scale were developed to analyze the impact of NAFTA (Francois and Shiells, 1994). Winters (1997) develops a model with imperfect competition to analyze terms-of-trade effects of an RTA, particularly on non-RTA members. He argues that it is incorrect to compute changes in exports from the non-member countries to RTA countries and draw welfare conclusions. “ROW exports are a very poor indicator of ROW welfare.” (p. 134). He develops a theoretical model to analyze the welfare impacts of an RTA on non-members and argues that they depend on changes in its terms of trade, levels of output, number of firms, existing trade restrictions and induced investment effects. His proposed analysis of price changes relies on new trade theory tools of imperfect competition and product differentiation.
Winters and Chang (2000) and Chang and Winters (2002) apply Winters’ (1997) theoretical model and examine the terms-of-trade effects of an RTA when there is Bertrand competition between the RTA partner and the non-partner. They use detailed price data and estimate the price effects when an RTA reduces tariffs faced by one exporter. Winters and Chang (2000) consider Spain’s accession to the EU. Analyzing data on the price of Spanish imports of finished manufactured goods from major OECD sources, they find that the preferred exporter will raise its pre-tariff price while the non-member will reduce its pre-tariff price. For example, they find that a 1% fall in the tariff on EC sales to Spain could reduce U.S. export prices to Spain by 0.34%. Using similar techniques to analyze the terms-of-trade effects of Brazil’s entry to MERCUSOR, Chang and Winters (2002) find that non-members’ export prices to Brazil fell relative to their export prices of the same commodities to other markets. As nonmember countries, including the U.S., Chile, and European suppliers, were forced to compete with preferred MECOSUR exporters to Brazil, they experienced a terms of trade decline. For the U.S., this relationship was small, but statistically significant, for agricultural commodities.

**Lobbying and Special Interests**

Perhaps the broadest framework to explain the role of interest groups in shaping regionalism is the analytical model developed by Grossman and Helpman (1995). Building on Grossman and Helpman (1994) which explains policy formation as the outcome of lobbying and contribution competition among industries, they describe free trade negotiations as a process of providing a sufficient balance among countries’ interest groups. Free trade areas are devised in order to ensure that each country will have a sufficient number of exporters that will benefit from the agreement and provide it with their political support. Grossman and Helpman argue that the need for political viability may ultimately contradict the social desirability of the agreement. Specifically, when export growth results from trade diversion, so that narrow producer interests can be served by the agreement at the expense of the diffuse group of taxpayers, the agreement will be welfare-reducing. Trade diversion enhances political viability, but makes it more likely that countries will craft trade-diverting agreements.

In an application of the Grossman-Helpman (1994) model to the case of foreign lobbying, Kee et al. (2003) study the relationship between foreign lobbying and the GSP tariff preferences that the United States grants to Latin American countries. His results show very high returns to Latin American exporters’ political contributions of above 50 percent. The reason for these large returns is due to the relatively low weight given to social welfare in the U.S. government’s objective function relative to Latin American exporters’ lobbying contributions for tariff preferences.

Baldwin’s (1997) domino theory uses the role of interest groups in nonmember countries to make a positive case for regionalism. He argues that trade and investment diversion within a trade pact will generate political economy forces in excluded nations to either accede to the pact—if its membership is open—or to create new trade pacts among themselves. This pressure will increase with the size of the trade pact. The momentum is created by asymmetric lobbying—losers tend to lobby harder than winners. Exclusion from trade blocs will therefore tend to strengthen the hand of pro-liberalization
forces that are adversely affected by trade or investment diversion. This sets in motion a process that causes bilateral trade barriers to fall like dominoes, regardless of the pace of progress in multilateral negotiations.  

Conclusion

The historical transition from old to new regionalism—a transition from shallow to deep integration, and the partnering of developed and developing countries—has been accompanied by developments in economic theory and empirical work in international trade. Old trade theory, based on the elegant Heckscher-Ohlin-Samuelson (HOS) and Viner-Meade theoretical frameworks, provided a powerful set of tools for analyzing issues arising from both global trade liberalization and the formation of regional trade agreements involving liberalizing commodity trade within the RTA (e.g., a customs union or free trade agreement). As both global liberalization and regional trade agreements moved beyond commodity trade to incorporate additional elements—deep versus shallow integration—new theoretical analysis also followed. New trade theory, however, is much more eclecti

There is a significant body of work using the methods of old trade theory to analyze the impact of new regionalism. The old paradigm is well developed, well understood, and comfortable, providing a body of conventional wisdom that facilitates analysis. Much of this work, however, is unsatisfactory, focusing on a narrow range of forces at work and missing a lot of the action arising from integration that goes well beyond commodity trade flows. It is time to move beyond this work and incorporate elements of new trade theory in empirical and theoretical analysis of new RTAs.

The state of knowledge concerning new regionalism is certainly in flux. There are many important hypotheses that are as yet highly tentative, calling for both theoretical and empirical work. A partial list includes:

- Given that new regionalism usually involves integrating developed and developing countries, what are the links between the formation of RTAs and successful development strategies in the developing countries?
- How do rules of origin in FTAs affect future multilateral reforms? Do they emphasize (exaggerate) the trade diversion effects of an FTA? Are they manageable administratively?
- What is the nature of trade productivity links, in both developed and developing countries? Are there differences in the nature of these links in developed and developing countries?

28 Schiff and Winters (2003) have a more critical view of the domino theory and argue that it is unlikely to lead all the way to global free trade. They cite incentives for RTA members to limit entry and exploit terms-of-trade benefits against outsiders. They also summarize the empirical evidence on expansion in RTAs and find that the number of RTAs that have accepted new members is approximately equal to the number of RTAs that have not expanded.
• Is there evidence of Smithian gains at the micro level? Do we see finer specialization in production following the formation of an RTA? To what extent does an RTA (which guarantees access to partner markets) make finer specialization in production more feasible?

• Does deeper integration among partners contribute to productivity gains? At the micro level, is there more harmonization in production? Are there changes in FDI following the creation of RTAs?

• Finally, an issue much studied but not yet resolved, is the extent to which the formation of RTAs impedes or supports continued global trade liberalization.

References


Annex: Analytical Tools in the Viner-Meade Framework

Trade creation, trade diversion and the terms-of-trade effects largely define the static welfare impacts of an RTA. Viner (1950) developed these concepts within a “second best” theoretical framework in which countries form a customs union. Much of the welfare analysis also pertains to the formation of RTAs. Viner described the changes in the allocation of production and consumption following the union. To focus on trade-creating and trade-diverting effects, consider first the case in which the RTA member, country A, is small relative to its RTA partner, country B, and the nonmember, country C—the export supply curves from the partner and the rest of the world are perfectly elastic. Trade creation occurs when A’s imports increase following the formation of an RTA. It corresponds to the production and consumption gains from additional imports at a lower domestic price—A imports more from its lower-cost RTA partner, B, while its own higher-cost domestic production declines.

Trade diversion occurs when RTA members’ imports from partner countries replace imports from more efficient non-member countries, as a result of the RTA’s tariff preferences. An RTA improves welfare for its members if the benefits of trade creation dominate the losses from trade diversion. In a partial equilibrium framework, the net welfare effect depends on the magnitude of the gains in consumer surplus resulting from a lower internal price and the loss of tariff revenue.

One can construct the case in which an RTA is purely trade-diverting. Described as the “small union case,” assume that both partners are small relative to the rest of the world. Assume that Country A initially imports a good from both B, which has an inelastic export supply curve, and Country C, which is large and has an elastic export supply curve to Country A. Suppose Country A forms an RTA with Country B, and that Country B cannot supply all of Country A’s import demand following the RTA. In this case, the domestic price of imports in Country A does not change; it is still at the price set by Country C. However, the share of imports changes as Country A buys more imports from Country B. Country B moves up its export supply curve until its tariff-free supply price in Country A equals the tariff-ridden price from Country C. Country A is worse off with the RTA, since its consumers still pay the same domestic price, and tariff revenue has been transferred to the partner through Country A’s terms-of-trade loss against Country B. The larger the initial share of Country A’s trade with its partner, the greater the welfare loss associated with the tariff revenue transfer.²⁹

When countries are large enough to affect world market prices as a result of changes in domestic trade policies, there are terms-of-trade effects in addition to the welfare effects from trade creation and diversion. When the price of a country’s export relative to its import increases, its terms of trade improve and there is an additional welfare gain. When there are terms-of-trade changes, even an RTA that is trade-diverting can result in a net welfare gain.³⁰

²⁹ Grossman and Helpman (1995) make the argument against RTAs using a similar theoretical framework.
³⁰ The case of terms-of-trade changes leading to welfare gains from trade-diverting RTAs is shown, using offer curves, by Wonnacott and Wonnacott (1981), using the importer’s utility, by Johnson (975) and using indifference curves by Lipsey (1957).
Mundell (1964) describes the changes in terms of each country’s trade balance. He considers the case where Country A removes tariffs on Country B. Country A will import more from B and buy less from C and produce fewer home goods. Country B’s exports increase and it experiences a trade surplus, while Country C experiences a trade deficit. The world price must adjust to offset the trade surplus in Country B and the trade deficit in Country C. For country B, the price of its export good relative to the price of its import good must increase to offset the initial trade surplus. Country B’s terms of trade improve. The opposite price adjustment occurs in Country C which experiences a terms-of-trade loss.

The effect on country A is ambiguous: its terms of trade decline relative to B but improve relative to C. Furthermore, even a small country can experience terms-of-trade effects due to an RTA. As Wonnacott and Wonnacott (1981) note, “When a (customs union) is being established, the terms of trade is a slippery concept; … the assumption that A is very small, and is faced by a large B and C does not mean that terms of trade can be ignored, since B’s agreement to cut tariffs will affects A’s terms of trade.” (p. 706).

Johnson (1960) identifies the following conditions for net welfare gains for RTA members, when there are terms-of-trade effects:

- A country is more likely to gain from the creation of tariffs resulting from a customs union the higher the initial level of its tariffs and the more elastic the partner’s export supply curve.
- A country is less likely to lose from trade diversion the smaller are the initial differences in cost between the partner and the foreign sources of supply for goods which they both can produce, the more elastic is the partner supply of such goods, and the less elastic is the foreign supply of them.
- A country is more likely to gain on its terms of trade with the foreign country the more inelastic the foreign export supply to and the more inelastic the foreign demand for its exports.

Terms-of-trade effects result from either large country assumptions or models of trade with differentiated products, which gives even small countries a degree of market power.\(^{31}\) In either case, when tariff elimination increases a member’s demand for its RTA partner’s good, its drives up the partner’s supply price and the member’s within-union terms of trade will deteriorate. Presumably, this loss could be offset by similar concessions in the partner country.

Most analysts seem to agree that countries are not primarily motivated by terms-of-trade impacts when considering RTA membership. However, Bagwell and Staiger (1999) argue for the primacy of terms-of-trade objectives. In their model of a world in

\(^{31}\) Computable general equilibrium (CGE) models, which are often used to empirically estimate the effects of RTAs, typically assume that products are differentiated by country of origin. In this framework, even small countries have a degree of market power. Schiff and Winters (2003) assert that CGE models are therefore biased towards finding benefits from RTAs. This characterization is incorrect because it focuses only on the terms-of-trade gains members experience at the expense of non-members. One needs to consider the changes in both intra-union terms of trade and terms of trade with the non-union countries, which is captured in CGE models.
which tariffs are set uncooperatively, unilateral reform leads to terms-of-trade losses if demand for, and the world price of the import rise. So, countries pursue regionalism in order to allow tariffs to be reduced without the terms-of-trade externality. Wonnacott and Wonnacott (1981) make a similar point—unilateral trade liberalization does not allow the liberalizing country the benefit of accessing another country’s market at its domestic price. In effect, unilateral trade liberalization leads to “missing tariff revenue.” A customs union provides incentives for participants because partners reciprocate and open their markets.

Welfare analysis based on trade creation, diversion, and terms-of-trade changes by sector use a partial equilibrium framework. A more appropriate, general equilibrium measure would consider change in real GDP, real absorption or a utility based measure such as equivalent or compensating variation. Kowalczk (2000) describes a general equilibrium measure from the literature on the theory of tariffs. He notes that the change in a country’s real income (assuming balanced trade) can be decomposed into a terms-of-trade effect and a volume of trade effect. Harrison et al. (1993) describe another general equilibrium measure in which they decompose a nation’s welfare into changes in producer surplus, consumer surplus, and the change in tariff revenue.

Kose and Reisman (2000) evaluate the welfare effects of an RTA using a stylized multi-country general equilibrium model in which tariffs are determined endogenously. They identify the terms of trade and the trade volume effects of unilateral tariff reform, a free trade area and a customs union. They find that a significant fraction of the welfare changes can be explained by the volume of trade effect. Using a more detailed multi-country computable general equilibrium (CGE) model, Hamilton and Whalley (1985) evaluate the effects of RTAs on global and regional trade patterns and welfare. They consider a series of different types of possible (although not necessarily proposed) bilateral RTAs involving the U.S. and other regional groupings to identify what contributes to welfare changes. They find that trade creation and trade diversion issues, which are emphasized in the theoretical literature, are not the important factors determining the gains or losses from an RTA. Instead, what matters is:

- Whether the initial levels of protection are symmetric.
- Relative sizes of the regions.
- The pattern of trade between participating and non-participating regions.

Wonnacott and Wonnacott (1981) also note that the emphasis on trade creation and diversion in trade theory is not relevant for empirical analysis:

“The standard assumption, that A is very small compared to C (the non-RTA partner), is not so reasonable as it seems at first glance; in particular, it is not nearly so reasonable in a many-good world as it seems in the common two- and three-good models of trade theory. In fact, no outside country or group of outside countries is likely to be predominant in the pricing of all goods.” (p. 706).