

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: International Trade in East Asia, NBER-East Asia Seminar on Economics, Volume 14

Volume Author/Editor: Takatoshi Ito and Andrew K. Rose, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-37896-9

Volume URL: http://www.nber.org/books/ito_05-1

Conference Date: September 5-7, 2003

Publication Date: August 2005

Title: The Trade and Investment Effects of Preferential Trading Arrangements

Author: Philippa Dee, Jyothi Gali

URL: <http://www.nber.org/chapters/c0193>

The Trade and Investment Effects of Preferential Trading Arrangements

Philippa Dee and Jyothi Gali

The number of preferential trading arrangements (PTAs) has grown dramatically over the last decade or so. By the end of March 2002, there were 250 agreements in force that had been notified to the World Trade Organization (WTO), compared with 40 in 1990 (WTO 2002).

The coverage of preferential trading arrangements has also tended to expand over time. The preferential liberalization of tariffs and other measures governing merchandise trade remains important in many agreements. But they increasingly cover a range of other issues—services, investment, competition policy, government procurement, e-commerce, labor, and environmental standards.

This paper examines, both theoretically and empirically, the effects of the trade and nontrade provisions of PTAs on the trade and foreign direct investment (FDI) flows of member and nonmember countries.

5.1 Theoretical Review

The first wave of PTAs in the 1950s to 1970s were generally limited in scope, with preferential liberalization of merchandise trade playing a central role (the European Union [EU] was an important early exception). In part, this was because general tariff levels were higher to start with.

The static analysis of first-wave PTAs challenged the presumption that

Philippa Dee is visiting fellow at the Asia-Pacific School of Economics and Government at Australian National University and formerly an assistant commissioner at the Australian Productivity Commission. Jyothi Gali is a senior research economist at the Australian Productivity Commission.

The views are those of the authors and do not necessarily reflect those of any institution with which they have been affiliated.

these were a step in the right direction.¹ It concluded that although PTAs eased one economic distortion, namely, the average tariff on imports in general, they exacerbated another, namely, the geographical disparity in import tariffs. This was a classic situation of second best, with no clear presumption in favor of gains to either PTA members or the world as a whole. The answer “depends,” and the devil is in the detail. The analysis is summarized in the appendix, using a diagrammatic exposition similar to that first developed by Johnson (1960).

The literature also recognized that if the answer depends, then the question is an empirical one. Various analysts examined the trade effects of various PTAs, trying to determine whether they have encouraged imports in general—trade creation—more than they have pushed the geographic source of imports in the wrong direction—trade diversion. There is a degree of apparent consensus (summarized later) about which PTAs have been beneficial and which have been harmful to members. There have also been recent generalizations that PTAs are relatively benign.

Interest in PTAs revived early in the 1980s as the United States reacted first to EU expansionism and the loss of EU markets and then to the uncertain prospects for launching the Uruguay Round, by selecting partners for bilateral and regional trade arrangements. The second-wave agreements were predominantly *free trade agreements*, where members retained their own external tariffs, as opposed to *customs unions*, which adopt a common external tariff. Hence rules of origin became important to prevent trade deflection, whereby imports would enter through the country with the lowest external tariff. The second wave of PTAs also saw the inclusion of nontariff barriers and other nontraditional areas, such as dispute resolution and competition policy. However, the sectoral focus remained on goods markets.

With the second wave, the focus of theoretical work shifted to the dynamic question of whether PTAs were “building blocks” or “stumbling blocks” to multilateral free trade. Bhagwati, Krishna, and Panagariya (1999) identified two distinct approaches. First, suppose a PTA expands its membership. Will that reduce or increase welfare? If expansion increases welfare, then PTAs are seen as building blocks. Second, will a PTA expand its membership? And if so, is there an incentive for expansion to eventually cover the entire world, with nondiscriminatory free trade for all, or will it stop short? This approach uses political economy considerations.

Some partial answers to these questions were provided by Krugman (1993), Deardorff and Stern (1994), Baldwin, (1996), Levy (1997), and

1. The seminal work is Viner (1950). Other early contributions came from Gehrels (1957), Lipsey (1957, 1958), Johnson (1960), Mundell (1964), Corden (1972), and Riezman (1979). Comprehensive surveys of the literature are available in Baldwin and Venables (1995), Pomfret (1997), Bhagwati, Krishna, and Panagariya (1999), and Panagariya (2000), among others. Two recent policy-oriented reviews are by the WTO (1995) and the World Bank (2000).

Krishna (1998). The most recent, comprehensive analyses by Zissimos and Vines (2000) and Andriamananjara (2002) acknowledge that joining a PTA is the best safe-haven strategy when other countries are doing so. But they find that because PTA membership confers a terms of trade gain to members at the expense of nonmembers, at least some members will be better off limiting PTA membership than allowing expansion to cover the world as a whole.² Any redesign of the WTO rules disciplining the formation of PTAs would need to recognize that reality.

During the 1990s the number of PTAs expanded dramatically. In addition to new preferential initiatives by the EU and the United States, the third wave now includes players such as Japan. Until 2002 Japan was one of only four WTO members not to participate in any PTA (although it was a member of nondiscriminatory Asia-Pacific Economic Cooperation [APEC]). Its first agreement, the Japan-Singapore Economic Agreement for a New Age Partnership, typifies many new age agreements. The provisions governing merchandise trade are very limited. Both countries already have zero or very low tariffs on imports of nonagricultural products, and trade in agricultural products between them is minimal, but because of the sensitivity of the trade in cut flowers and goldfish, agricultural and fishery products (along with some petrochemical and petroleum goods) have been excluded from the bilateral agreement altogether. Instead, the agreement focuses on new age issues—especially e-commerce and services. Other such agreements include FDI, competition policy, government procurement, labor, and environmental standards.

Despite the evolution of third-wave or new age agreements, there has been little literature dealing with the effects of preferential nontariff provisions. Two exceptions are Pomfret (1997, chapter 10) and Ethier (1998a,b, 1999, 2001), who deal primarily with effects on investment.

Pomfret (1997) does not discuss in detail the economic welfare effects of discriminatory provisions governing foreign direct investment, but his discussion of the welfare effects of preferential nontariff barriers to trade is suggestive. Pomfret (1997) notes that the critical distinction is whether nontariff barriers are rent generating—allowing a markup of price over cost—or whether they are cost escalating—increasing the real resource costs of doing business.

The analogy with preferential liberalization of *investment provisions* is as follows:

- If investment barriers are of the sort to generate rents, then preferential liberalization will generate gains from investment creation as production is moved from a high-cost domestically-owned producer to a

2. These are further developments of the arguments about the negative externalities from terms of trade changes developed by Bond and Syropoulos (1996) and Bagwell and Staiger (1998, 1999), among others.

lower-cost member's affiliate. But it will also generate losses from investment diversion as production is moved from a low-cost nonmember affiliate (located somewhere in the world) to a higher-cost member affiliate.

- If investment barriers are of the sort to escalate costs, then preferential liberalization will unambiguously save real resources and increase welfare, irrespective of whether the partner is the least-cost location (see also Baldwin 1994).

Thus the welfare implications of preferentially liberalizing investment provisions are more positive than they were for preferential tariff liberalization because of the possibility of saving real resources. But the potential for losses from investment diversion also remains.

In a series of papers, Ethier (1998a,b, 1999, 2001) develops variants of a model in which investment responds in "beachhead" fashion to the preferential *trade provisions* of PTAs.

This model is an explicit attempt to capture some of the salient features of third-wave PTAs. Many third-wave agreements are between small and larger countries. The small countries want to reform their internal economies so that they can be accepted as members of the global trading system. The sign of successful reform is whether these countries attract FDI. The small countries use (often asymmetrical) trade concessions to large countries as a way of signaling a credible commitment to reform.

There is no presumption in Ethier's (1998b) framework that the investment they attract comes from the large PTA partner. The aim of these small reforming countries is often to divert investment from nonmember countries. Ethier (2001) also examines in detail the incentives for the larger country to accede to such an arrangement, even in preference to pursuing further multilateral reform. Finally, he shows that a world equilibrium in which small countries compete for investment in this fashion is beneficial because it internalizes an externality associated with agglomeration economies.

Ethier's (1998a,b, 1999, 2001) positive outlook on PTA formation comes from this benign view of competition for investment, rather than from the characteristics of PTAs per se. As he acknowledges, his model of PTA formation is consistent with massive amounts of investment diversion to take advantage of trade beachheads and subsequent trade diversion from those beachhead positions. But in his model, there is sufficient symmetry between countries for this trade and investment diversion to have no adverse welfare consequences—every country is the lowest-cost source of imports and the best host for FDI. With more diversity, this massive diversion is no longer benign.

Ethier's (1998a,b, 1999, 2001) positive view also depends on the competition for investment occurring through reform, which is seen as a good thing. If it were to occur through the competitive granting of investment incentives, or if reform involved inappropriate concessions forced by a larger

hegemon (as Bhagwati [1999] fears), the competition for investment may itself be less benign.

At first sight, the focus of third-wave agreements on nontariff issues may suggest that traditional concerns about trade diversion are outmoded. But the theoretical literature suggests otherwise. Investment barriers can be used as a protective device, and preferential liberalization of investment provisions can generate investment diversion, with adverse consequences, as well as beneficial investment creation. Even where investment is attracted in beachhead fashion in response to trade liberalization provisions, both the investment and subsequent trade from the beachhead position may be diversionary. Thus the nontariff focus of third-wave agreements cannot shake the first-wave concerns about the adverse second-best effects of preferential liberalization.

The second section of this paper summarizes the trade and nontrade provisions included in a number of recent PTAs. The third section empirically estimates the effects of recent trade and nontrade provisions on bilateral trade and investment flows. The fourth section recapitulates the key findings.

5.2 Breadth of Coverage of PTAs

Figure 5.1 shows the discernible upward trend in the breadth of coverage of PTAs over recent times. On the vertical axis is an index measure of

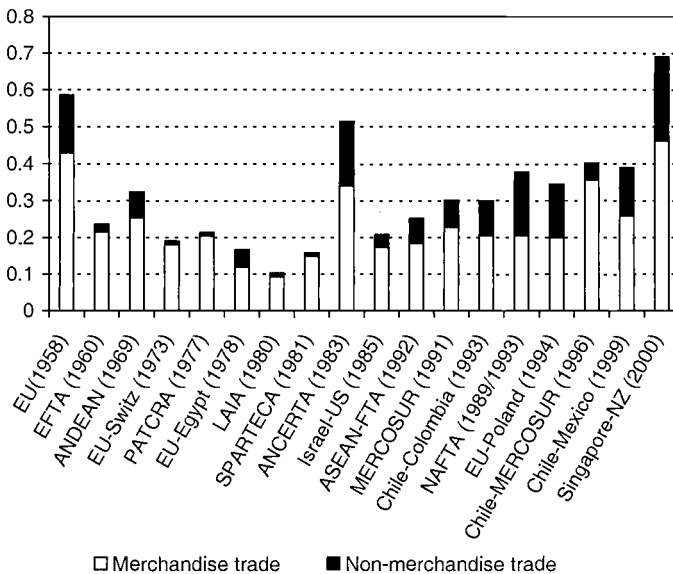


Fig. 5.1 Member Liberalization Index (MLI) for selected PTAs, index score ranges between zero and one

Source: Adams et al. (2003).

breadth of coverage, with provisions governing merchandise and nonmerchandise trade scored separately. The Member Liberalization Index (MLI) is described in detail in Adams et al. (2003). On the horizontal axis is the date of establishment. Coverage has clearly tended to increase in the more recently established or expanded PTAs, and this has generally been because of an expansion in the coverage of nonmerchandise trade issues.

The index includes provisions covering the following:

- Agricultural products—including domestic support measures, tariff quotas, sanitary and phytosanitary measures, tax exceptions, export incentives, and technical barriers to trade, among others
- Industrial products—including coverage and restrictiveness of rules of origin, safeguards, antidumping, coverage and timing of tariff preferences, among others
- Services—including provisions governing market access and national treatment in services
- General measures—general national treatment provisions, investment rules, domestic competition policy, government procurement, intellectual property rights, and general provisions covering the temporary and permanent movement of people

These provisions are classified into two subindexes for quantitative analysis. The merchandise MLI includes the provisions covering agriculture and industrial products—an index of traditional provisions. The nonmerchandise MLI, covering third-wave issues, includes the services provisions, plus the general measures covering all trade.

The coverage varies from one PTA to another. Some involve only a few products or sectors, while others stretch well beyond the traditional tariff elimination. Note that the scores are based on how the language of the agreements is written, not on whether or how the provisions are used. A high index for nonmerchandise trade indicates that a PTA is more liberal to members in its services trade, investment, and related provisions. This index takes a high value for Singapore-New Zealand (NZ), followed by Closer Economic Relations (CER) (between Australia and New Zealand), the North American Free Trade Agreement (NAFTA) and the EU.

The provisions indexed in the MLI are treated as additive to and independent of each other. In reality some provisions might interact to strengthen or weaken other provisions. For example, the time schedule for preferential tariff liberalization is closely related to the restrictiveness of rules of origin. The impact of interaction effects among the provisions in various PTAs is potentially an empirical question, but interaction effects have not been allowed for specifically in the construction of this index nor in the subsequent econometric analysis. For this reason, the econometrics may understate (where interaction effects reinforce) or overstate (where interaction effects cancel) the overall effects of PTAs.

The estimated relationship between these provisions and the level of trade (or investment) provides an indication of whether provisions included in PTAs have any effect collectively on trade (or investment) flows with member or nonmember countries. Because PTAs are by definition exclusive and discriminatory against nonmembers, trade and nontrade provisions that are favorable to the intra-PTA trade (or investment) may become barriers to nonmember countries.

5.3 Empirical Analyses

The key empirical task is to disentangle the effects of PTA formation or expansion from all other influences on trade and investment flows. There are two main approaches available in the literature.

First, *ex ante* studies have used counterfactual analyses based on partial or general equilibrium models. These models assume a certain model structure, with specific functional forms and parameter values to represent the countries in a base year prior to the formation of the PTA. Those models with a sufficiently tight theoretical structure can also be used to draw direct inferences about welfare. The model is then subjected to the preferential removal of tariffs alone, and the welfare effects are calculated. Surveys of assessments of PTAs using general equilibrium models can be found in De Rosa (1998) and Robinson and Thierfelder (2002). Scollay and Gilbert (2000) survey computable general equilibrium (CGE) assessments of APEC. Most of these studies find that PTAs create additional trade for both members and nonmembers. Most also find that PTAs improve welfare, at least among member countries.

However, these CGE analyses suffer from a number of theoretical and practical difficulties. Some (in particular, many of those covered by the Robinson and Thierfelder [2002] survey) assume fixed terms of trade. As noted by Panagariya and Duttagupta (2002), this is inconsistent with one of their other key assumptions, namely, product differentiation at the national level. Deardorff and Stern (1994) note how the assumption of national product differentiation can itself leave an “idiosyncratic stamp” on examinations of PTAs, in particular helping to explain Krugman’s (1993) finding of welfare losses in a world of three trading blocs, a result that does not appear to carry over to empirical CGE analyses. But in addition, the assumption of fixed terms of trade rules out one of the key effects of PTAs, namely, terms of trade changes.

Further, the CGE studies typically use a very simple characterization of PTAs. Most assume comprehensive across-the-board elimination of tariffs (and sometimes nontariff barriers) among members, although most real-world PTAs have complex patterns of exemptions. In addition, the studies typically ignore many of the potentially trade-restrictive nontariff measures, such as rules of origin or local content requirements, that typically

accompany the merchandise trade measures. Finally, they typically ignore provisions affecting nonmerchandise trade (although a notable exception is Hertel, Walmsley, and Itakura 2001).

This is not to deny that particular CGE models, when used with appropriate assumptions (such as variable terms of trade), can give valuable insights into the possible effects of important tariff provisions of PTAs. But conclusions drawn from *surveys* of CGE studies should be treated cautiously. And the results from CGE studies should not be generalized to draw conclusions about the effects of nonmerchandise trade provisions of PTAs.

By contrast, ex post studies of PTAs measure their trade creation and trade diversion effects by using econometric methods to establish a link between actual PTA formation and actual trade outcomes, controlling for the effects of all other influences. Because welfare is unobservable, these econometric studies cannot establish welfare effects directly. And as noted in the appendix, the link between trade outcomes and welfare is weak. But the studies do examine *actual* PTAs, in all their complexity, including nonmerchandise trade provisions. The present study is an ex post evaluation of the effects of PTAs.

5.3.1 Gravity Model

The gravity model is the key ex post econometric technique used to examine the determinants of bilateral trade flows. It is a model of trade flows based on an analogy with the law of gravity in physics. Trade between two countries is positively related to their size and inversely related to the distance between them. A number of other explanatory variables are added to this model to analyze various bilateral trade policy issues. In the augmented gravity model, trade between two countries is determined by supply conditions at the origin, demand conditions at the destination, and various stimulating or restraining forces. This specification has recently been shown to be consistent with a number of theoretical models of international trade.³

The standard way of assessing the impact of PTAs is to add PTA-specific binary dummy variables to the augmented gravity model to capture effects not captured through normal bilateral trade determinants. Studies adding PTA-specific dummy variables to capture the trade creation and diversion effects of PTAs date back to the 1970s. Aitken (1973) initially added one

3. The gravity model can be derived theoretically as a reduced form from a general equilibrium model of international trade in goods. Baier and Bergstrand (2001) derived it from a model of monopolistic competition. Feenstra, Markusen, and Rose (2001) derived it from a reciprocal dumping model of trade with homogeneous goods. Deardorff (1998) derived it from a model with perfectly competitive markets. Evenett and Keller (2002) showed empirically that the monopolistic-competition-based theory of trade fits the trade flows among industrialized countries well. Anderson and van Wincoop (2003) nevertheless showed that many empirical implementations have strayed from the theoretically derived reduced form.

dummy variable to his gravity model to capture the intrabloc effect of a PTA—“a gross trade effect” of Balassa (1967). Bayoumi and Eichengreen (1995) and Frankel (1997) added two dummy variables for each PTA to capture the separate effects on intrabloc and extrabloc trade. The first dummy variable takes a value of one when the two countries are members of the same PTA. The second dummy variable is one if either country in a particular pair belongs to the PTA. If a positive coefficient on the first dummy exceeds a negative coefficient on the second, then trade creation may be said to outweigh trade diversion.

Soloaga and Winters (2001) added three dummy variables for each PTA to distinguish an intrabloc effect, an extrabloc effect on imports, and an extrabloc effect on exports (see figure 5.2). The second and third dummy variables in their study measure the extent of import diversion and export diversion, respectively. They argued that both are needed because bloc members’ imports and exports could follow different patterns after the formation of a PTA.

The current analysis also uses three dummy variables for each PTA. But instead of taking simple zero-one values, irrespective of the scope or coverage of the PTA provisions, the dummies take the value of the MLI index (or subindex) whenever the PTA is in force.

Because the gravity model is estimated using panel data, the PTA-specific dummy variables are introduced in one of two ways. Firstly, *dynamic* PTA-specific indexes are defined to take a nonzero value only for the years in which a PTA between the two countries is in force and to take a value of zero otherwise. These indexes capture the effect of the formation,

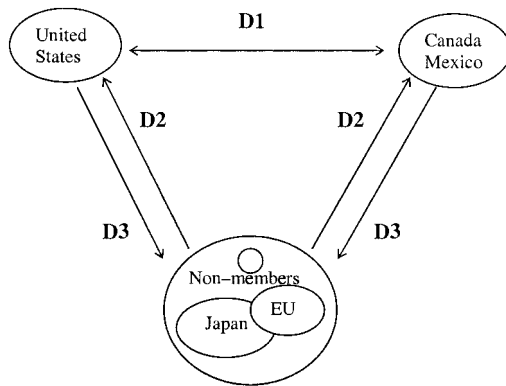


Fig. 5.2 Use of PTA-specific dummy variables in a gravity model—the example of NAFTA

Notes: D1 captures the effects of NAFTA on intrabloc trade. D2 captures the effects of NAFTA on members’ imports from nonmembers—import diversion if the coefficient is negative. D3 captures the effects of NAFTA on members’ exports to nonmembers—export diversion if the coefficient is negative.

expansion, and contraction of a PTA on trade and investment only after it occurs. In contrast, *antimonde* PTA-specific indexes take a nonzero value for all the years in the sample, irrespective of when the PTA was formed. These *antimonde* indexes are used as the panel analogue to the nondynamic indexes of previous cross-sectional gravity model studies. They have the same disadvantage as these studies of allowing the formation, expansion, and contraction of a PTA to affect trade and investment before the event.

5.3.2 Model Specification

Effects on Trade

The panel data includes information on all potential trading partners, even when a country has no exports to some partners in some years. Because the nature of trade relations in many countries in the World Trade Flows (1997, 2000) database is such that each country trades with a relatively small number of partners, the dependent variable contains a significant number of zero observations as well as many positive observations.⁴ As a result, a Tobit estimation procedure is used to appropriately account for the censored nature of the dependent variable—the natural log of exports between country i and country j in year t .⁵ Adams et al. (2003) describe the data sources in detail. The panel has a relatively long time dimension, covering 1970–1997 for the estimation of trade effects (1988–1997 for the estimation of investment effects). While this risks the problem of subsample instability, it has the advantage that it helps to overcome the problem of nuisance parameters in the estimation of fixed effects in a Tobit context (Greene 2002).

The gravity model estimated here allows for *product differentiation* at the country level. Much of the recent literature on PTAs has focused on imperfectly competitive behavior. Recognizing this is important for two reasons:

- Some economic integration has occurred among economies with almost similar structures and large volumes of intraindustry trade.
- There is a positive interaction between market structure and the gains from integration, often called the procompetitive effects of PTAs, which the new age agreements aim to capture.

4. For the full sample trade model, the number of observations is 116 countries \times 115 partners \times 28 years = 373,520, with about 44 percent having zero values.

5. There are a variety of alternate approaches to this problem. The zero values can be simply omitted as in the case of Frankel (1997), which leads to the possibility of selectivity bias. Arbitrarily small numbers can be used in place of zeros. Eichengreen and Irwin (1995) expressed the dependent variable as $\ln(1 + Y_{ij})$. Clark and Tavares (2000) and Soloaga and Winters (2001) used a Tobit specification for their cross-sectional gravity model.

The product differentiation model of Helpman and Krugman (1985) and Helpman (1987) is integrated into the current gravity model specification. In their models, one of the two goods is differentiated and the other is homogenous. The bilateral trade of each country is the sum of interindustry and intraindustry trade flows, with the latter being trade in the differentiated product.

The corresponding reduced form of gravity model for trade is

$$\begin{aligned}
 (1) \quad \text{Ln } Y_{ijt}^* &= a + \alpha_i + \gamma_j + \lambda_t + \beta_1 \text{Ln SGDP}_{ijt} + \beta_2 \text{Ln RLFA}_{ijt} \\
 &+ \beta_3 \text{Ln Similar}_{ijt} + \beta_4 \text{Ln Dis}_{ij} + \beta_5 \text{Ln RER}_{ijt} + \beta_6 \text{Ln TAR}_{ijt} \\
 &+ \beta_7 \text{Ln}_{ij} + \beta_8 \text{Bor}_{ij} + \beta_9 \text{Col}_{ij} + \beta_{10} \text{Cur}_{ij} + \beta_{11} \text{Is}_i + \beta_{12} \text{Is}_j \\
 &+ \beta_{13} \text{lock}_i + \beta_{14} \text{lock}_j + \beta_{15} \text{3wave}_{ij} + \sum_j \text{MRTA}_{ij} \\
 &+ \sum_j \text{MRTA}_{i-j} + \sum_{-i} \text{MRTA}_{j-i} + \text{Ln } \varepsilon_{ijt}
 \end{aligned}$$

where

- Ln is natural logarithmic transformation;
- Y_{ijt}^* is the value of exports from country i to j in year t ; using exports as a dependent variable rather than total bilateral trade allows the identification of export and import diversion separately;
- α_i is unobserved specific effects in exporting country i ;
- γ_j is unobserved specific effects in importing country j ;
- λ_t is unobserved specific effects in time period t ;
- SGDP_{ijt} is the sum of bilateral gross domestic products (GDPs) of i and j in year t ;
- RLFA_{ijt} is the absolute differences in GDP per capita of i and j in year t ;
- Similar_{ijt} is similarity in country size between i and j in year t in terms of aggregate GDP;
- Dis_{ij} is distance between the two largest or capital cities of countries i and j ;
- RER_{ijt} is the bilateral real exchange rate between i and j in year t ;
- TAR_{ijt} is an average tariff rate in importing country j on goods from country i in year t ;
- Ln_{ij} is a measure of linguistic similarity between i and j ;
- Bor_{ij} is a dummy that takes a value 1 if i and j share a land border and 0 otherwise;
- Col_{ij} is a dummy that takes a value 1 if i and j have colonial linkages and 0 otherwise;
- Cur_{ij} is a dummy that takes a value 1 if i and j have the same currency and 0 otherwise;
- Is_i is a dummy that takes a value 1 when i is island nations and 0 otherwise;

- $Lock_i$ is a dummy that takes a value 1 when i is a landlocked nation and 0 otherwise;
- $3wave_{ij}$ is an index capturing the third wave provisions of a PTA, that takes a value of the nonmerchandise MLI index if the i and j are participants of a specific PTA in the sample and 0 otherwise; it also has a time dimension when defined in dynamic rather than antimonde form;
- $MRTA_{ij}$ is an index capturing the merchandise trade provisions of a PTA, that takes the value of the merchandise MLI if both countries i and j belong to the same PTA and 0 otherwise; it also has a time dimension when defined in dynamic rather than antimonde form;
- $MRTA_{i-j}$ is an index that takes the value of the merchandise MLI when the importing country j belongs to that particular PTA and 0 otherwise; it also has a time dimension when defined in dynamic rather than antimonde form;
- $MRTA_{j-i}$ is an index that takes the value of the merchandise MLI when the exporting country i belongs to that particular PTA and 0 otherwise; it also has time dimension when defined in dynamic rather than antimonde form; and
- ε_{ijt} is an error term.

From an econometric point of view, the α_i , γ_j and λ_i specific effects are treated as fixed unknown parameters. The use of three separate fixed effects is advocated by Matyas (1997, 1998) and avoids the omitted variable bias identified by Haveman and Hummels (1998) and Anderson and van Wincoop (2003).

The expected relationship of the observed explanatory variables with bilateral exports is discussed in detail in Adams et al. (2003). In a model of product differentiation, countries similar in size will trade more, and the trade will be of an intraindustry nature. The index of size similarity (Similar) captures this effect. By contrast, traditional trade theory says that countries with dissimilar levels of per capita GDP will trade more than the countries with similar levels. The absolute difference in the per capita GDP between exporting and importing countries (RLFA) is included as an explanatory variable in the gravity model as a way of distinguishing the traditional from the differentiated product approaches.⁶

The preceding gravity model specification includes the real exchange rate (RER) as a relevant price variable in order to control for fluctuations in relative prices among trading partners.

The average bilateral tariff rate (Tar) is expected to show a negative rela-

6. The specification based on product differentiation in the preceding also differs from traditional gravity model specifications by including the sum of importing and exporting country GDPs, rather than including each separately. This small loss of generality means that the product differentiation version does not encompass the traditional model fully.

relationship with trade. The PTA-specific indexes capture the extent of traditional and new age provisions of a PTA but not the size of the tariff preferences thereby created. Because the bilateral tariff variable includes preferential tariffs,⁷ the overall measured effect of PTAs on trade will be split between the tariff variable and the PTA-specific indexes in specifications where both occur. To test whether the coefficients of the PTA-specific indexes are sensitive to the inclusion of the tariff variable, the gravity model is estimated initially without the tariff variable. It is then reestimated with the tariff variable for that subset of countries and time periods for which bilateral tariff data are available.

Effects on Investment

The product differentiation specification also provides a rationale for applying the gravity model to investment flows (Egger 2001).

The raw FDI data for this analysis are sourced from the United Nations Conference on Trade and Development (UNCTAD) and the Organization for Economic Cooperation and Development (OECD) for the period between 1988 and 1997 for about seventy-seven countries (see Adams et al. 2003). As there are some deficiencies in this data, the qualitative aspects of the analytical results rather than the precise magnitude of the investment estimates are of main interest.

The dependent variable in the gravity model is the natural logarithm of the stock of outward investment from home country to host country. The stock of outward investment is used as the dependent variable, rather than outflows, for two reasons. First, more outward stock than outflow data are available in the source documents. For many countries in the late 1980s and for some Latin American countries in early 1990s, the bilateral FDI flow data are not fully reported in the UNCTAD investment directories. Second, statistical tests suggested that a gravity model based on the stock of outward investment was preferred to a model based on outflows.

Apart from the more limited number of years and countries analyzed, the investment model is similar to the trade model. In particular, the key bilateral determinants are the same for trade and investment (see also Egger 2001), although the sign and magnitude of the impact of some of these explanatory variables differs.

For each PTA, three merchandise MLIs and three indexes of third-wave provisions are included in the investment gravity model to test how the investment to members, and to and from nonmembers, responds to the traditional and third-wave provisions embedded in each PTA. Three merchandise MLIs and only one (intrabloc) nonmerchandise MLI were included in the trade gravity model. Because of model convergence prob-

7. The bilateral tariff data are applied rates obtained from UNCTAD's Trade Analysis and Information System (TRAINS) database. As such they incorporate tariff preferences.

lems, the effects of new age provisions on exports to nonmembers and imports from nonmembers could not be analyzed separately in the trade model.

Two additional variables are added to the investment model because in addition to the investment provisions of PTAs, countries also negotiate bilateral investment treaties (BITs). About 191 PTAs were in force in 2000, with only a few covering investment provisions, while 1,941 BITs were in place then. The specification controls for whether an investment treaty is either signed or enacted between a pair of countries.

There are clearly some interdependencies between the trade equation and the investment equation, but in the absence of the remaining elements of the balance of payments, there is an insufficiently tight link to warrant seemingly unrelated regression or some other systems-estimation technique. Similarly, there are no obvious cross-equation restrictions that would improve the efficiency of estimation. One option would have been to include trade flows in the investment equation and investment flows in the trade equation, but this would have created a severe simultaneity problem. In the current PTA context, it was felt that including trade PTA dummies in the investment equation and vice versa was a minimalist approach to capturing the interdependencies.

The expected effects of traditional and new age provisions on investment are not straightforward.

If trade liberalization makes exporting from the home country relatively more attractive than FDI as a way to serve the regional market, then the trade provisions of a PTA could cause a reduction in intrabloc investment. But the trade provisions could also enable transnational corporations to operate vertically in a PTA area, stimulating intra-FDI flows among the relevant partners. The structure and motivation of investment will determine the net impact of trade provisions of PTAs on intra-PTA investment. So too will the structure and motivation of intrabloc trade (Markusen 1983).

According to Ethier (1998b, 2001), the inflows of FDI from nonmember countries into the PTA region are likely to go up in response to the trade provisions of PTAs, as nonmembers establish beachhead positions in one PTA member country in order to serve the market of the others. Alternatively, if multinationals are initially operating in member countries to serve the protected local market (the tariff jumping motivation for investment), then these multinationals may rationalize their network of affiliates after the formation of the PTA, and, as a result, some member countries could lose investment to nonmember countries.

Thus, the response of investment to the merchandise trade provisions of a PTA is an empirical question. The various possibilities can be tested in the following way.

If investment responds in beachhead fashion to the trade provisions of PTAs and in turn stimulates intrabloc trade, this can be identified by the combination of a positive and significant effect of trade provisions on intra-

bloc trade and a positive and significant effect of trade provisions on investment from nonmember countries.

Alternatively, a reversal of tariff-jumping investment can be identified by a positive and significant effect of trade provisions on investment to nonmember countries.

Investment may also respond to the nontrade provisions of PTAs. If, as a result, production is moved from a high-cost domestically owned producer to a lower-cost member's affiliate, this investment creation is likely to benefit members of the PTA. But if production is moved from a low-cost nonmember affiliate to a higher-cost member affiliate, this investment diversion may not benefit members.

Measures of net investment creation or diversion can be obtained by summing the significant coefficients of the three separate nonmerchandise MLI variables in parallel fashion to the trade equation. One further qualification to the welfare implications of investment is that if the initial nontrade restrictions are of the sort to raise costs rather than generate rents, then any investment relocation in response to their preferential removal will unambiguously benefit members.

5.3.3 New Evidence on Trade Creation and Diversion

The observable effects—normal bilateral trade determinants and trade provisions of PTAs—and unobservable country- and time-specific effects all significantly influence the bilateral trade flows. The signs and significance of the coefficients on the observable effects are generally as expected (see table 5.1). Interestingly, they support both traditional and product differentiation theories of trade because similarity in size and differences in income per head are both associated with higher bilateral exports. In the preferred specification with dynamic PTA variables and fixed effects, the coefficient on the sum of GDPs is about 2, as expected.

The new estimates of trade creation and diversion tend to be different from past estimates for most PTAs. Past estimates showed most PTAs to be trade creating in net terms. By contrast, the results here suggest most PTAs do not create additional trade, either for members or for nonmembers of the agreement. The net trade effects of preferential agreements found in this study are compared with past estimates in table 5.2, which shows whether the net effects are positive or negative.

Nearly all PTAs are found to have caused net trade diversion in the new assessment. The PTAs found to have inconclusive effects in past analysis drifted either way in the new assessment, but the Southern Common Market (MERCOSUR) was found here to have caused net trade diversion.⁸

8. The assessment of net trade effects is based on the marginal effects of PTAs reported in Adams et al. (2003), rather than the raw maximum likelihood Tobit estimates in table 5.1, for the preferred specification with dynamic PTA variables and fixed effects. Although the marginal effects and raw Tobit estimates are not equal (as explained in Adams et al. 2003), in practice the assessment of the direction of net trade effects is the same, whichever is used.

Table 5.1

Gravity model of trade: Econometric results from full sample

Variable name	Dynamic PTA indexes		Antimonde PTA indexes	
	Without fixed	With fixed ^a	Without fixed	With fixed ^{a,b}
Ln sum of exporting and importing countries GDP (SUM)	2.841***	2.008***	2.185***	2.066***
Similarity in exporting and importing country's GDPs	1.245***	0.637***	0.965***	0.665***
Ln of absolute differences in per capita GDPs of exporting and importing country	0.361***	0.310***	0.217***	0.251***
Ln distance	-1.729***	-2.193***	-2.292***	-2.306***
Ln bilateral real exchange rate	0.162***	0.054***	0.023**	0.049***
Linguistic similarity	0.000***	0.000***	0.000***	0.000***
Colonial	1.167***	1.759***	1.628***	1.575***
Border	-0.088	-0.571***	-0.529***	-0.626***
Currency union	1.201***	3.136***	1.148***	3.025***
Exporting country is an island	0.684***	-2.250***	0.670***	0.289
Importing country is an island	1.070***	-3.369***	1.163***	-3.268***
Exporting country is landlocked	-2.292***	-3.456***	-1.869***	-0.648*
Importing country is landlocked	-2.052***	3.276***	-1.929***	4.515***
Third-wave provisions of PTAs	20.074***	13.899***	-10.760***	-8.748***
Andean1	3.135*	4.544***	3.871***	2.774***
Andean2	2.496***	-0.600	11.257***	
Andean3	-0.943***	-3.088***	13.716***	
APEC1 ^c	-2.081***	-2.727***	-0.052	0.091*
APEC2 ^c	-0.240***	0.583***	2.118***	-0.666***
APEC3 ^c	1.245***	0.486***	4.404***	1.941***
EFTA1	-6.252***	-7.023***	-1.972	-0.690
EFTA2	12.322***	0.252	9.111***	
EFTA3	17.195***	3.141***	15.189***	
EC/EU1	-16.129***	-16.022***	8.763***	9.608***
EC/EU2	5.344***	-1.209***	-8.208***	10.632***
EC/EU3	6.343***	-0.486*	-7.920***	18.188***
GCC1 ^c	-0.400	-1.782***	-0.135	-0.341*
GCC2 ^c	-0.498***	0.139*	0.950***	0.855***
GCC3 ^c	-2.098***	-0.600***	0.118***	2.379***
LAFTA/LAIA1	30.591***	17.419***	28.057***	26.432***
LAFTA/LAIA2	-20.659***	-6.517***	-22.841***	
LAFTA/LAIA3	-5.267***	-0.635	-32.910***	
MERCOSUR1	-6.894**	-9.376***	0.800	1.075
MERCOSUR2	1.451***	1.929***	-5.002***	
MERCOSUR3	2.917***	-1.306***	23.916***	
NAFTA1	-17.152***	-14.970***	-2.072	-0.966
NAFTA2	5.195***	1.166**	7.310***	
NAFTA3	-2.720***	-0.790	-1.938***	
SPARTECA1	42.499***	35.093***	31.956***	31.573***
SPARTECA2	-9.865***	-0.402	-12.250***	
SPARTECA3	-13.312***	0.557	-18.496***	
CER1	-28.857***	-24.283***	-16.504***	-17.251***
CER2	3.329***	-2.229***	2.285***	
CER3	8.040***	-2.073***	7.650***	
EU-Switzerland1	-24.872***	-32.320***	-28.599***	-27.680***

Table 5.1 (continued)

Variable name	Dynamic PTA indexes		Antimonde PTA indexes	
	Without fixed	With fixed ^a	Without fixed	With fixed ^{a,b}
EU-Switzerland2	9.457***	5.339***	25.975***	
EU-Switzerland3	11.542***	5.076***	26.380***	
Chile-Colombia1	-17.149**	-14.407**	4.525*	3.281
Chile-Colombia2	2.234***	4.116***	-4.483***	
Chile-Colombia3	-0.251	2.237***	3.093***	
Chile-Mexico1			-4.187*	-4.096***
Chile-Mexico2			-4.933***	
Chile-Mexico3			1.400***	
US-Israel1	15.060***	10.984***	14.783***	14.185***
US-Israel2	-5.774***	-2.725***	1.888***	
US-Israel3	1.112**	-1.435**	9.056***	
Australia-PNG1	0.669	-6.200**	-10.816***	-10.797***
Australia-PNG2	0.784*	-1.202*	2.390***	
Australia-PNG3	1.487***	-0.946	-2.706***	
Singapore-New Zealand1			2.186*	1.802*
Singapore-New Zealand2			4.390***	
Singapore-New Zealand3			3.587***	
Chile-MERCOSUR1	-7.199**	-11.064***	-3.888***	-4.124***
Chile-MERCOSUR2	2.632***	2.136***	11.792***	
Chile-MERCOSUR3	0.328	1.145***	1.603***	
EU-Egypt1	-4.724	-8.702***	0.622	-0.055
EU-Egypt2	-12.498***	3.048***	3.582***	
EU-Egypt3	-15.582***	4.185***	-2.505***	
EU-Poland1	-19.307***	-27.309***	-9.699***	-11.991***
EU-Poland2	-4.386***	-0.834***	13.716***	
EU-Poland3	-2.186***	-0.741**	21.851***	
AFTA1	-3.783	-9.232***	-5.953***	-5.597***
AFTA2	7.170***	4.191***	0.492*	
AFTA3	7.375***	4.826***	-2.095***	
Constant	-12.101	-2.067***	-1.962***	-1.910***
LR χ^2 (chi2)	307,352.95	422,218.2	351,561.7	424,106.7
Log likelihood	-686,398.4	-629,010.8	-664,294.0	-628,021.5
σ (standard deviation of the error term)	4.407	3.547	4.036	3.542

Source: Authors' estimates.

Notes: Dependent variable = Ln exports; time period = 1970–1997; unbalanced panel; Tobit maximum likelihood estimates.

^aTo save space the fixed effect coefficients associated with exporting country, importing country, and time are not reported here.

^bIn the antimonde specification, some PTA indexes are dropped because of high multicollinearity between the country-fixed effects and the PTA indexes.

^cWhile a Member Liberalization Index has not been calculated for APEC (a nonpreferential arrangement) or for the Gulf Cooperative Council (a preferential one), their possible effects on the trade flows of their members have been controlled for through a set of three conventional zero-one dummy variables.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 5.2 **New evidence on PTAs as causing net trade creation or diversion**

Net trade creation	Past estimates		New estimates	
	Inconclusive	Net trade diversion	Net trade creation	Net trade diversion
Andean CER AFTA EEC/EU? EFTA?	LAIA MERCOSUR	NAFTA	Andean LAFTA/LAIA United States-Israel SPARTECA	AFTA EFTA EC/EU MERCOSUR NAFTA CER EU-Switzerland Chile-Colombia Australia-PNG Chile-MERCOSUR EU-Egypt EU-Poland

Sources: Past estimates assessment based on the findings from a majority of the following studies: Bayoumi and Eichengreen (1995), Frankel, Stein, and Wei (1995), Boisso and Ferrantino (1997), Frankel (1997), Fink and Primo Braga (1999), Krueger (1999b), Li (2000), Clark and Tavares (2000), Freund (2000), Gilbert, Scollay, and Bora (2001), and Soloaga and Winters (2001); table 5.1 source.

Overall, the main PTAs—NAFTA, the European Community (EC/EU), MERCOSUR, and CER—as well as many bilateral agreements not considered previously, are found here to have created negative net trade effects. However, there is a qualification to this finding. In agreements with a small number of members, the intra-PTA effect is estimated imprecisely, with a large standard error, while the extra-PTA effect can be estimated more accurately.⁹ Thus, the findings for those PTAs, such as CER, with a small number of members are less robust than those for larger PTAs. In addition, the measures of distance used in this study are unlikely to capture fully the ways in which changes in trading patterns and reductions in transport costs have raised the attractiveness of extrabloc as opposed to intrabloc trade for CER members over time.

As noted, the net trade effect criterion has limitations in assessing the effects on economic welfare. Nonetheless the new evidence suggests negative net trade effects for many PTAs, controlling for other factors.

A number of factors have contributed to the more negative findings in this study. These are now considered in turn.

All the past gravity model studies surveyed here estimated the PTA effects using PTA dummies defined in antimonde form.¹⁰ The comparable

9. For example, the intra-CER dummy has positive values only for $14 \times 15 = 210$ observations and zero for remaining observations.

10. A more recent study with panel data and dynamic dummies is by Fukao, Okubo, and Stern (2003).

dynamic and antimonde estimates in this analysis are reported in table 5.1. They show that when PTA dummies are defined in antimonde form, the net trade effects are mainly positive, in contrast to the negative effects obtained for dynamic PTA variables.

In essence, when dummies are defined in dynamic form, the test for significance of their coefficients is a statistical test for whether the trade effects they capture are stronger after the formation or expansion of the PTA than before. In the past, this question has been assessed, at best, only by reference to the point estimates from various cross sections. Defining PTA dummies in dynamic form provides a more stringent statistical test of whether it was PTA formation, rather than some other set of factors specific to the bilateral country pair, accounting for the observed trade effects. The power of the test is further strengthened by the fact that individual country- and time-specific effects are controlled for separately, through the fixed effects. The more stringent test of the before and after effects of PTAs is the major factor accounting for the more negative findings of this study.

The differences are more prominent for the EC/EU and MERCOSUR agreements, where membership dynamics play an important role in their trade creation and diversion effects. For example, a significant negative intra-EU effect is found when using dynamic PTA specific indexes, compared to a significant positive effect found using antimonde indexes. The dynamic dummies account for individual countries switching from the European Free Trade Agreement (EFTA) to the EU.

Another reason for the more negative findings in this study is the use of panel analysis, which allows unobservable heterogeneity to be controlled for. Without allowing for country-specific effects, the coefficients on both the PTA variables and the other explanatory variables tend to be upward biased as are the test statistics for the significance of these variables. The likelihood-ratio test confirms the joint significance of the fixed effects.¹¹ This suggests that inferences based on past gravity model estimates without fixed effects suffer from omitted variable bias. Controlling for unobservable heterogeneity is another reason for the more negative findings in this study.¹²

A model with the average bilateral tariff variable as an additional determinant of trade is estimated on a restricted data set,¹³ and the results are shown in table 5.3. The average tariff rate in the importing country has a

11. The calculated test statistic of 114,775.3 clearly rejects the null hypothesis, as is expected given the individual significance of most of the country- and time-specific fixed effects in the model.

12. The findings here are also more negative than those in a recent panel study by Clarete, Edmonds, and Wallack (2003). However, their study uses antimonde dummies and fails to control for unobserved country- and time-specific heterogeneity. It also controls for fewer observable factors than here.

13. The data set is restricted because of the lack of bilateral tariff data for a number of countries and for a number of years.

Table 5.3

Gravity model of trade: Results from limited sample with tariff variable included

Variable name	Dynamic PTA specific variables—fixed effects ^a		Antimonde PTA specific variables—fixed effects ^{a,b}	
	Without tariff	With tariff	Without tariff	With tariff
Ln sum of exporting and importing countries GDP (SUM)	2.063***	2.036***	2.838***	2.869***
Similarity in exporting and importing country's GDPs	0.562***	0.542***	1.534***	1.557***
Ln of absolute differences in per capita GDPs of exporting and importing country	-0.154***	-0.086***	-0.100***	-0.031
Ln distance	-1.404***	-1.393***	-1.458***	-1.469***
Ln bilateral real exchange rate	0.494***	0.513***	0.448***	0.454***
Ln tariff		-0.134***		-0.142***
Linguistic similarity	0.000***	0.000***	0.000***	0.000***
Colonial	1.141***	1.087***	1.170***	1.113***
Border	-0.099	-0.052	-0.138	-0.079
Currency union	0.230	0.425	0.203	0.421
Exporting country is an island	-0.718***	-0.746***	-0.472***	-0.492***
Importing country is an island	1.177***	1.338***	4.027***	4.380***
Exporting country is landlocked	-1.402***	-1.315***	-1.251***	-1.166***
Importing country is landlocked	-0.236	-0.300	0.421**	0.429**
Third-wave provisions of PTAs	1.222	1.748	-1.328	-2.067
Andean1	5.702***	5.545***	5.123***	5.123***
Andean2	-8.696***	-8.485***	2.428**	3.014***
Andean3	-1.503***	-1.461***	9.368***	9.130***
APEC1 ^c	0.929***	1.032***	0.931***	1.029***
APEC2 ^c	-0.839***	-0.817***	-6.988***	-7.133***
APEC3 ^c	3.062***	2.994***	3.500***	3.414***
EFTA1	0.203	0.796	3.117	2.963
EFTA2	12.657***	14.513***		
EFTA3	11.631***	11.115***	8.552***	8.364***
EU1	-3.231***		3.626***	5.341***
EU2	5.184***	6.189***	1.118**	2.207***
EU3	-0.572	-0.289	-5.442***	-5.031***
GCC1 ^e	2.348	2.406	2.210	2.217
GCC2 ^e	-0.537	0.173		-1.452
GCC3 ^e	-0.186**	-0.267***	0.051	-0.026
LAIA1	16.159***	16.046***	17.917***	17.605***
LAIA2	23.259***	24.735***		
LAIA3	15.478***	14.820***	-16.838***	-16.752***
MERCOSUR1	-1.432	-1.132	3.227	3.169
MERCOSUR2	-8.512***	-8.380***		
MERCOSUR3	1.793***	1.969***	18.830***	18.874***
NAFTA1	-2.800	-2.692	2.551	3.315
NAFTA2	13.591***	14.416***	29.870***	31.170***
NAFTA3	-8.453***	-8.562***	-8.137***	-8.385***
SPARTECA1	17.033***	18.523***	17.702***	19.271***
SPARTECA3	-11.930***	-12.128***	-14.163***	-14.880***
CER1	-8.356**	-8.976	-6.959**	-7.056**

Table 5.3 (continued)

Variable name	Dynamic PTA specific variables—fixed effects ^a		Antimonde PTA specific variables—fixed effects ^{a,b}	
	Without tariff	With tariff	Without tariff	With tariff
CER2	4.704***	5.267**	9.138***	9.382***
CER3	8.616***	8.523***	7.378***	7.396***
EU-Switzerland1	-15.405***	-14.934***	-15.718***	-15.541***
EU-Switzerland3	17.254***	16.723***	21.229***	20.401***
Chile-Colombia1	0.079	0.416	-1.304	-0.673
Chile-Colombia2	3.275***	3.013***	-18.101***	-18.457***
Chile-Colombia3	-0.705	-0.545	4.641***	4.742***
Chile-Mexico1			-4.612	-4.405
Chile-Mexico2			49.545***	50.383***
Chile-Mexico3			4.116***	4.243***
US-Israel1	7.763	10.386*	10.255*	13.218**
US-Israel3	11.120***	10.649***	10.811***	10.341***
Australia-PNG1	-1.166	-2.097	-3.080	-5.104
Australia-PNG3	-1.492*	-1.382	-3.609***	-3.157***
Singapore-NZ1			-0.235	1.892
Singapore-NZ3			1.307***	1.455***
Chile-MERCOSUR1	-3.654*	-3.440	-3.682***	-3.478***
Chile-MERCOSUR2	1.403***	1.321***	-4.953***	-4.780***
Chile-MERCOSUR3	-1.171**	-1.148**	-2.080**	-2.222**
EU-Egypt1	5.573	5.454	5.688**	4.870*
EU-Egypt3	0.682	0.144	0.336	0.006
EU-Poland1	2.517	1.837	-3.763**	-4.049**
EU-Poland2	-1.457***	-0.600*		
EU-Poland3	2.099***	1.956***	10.075***	9.820***
AFTA1	-9.693***	-8.204*	-9.629***	-8.946***
AFTA2	-0.484	0.354		
AFTA3	3.060***	2.827***	0.572	0.222
LR χ^2 (chi2)	32,088.3	28,186.5	33,800.9	29,892.8
Log likelihood	-77,034.8	-71,010.6	-76,178.6	-70,157.4
σ (standard deviation of the error term)	2.859	2.855	2.776	2.766

Source: Authors' estimates.

Notes: Dependent variable = Ln exports; time period = 1998–1997; unbalanced panel; Tobit maximum likelihood estimates.

^aTo save space, the fixed effect coefficients associated with exporting country, importing country, and time are not reported here.

^bIn the antimonde specification, some PTA indexes are dropped because of high multicollinearity between country-fixed effects and PTA indexes.

^cWhile a Member Liberalization Index has not been calculated for APEC (a nonpreferential arrangement) or for the Gulf Cooperative Council (a preferential one), their possible effects on the trade flows of their members have been controlled for through a set of three conventional zero-one dummy variables.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

significant and negative effect on its imports, as expected. The PTA-specific indexes also show a significant effect. This is because they capture not only the existence of tariff preferences (as also captured in the tariff variable) but also the effects of nontariff measures affecting merchandise trade, such as rules of origin.

The coefficients on the PTA-specific indexes are generally not sensitive to the inclusion or exclusion of the tariff variable. So the inclusion of a tariff variable makes little difference to the main findings of this study. But in some cases, negative intrabloc trade effects in the full sample become positive in the smaller sample. And perhaps not surprisingly, while the comparative advantage motivation for trade showed as significant in the full sample (with a positive and significant coefficient on the difference in per capita GDP), this is not the case in the restricted sample. These differences also show that what constitutes “normal” trade is conditioned by how many countries and years are in the sample—those studies with restricted time and country coverage, particularly where it is restricted to high-income developed countries, are likely to have results biased accordingly.

The nonmerchandise provisions show a positive (complementary) relationship with trade when PTA indexes are defined dynamically. Thus favorable investment and services trade provisions in PTAs can enhance merchandise trade between member countries once the agreement is in operation.

In summary, the main result is that PTAs are not as relatively benign as previous studies have indicated. After controlling for country- and time-specific effects and the degree of liberalization of merchandise trade provisions in an unrestricted sample and testing explicitly for whether the trade effects are significantly different after PTA formation than before, most PTAs were estimated to have negative trade creation. Other recent empirical assessments have shown a more optimistic outlook for trade in preferential agreements by ignoring these analytical issues.

One potential puzzle is that the intrabloc effect is found to be negative for apparently more comprehensive and liberal PTAs—EU, CER, NAFTA, MERCOSUR and some of the recent bilateral agreements. One possible reason is that total elimination of tariffs among members, as required by the General Agreement on Tariffs and Trade (GATT) Article XXIV, may not be optimal for members. For example, Frankel, Stein, and Wei (1995) found that a 22 percent reduction in tariffs below multilateral tariff levels may instead be optimal. However, this finding is driven by welfare effects, not by trade volume effects.

A more likely explanation is that, although the merchandise MLI used in this study has attempted to capture the potentially trade-restrictive effects of the nontariff merchandise trade provisions embodied in PTAs, it has not always captured them adequately. For example, as noted earlier, the merchandise MLI has treated the trade restrictive effects of rules of origin

as being additive to and independent from the other provisions of PTAs. In reality, not only are rules of origin restrictive, they are also likely to neutralize or even reverse the trade effects of other provisions that are apparently quite liberal.¹⁴

The way in which rules of origin can operate in practice to counter the effects of other provisions that are apparently quite liberal can be seen most clearly in the case of NAFTA. There, the rules of origin are relatively complex—the specification of requirements for minimum change in tariff heading vary product by product and take up several hundred pages. Further, they are strictly enforced. The domestic content rules applied in the EU are also relatively complex. Even if the tariffs on each product are eliminated entirely (an apparently quite liberal provision), the complex rules of origin governing the sourcing of inputs to qualify for the tariff concession on output can undo the liberal effect of the tariff concession on output. This is not recognized in the MLI, which treats tariff provisions and rules of origin additively, not interactively. Thus, the MLI may overstate the effective amount of liberalization in agreements with complex rules of origin, explaining why it was that the *apparently* more comprehensive and liberalized PTAs were found to have a negative intrabloc effect, relative to average trade patterns in the sample.

Sensitivity Analysis

The preceding model specification differs from standard specifications in several respects. One is the inclusion of the sum of importing and exporting country GDPs, rather than each country's GDP separately. As noted, this comes from a model of product differentiation originating with Helpman and Krugman (1985) and Helpman (1987). But in the original model, the dependent variable was total bilateral trade—the sum of exports in both directions—rather than bilateral exports. This paper has followed Egger (2001) and others by using bilateral exports as the dependent variable to allow a more refined examination of the trade diversion issue. And the resulting estimate of the coefficient on the sum of GDPs is very similar to that of Egger (2001). But the question arises whether the results are sensitive to this treatment of GDPs, especially given the redefinition of the dependent variable. To test this, the preceding model was reestimated with (the log of) GDPs of importing and exporting country entered separately. The coefficient on the exporting country's GDP was 1.118, and the coefficient on the importing country's GDP was 0.766, with the sum being close to the result in table 5.1. In all other respects, the results were similar to those shown in table 5.1. The results are available on request from the authors.

14. For analyses of the welfare effects of rules of origin, see Duttagupta and Panagariya (2003), Krueger (1999a), Ju and Krishna (1998), and Krishna and Krueger (1995).

The specification also differs from some others by using triple-indexed fixed effects, controlling separately for importing country, exporting country, and time-related unobservable effects. This is in contrast to specifications that use a single country-pair fixed effect as well as a time effect. As noted, the triple-indexed approach was advocated by Matyas (1997, 1998) for econometric reasons. It is also the approach needed to control for the misspecifications identified by Haveman and Hummels (1998) and Anderson and van Wincoop (2003). Haveman and Hummels note that total exports are likely to be a better measure of “economic mass” than GDP. Gravity models are likely to be misspecified when bilateral exports grow faster than GDP simply because total exports grow faster than GDP, not because of some PTA effect. The extent to which total exports grow faster than GDP is an individual country effect, not a country-pair effect. Similarly, Anderson and van Wincoop (2003) note that in the correct theoretical specification, bilateral trade flows should depend on three measures of trade barriers—the bilateral trade barrier between the two countries and each country’s resistance to trade with all regions. Again, the two latter resistance effects are country effects, not country-pair effects.

It was not possible to test the sensitivity of the preceding Tobit specification to the inclusion of a country-pair fixed effects because there were too many country-pair groups, preventing estimation. The approach also risks overspecification, with the country-pair effects duplicating much of the work of the PTA dummies.

A final piece of sensitivity analysis is the inclusion of a lagged dependent variable. Appropriate econometric estimation of such a specification in a Tobit context with fixed effects is unlikely to have been possible on a data set of this size. In many other contexts, ordinary least squares (OLS) would be an acceptable alternative to Tobit estimation in practice, with the results not differing greatly between the two estimation methods. Were this the case here, it would have been possible to test sensitivity to the inclusion of a lagged dependent variable in an OLS context. However, Tobit estimation matters greatly here, as would be expected with 44 percent of the observations on the dependent variable being zero. The OLS estimation of the original triple indexed specification on the full sample (with zero export values replaced by small positive numbers) led to results with a coefficient on the sum of GDPs being unreasonably low, at 0.555, and similar downward bias in other coefficients, including those on distance, currency union, and the PTA dummies. Thus the results of further sensitivity analysis on the OLS specification were judged unreliable. For what it is worth, adding a lagged dependent variable produced short-run coefficients even lower than the already low OLS estimates, while the value of the coefficient on the lagged dependent variable itself implied long-run coefficients somewhat higher than the OLS estimates. In all other respects, the results were unchanged.

5.3.4 New Evidence on Investment Creation and Diversion

The observable effects—normal bilateral investment determinants and traditional and third-wave provisions of PTAs—and unobservable country- and time-specific effects all significantly influence the bilateral stock of outward investment (see table 5.4).¹⁵

The signs of coefficients on the normal bilateral investment determinants are generally as expected. Larger absolute differences in per capita GDP are associated with outward FDI being lower than otherwise. Because absolute differences in per capita GDP also boost bilateral exports, the results lend some support to the idea that trade and FDI are substitutes when trade is motivated by differences in factor endowments (Markusen 1983). Only if investment treaties are enacted between countries do they have a significant positive effect on outward investment. When they are signed but not enacted, they tend to suppress outward investment. The presence of a currency union has no significant effect on outward FDI, although it had a significant and positive effect on bilateral exports.

PTAs have been categorized in table 5.5 according to whether investment responds in either tariff-jumping or beachhead fashion to the trade provisions or whether it responds instead primarily to the nonmerchandise trade provisions. A single PTA can fall into more than one category.

Only the South Pacific Regional Trade and Economic Cooperation Agreement (SPARTECA) and Andean agreements showed weak evidence of investment behavior responding in beachhead fashion to the trade provisions of the agreement. SPARTECA is a nonreciprocal agreement between Australia, New Zealand, and selected South Pacific Island countries with few nontrade provisions and with trade provisions only for selected products. But the nonreciprocal tariff preferences may have allowed the Pacific island countries to attract investment, not only from Australia and New Zealand, but also from other countries, to gain preferential access to the CER market.

Empirical evidence is weak for Ethier's more general view that PTA members can attract investment from nonmember countries, once other observable and unobservable factors are controlled for. In four agreements the trade provisions appear to have encouraged inward FDI from third parties—the Andean Pact, SPARTECA, ASEAN Free Trade Agreement (AFTA) and NAFTA—but in no case was the effect significant, and only in the first two agreements was the effect also associated with an increase in intrabloc trade (defining beachhead investment). In three agreements, it was the nontrade provisions that encouraged inward FDI from third parties—NAFTA, EU, and SPARTECA—but only in the first case was the effect significant. Thus when a PTA had a significant effect on inward FDI

15. The results in table 5.4 differ slightly from those in earlier versions of this paper.

Table 5.4 Gravity model of investment

Variable name ^c	Dynamic PTA indexes	
	Without fixed ^b	With fixed ^{a,b}
Ln sum of exporting and importing countries GDP (SUM)	1.496***	1.152***
Similarity in exporting and importing country's GDPs	0.086	-0.058
Ln of absolute differences in per capita GDPs of exporting and importing country	-0.487**	-0.603***
Ln distance	-0.682***	-0.572***
Ln bilateral real exchange rate	0.335**	-0.242*
Ln tariff	0.013	0.001
Linguistic similarity	0.000***	0.000***
Colonial	1.601***	1.285***
Border	0.615**	0.595**
Currency union	-0.716	-1.162
Home country is an island	-3.577***	-3.212***
Host country is an island	-1.004***	-1.174***
Home country is landlocked	-2.842***	-3.122***
Host country is landlocked	0.396	0.583**
Investment treaties signed	-1.525***	-1.782***
Investment treaties enacted	1.029**	1.055**
M-ANDEAN1	-4.664	-6.121
M-ANDEAN2	0.410	
M-ANDEAN3	0.557	
M-APEC1 ^d	-0.117	-0.179
M-APEC2 ^d	-0.555**	-0.259
M-APEC3 ^d	-1.339***	-2.255***
M-EFTA1	3.090	1.433
M-EFTA2	1.139	
M-EFTA3	1.837*	
M-EU1	0.537	1.662
M-EU2	-0.559	-2.172
M-EU3	0.111	
M-NAFTA2	2.277	
M-NAFTA3	2.995	
M-SPARTECA1	11.279	10.430
M-SPARTECA2	3.309	
M-SPARTECA3	-8.729***	
M-CER1	-1.720	-0.026
M-CER2	-2.385	
M-CER3	10.641***	
M-US-Israel2	-6.742**	
M-US-Israel3	-5.725	
F-ANDEAN1	1.234	5.590
F-ANDEAN2	-11.326**	-10.022**
F-ANDEAN3	-8.575	-19.530*
F-APEC1 ^d	0.095	0.003
F-APEC2 ^d	0.813***	0.976***
F-APEC3 ^d	0.553*	1.820***
F-EFTA1	14.308	-12.698
F-EFTA2	-16.474	-39.673***
F-EFTA3	138.905***	170.216***
F-EU1	1.812	2.471
F-EU2	2.136	0.795

Table 5.4 (continued)

Variable name ^c	Dynamic PTA indexes	
	Without fixed ^b	With fixed ^{a,b}
F-EU3	12.840***	17.634***
F-NAFTA1	-9.503*	-5.457
F-NAFTA2	5.802***	3.147*
F-NAFTA3	13.650***	10.317***
F-SPARTECA2	126.591	148.125
F-CER1	-0.647	3.999
F-CER2	1.617	-2.509
F-CER3	31.455***	23.847***
F-United States-Israel1	6.865	-8.098
F-United States-Israel2	-0.754	8.118
F-United States-Israel3	21.867*	54.425***
LR χ^2 (chi2)	1,113.8	1,444.7
Log likelihood	-2,192.9	-2,027.5
σ (standard deviation of the error term)	1.875	1.609

Source: Authors' estimates.

Notes: Dependent variable = Ln stock of outward investment; time period = 1988–1997; unbalanced panel; Tobit maximum likelihood estimates.

^aTo save space, the fixed effect coefficients associated with home country, host country, and time are not reported here.

^bSome PTA indexes are dropped because of high multicollinearity among explanatory variables.

“M” before each PTA name denotes index of traditional merchandise trade provisions, and “F” before each PTA name denotes index of new age provisions.

^dWhile a Member Liberalization Index has not been calculated for APEC (a nonpreferential arrangement), its possible effects on the trade flows of its members has been controlled for through a set of three conventional zero-one dummy variables.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 5.5 Main drivers of investment in PTAs

No measurable impact	Tariff-jumping effects of trade provision	Beachhead effects of trade provisions	Nontrade provisions
MERCOSUR ^a AFTA ^a	EFTA CER	SPARTECA ^b Andean ^b	Andean EFTA EU NAFTA CER United States-Israel

Source: Table 5.4.

^aPTA indexes for these agreements were dropped because of high multicollinearity among the explanatory variables.

^bOnly weak evidence for this characterization because the coefficients are not significant.

from third parties, it was in response to the nontrade rather than the trade provisions. It appears that Ethier's beachhead investment is not an important phenomenon empirically.

The EFTA and CER agreements showed evidence of investment behavior consistent with an unwinding of tariff-jumping behavior. But both of these agreements also showed significant evidence of investment responding to the nontrade provisions of the agreement. A total of five of the nine PTAs examined for investment effects showed significant evidence of investment responding to the nontrade provisions of the agreements.

While table 5.5 indicates investment responses to the traditional and third-wave provisions of PTAs, it does not indicate whether PTAs caused investment creation or investment diversion per se. A summary of the signs of the significant coefficients is reported in table 5.6.

As noted, the trade provisions of PTAs did not result in a significant increase in investment from nonmembers in any PTA. Trade provisions caused a reduction in outward investment (investment diversion) in SPARTECA but an increase in outward investment in EFTA and CER consistent with the unwinding of tariff-jumping investment.

The new age provisions in various PTAs have had a more widespread impact on investment than the trade provisions. The NAFTA agreement was estimated to have reduced investment among members. All other agreements considered had no significant effects on investment among members.

While NAFTA attracted investment from nonmembers, particularly into Mexico, the new age provisions in the Andean Pact and EFTA were unable to attract investment from nonmembers. For EFTA, the loss of membership to EU made it a less attractive place for foreign direct investment. This is similar to the findings of Baldwin, Forslid, and Haaland (1995), who found that EC caused a diversion of third-country capital

Table 5.6 New evidence on investment creation and diversion

Trade provisions		Third-wave provisions		
Extra-PTA (inward)	Extra-PTA (outward)	Intra-PTA	Extra-PTA (inward)	Extra-PTA (outward)
United States-Israel(-)	EFTA(+) CER(+) SPARTECA(-)	NAFTA(-)	Andean(-) EFTA(-) NAFTA(+)	Andean(-) EFTA(+) EU(+) NAFTA(+) CER(+) United States-Israel(+)

Source: Table 5.4, significant coefficients.

Note: Positive (+) symbol denotes investment creation, and negative (-) symbol denotes investment diversion.

from EFTA to the EU. The Andean Pact was similarly affected by the loss of Peru, although the findings for Latin American countries are also affected by incompleteness of the FDI outstock data.

The sum of significant coefficients on the indexes of third-wave provisions for each PTA can provide an indicative measure of the impact of these new age provisions on net investment creation (see table 5.7). Of the nine PTAs examined for investment effects, five showed positive net investment effects. Only the Andean Pact caused net investment diversion. This agreement apparently caused a reduction in both inward and outward investment with third parties, without succeeding in causing a significant boost in intrabloc investment. As noted, this result may in part reflect incompleteness in the FDI outstock data for Latin America.

The North American Free Trade Agreement, EU, EFTA, CER, and the US-Israel agreement caused net investment creation, not because they stimulated investment among members, but primarily because they appear to have stimulated outward investment from member to nonmember countries. This is consistent with some of the nontrade provisions of these agreements being nonpreferential in nature. It is also consistent with some of these regions being major sources of FDI, but this suggests that the estimated effects may well reflect the influence of causal factors not controlled for in the analysis that make these countries net capital exporters, rather than the effects of PTA formation and expansion per se. One example is financial deregulation and the growth of superannuation funds that have encouraged Australia to become a much more important capital exporter recently (Battellino 2002).

The South Pacific Regional Trade and Economic Cooperation Agreement had no significant impact on net investment creation, while the effects of MERCOSUR and AFTA on investment could not be distinguished because of problems of multicollinearity.

Though the investment results appear to be more positive than the results reported for trade, there are number of qualifications that need to be considered. Winters (1997) argued that new FDI from any source could go

Table 5.7 Net impact of PTAs' third-wave provisions on investment

Net investment creation	Net investment diversion	No measurable impact
EFTA	Andean	MERCOSUR ^a
EU		AFTA ^a
NAFTA		SPARTECA
CER		
United States-Israel		

Source: Table 5.4, fixed effects estimates, significant coefficients.

^aPTA indexes for these agreements were dropped because of high multicollinearity among the explanatory variables.

into the production of goods for trade diversion and thus worsen the PTA's welfare overall. In similar tone, McLaren (2002) argued that

A regional trade regime can plausibly be interpreted as a coordination failure, in which the anticipation that the world will break into regional trade blocs induces sunk private sector investments that then lead to a demand for regionalism. Under this argument, regionalism can be Pareto-worsening even though once sunk investments have been made it is, *ex post*, a relatively efficient compromise: hence, regionalism is "insidious," the damage it does to efficiency is hidden in the distortion of *ex ante* investments. (McLaren 2002, 572)

The gravity model estimates provide indications of the positive net investment effects of PTAs but do not consider whether the resulting investment contributes to trade diversion.

Further, as noted before, a finding of net investment creation is a weak indicator of whether the welfare gains from investment creation outweigh the costs of investment diversion. Investment diversion may dominate creation in welfare terms, even if it does not in "volume of investment" terms. On the other hand, if the nontrade provisions reduce restrictions that raise costs, member countries can gain in welfare terms, despite investment diversion. But in either case, members could well gain even more from multilateral liberalization of nontrade restrictions.

5.4 Summary

Theoretical work has always highlighted that while the merchandise trade provisions of PTAs can boost trade among members, this is often at the expense of nonmembers. So whether it benefits a country to join a PTA depends on the cost structures in partner countries, compared with the cost structures in third parties. If a preferential trade arrangement diverts a country's imports from a low-cost third party to a higher-cost preferential trade partner, it can be made worse off. Conversely, the opportunity for benefits is greater where the PTA partner is at world-best competitiveness and where liberalization under the PTA encourages imports from that source.

The new empirical work outlined in this paper suggests that of the eighteen recent PTAs examined in detail, twelve have diverted more trade from nonmembers than they have created among members. What is more, some of the apparently quite liberal PTAs—including EU, NAFTA and MERCOSUR—have failed to create significant additional trade among members (relative to the average trade changes registered among countries in the sample).

Part of the reason for this more negative finding than in previous studies is the rigorous statistical test that has been applied to ascertain whether

intrabloc trade is significantly greater after bloc formation (or expansion) than before. In the past, this was assessed, at best, only by reference to the point estimates from various cross sections. But the finding is also consistent with the observation that many of the provisions needed in preferential arrangements to underpin and enforce their preferential nature—such as rules of origin—are in practice quite trade restricting.

While the increasing focus of PTAs on nontrade provisions may suggest that conventional concerns about trade diversion are outmoded, some theoretical literature suggests this conclusion would be premature.

On the one hand, in an increasingly integrated world economy, even minor *trade* concessions can have a significant impact on *investment* flows. And if investment is attracted into one PTA partner in order to serve the markets of the others, then the trade from such beachhead positions can constitute traditional trade diversion.

On the other, the *nontrade* provisions of PTAs, particularly those related to investment and services, can also have a significant impact on *investment* flows. But the preferential nature of the PTA provisions may mean that investment is diverted from a low-cost to a higher-cost host country, and such investment diversion can also be harmful.

The empirical work in this paper finds little evidence of beachhead investment, or an unwinding of tariff-jumping investment, in response to the trade provisions of PTAs. Only for SPARTECA and the Andean Pact, for example, is there (weak) evidence of FDI responding in beachhead fashion to trade provisions. And only for EFTA and CER is there some evidence of an unwinding of tariff-jumping investment.

There is evidence that FDI responds significantly to the nontrade provisions of PTAs. And in five of the nine PTAs examined for investment effects, the nontrade provisions led to net investment creation.

Although it is a weak test, this suggests that on balance, the nontrade provisions of these PTAs have created an efficient geographic distribution of FDI. This is consistent with the fact that at least some of the nontrade provisions (e.g., commitments to more strongly enforce intellectual property rights) are not strongly preferential in their nature.

Further, the theoretical literature has stressed that if the nontrade barriers are of the sort to raise the real resource costs of doing business, rather than simply to create rents that raise prices above costs, then preferential liberalization will be beneficial, even in the absence of net investment creation.

However, the trade that may be generated from the new FDI positions may still be diverted in the wrong direction in response to the trade provisions of PTAs and may therefore contribute to the net trade diversion also found here.

Thus the results of this paper suggest that there may be economic gains from the nontrade provisions of third-wave PTAs, but they also suggest

that there are still economic costs associated with the preferential nature of the trade provisions. And these costs could be magnified in a world of increasing capital mobility.

Thus the findings of this research on the effects of the nontrade provisions of PTAs are more positive than those on the trade provisions. This suggests there could be real benefits if countries could use regional negotiations to persuade trading partners to make progress in reforming such things as investment, services, competition policy, and government procurement, especially if this is done on a nonpreferential basis.

Appendix

The Static Welfare Effects of PTAs

In figure 5A.1, S_a and D_a are the domestic supply and demand curves in country A . S_b is the supply curve of imports from the PTA partner country, showing that any quantity can be supplied from there at the price P_b . S_w is the supply curve of imports from the rest of the world, showing that any quantity can be supplied from there at price P_w . P_a^* is the initial, tariff-inflated price in country A , with the tariff t equal to $P_a^* - P_w$. Initially all imports $Q_c - Q_p$ come from the rest of the world, as with the same tariff t placed on imports from B , the local price in country A would exceed P_a^* . The tariff revenue on the imports from the rest of the world is $AEJF$. The quantity produced domestically is Q_p , and domestic consumption is Q_c .

Now suppose that country A eliminates its tariff on imports from B but

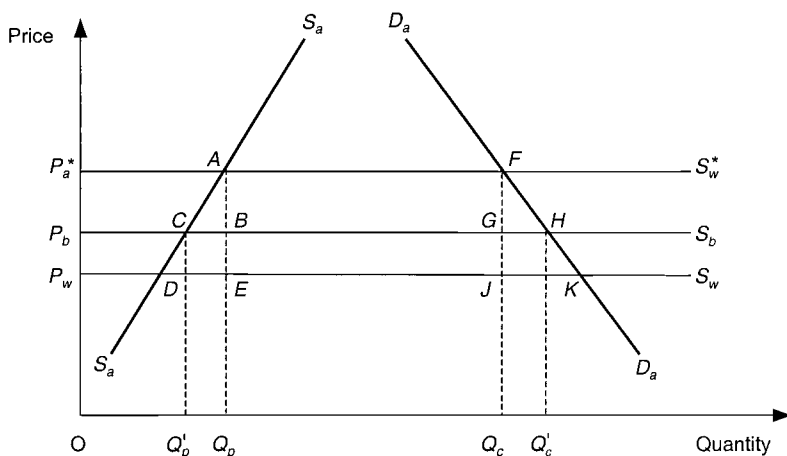


Fig. 5A.1 An illustration of trade creation and diversion effects of a PTA

retains it on imports from the rest of the world. With imports now available from B at P_b , the import quantity expands to $Q'_c - Q'_p$, with country B rather than the rest of the world becoming the source. Tariff revenue shrinks to zero. Domestic production shrinks to Q'_p , and domestic consumption expands to Q'_c .

The net effect of PTA formation on economic well-being in country A is given by $ABC + FGH - BEJG$. The first effect, the gain of $ABC + FGH$, is the net benefit to consumers and the net resource saving in production from having domestic production shrink from Q_p to Q'_p and consumption expand from Q_c to Q'_c . This is the trade creation gain from shifting high-cost domestic production to a lower-cost partner.¹⁶ The second effect, the loss of $BEJG$, is that portion of the tariff revenue lost by shifting imports from the rest of the world to the higher-cost partner that is not recouped in lower domestic prices to consumers. It is the welfare loss from trade diversion and arises essentially because forgone domestic tariff revenue accrues instead as profit to producers in the partner country. The effect on country A is ambiguous a priori. Strictly speaking, only if the partner country is already at world-best production cost is a welfare gain to country A assured. But then A 's economic motive for preferential rather than nondiscriminatory trade liberalization is unclear.

What about the welfare effects on the country receiving the preferential tariff concession and the effects on the rest of the world? Both face a change in demand for their product from country A , but because of the assumption of constant costs, there is no induced change in unit costs that can flow on to benefit domestic consumers or drive an improvement in resource allocation in those countries.¹⁷ Thus, the effect on country A , the country granting the tariff preference, is the same as the welfare effect on the PTA and the world as a whole. This highlights one of the key weaknesses of the simple analysis—its assumption of constant costs of production in the partner country and in the rest of the world.

The simple analysis is nevertheless useful for outlining the nature of empirical tests for trade creation and trade diversion. Typically, these tests measure the amount by which the volume (or more often, the value) of trade increases with partner countries— $Q'_c - Q'_p$ in the preceding example—and compare it with the amount by which trade with the rest of the world is reduced— $Q_c - Q_p$ in the preceding example. If the net effect is positive, it is still only a weak test of whether the gains from trade creation outweigh the costs of trade diversion. It establishes that there is some pos-

16. Viner's (1950) original analysis omitted the consumption gain FGH . Johnson (1960) was the first to include it as part of the gains from trade creation, thereby ending unproductive debates about the possibility of welfare-increasing trade diversion (Gehrels 1957; Lipsey 1957; Michaely 1976).

17. If there is a preexisting distortion in the exporting sector of the exporting country, then an expansion of that sector could worsen the allocation of resources.

itive width to the triangles ABC and FGH , but it does not establish that their areas exceed that of $BEJG$. This also depends on the reduction in costs per unit of newly created trade and the increase in costs per unit of diverted trade. What can be concluded in this model is that if the empirical tests establish net trade creation in a volume or value sense, then the PTA may still have generated welfare losses, but if the empirical tests establish net trade diversion, then the PTA cannot have created welfare gains.

The assumption of constant costs in the partner country and in the rest of the world is consistent with perfect competition in those two markets. There has been a great deal of analysis examining the welfare effects of allowing unit production costs to vary in those two markets (see Panagariya [2000] for a summary), although the analysis has not always been explicit about the nature or source of the less-than-perfect competition in those markets.

The easiest way to see the dramatic effects that less-than-perfect competition can have is to imagine in figure 5A.1 that the producers in country B form a cartel and price up to the world price plus external tariff after they are granted the tariff preference. Their price would remain at P_a^* , the losses to country A from trade diversion would expand to $AEJF$, and the gains to A from trade creation would disappear completely! On the other hand, country B would now have a net gain in rent of $ABGF$ that was previously tariff revenue accruing to A . The net loss to the PTA and the world as a whole would be $BEGJ$. Thus, less-than-perfect competition can preserve the losses from trade diversion but destroy the gains from trade creation.

One of the most general treatments of less-than-perfect competition is by Mundell (1964). He draws the following conclusions on the effects of a customs union, assuming that all goods are gross substitutes and initial tariffs are low:

(1) The discriminatory tariff reduction by a member country improves the terms of trade of the partner country with respect to both the tariff reducing country and the rest of the world, but the terms of trade of the tariff-reducing country might rise or fall with respect to third countries.

(2) The degree of improvement in the terms of trade of the partner country is likely to be larger the greater is the member's tariff reduction; this establishes the presumption that a member's gain from a free-trade area will be larger the higher are initial tariffs of partner countries. (Mundell 1964, 8)

A key to this result is the revenue transfer effect that can arise with less-than-perfect competition. It is also the basis for Panagariya's (1999) conclusion, for example, that the United States is likely to gain, but that Mexico could lose, from NAFTA.

References

- Adams, R., P. Dee, J. Gali, and G. McGuire. 2003. The trade and investment effects of preferential trading arrangements—Old and new evidence. Productivity Commission Staff Working Paper. Canberra, Australia: Australian Government, Productivity Commission, May.
- Aitken, N. 1973. The effect of the EEC and EFTA on European trade: A temporal cross-section analysis. *American Economic Review* 63 (5): 881–92.
- Anderson, J., and E. van Wincoop. 2003. Gravity with gravitas: A solution to the border puzzle. *American Economic Review* 93 (1): 170–92.
- Andriamananjara, S. 2002. On the size and number of preferential trading arrangements. *Journal of International Trade and Economic Development* 11 (3): 279–95.
- Bagwell, K., and R. Staiger. 1998. Will preferential arrangements undermine the multilateral trading system? *Economic Journal* 108 (449): 1162–82.
- . 1999. An economic theory of the GATT. *American Economic Review* 89 (1): 215–48.
- Baier, S. L., and J. H. Bergstrand. 2001. The growth of world trade: Tariffs, transport costs and income similarity. *Journal of International Economics* 53 (1): 1–27.
- Balassa, B. 1967. Trade creation and trade diversion in the European common market. *Economic Journal* 77 (1): 1–21.
- Baldwin, R. 1994. Towards an integrated Europe. London: Centre for Economic Policy Research.
- . 1996. A domino theory of regionalism. In *Expanding European regionalism: The EU's new members*, ed. R. E. Baldwin, P. Haapraanta, and J. Kiander, 25–48. Cambridge, UK: Cambridge University Press.
- Baldwin, R. E., R. Forslid, and J. Haaland. 1995. Investment creation and investment diversion: Simulation analysis of the single market program. CEPR Discussion Paper no. 1308. London: Centre for Economic Policy Research.
- Baldwin, R. E., and A. J. Venables. 1995. Regional economic integration. In *Handbook of international economics*. Vol. 3, ed. G. Grossman and K. Rogoff, 1597–1644. Amsterdam: Elsevier.
- Battellino, R. 2002. Australia as a capital exporter. Paper presented at conference, The Impact of an Australian–US Free Trade Agreement: Foreign Policy Challenges and Economic Opportunities, 29–30 August, Canberra, Australia.
- Bayoumi, T., and B. Eichengreen. 1995. Is regionalism simply a diversion? Evidence from the evolution of the EC and EFTA. CEPR Discussion Paper no. 1294. London: Centre for Economic Policy Research.
- Bhagwati, J. 1999. Regionalism and multilateralism: An overview. In *Trading blocs: Alternative approaches to analyzing preferential trade agreements*, ed. J. Bhagwati, P. Krishna, and A. Panagariya, 3–32. Cambridge, MA: MIT Press.
- Bhagwati, J., P. Krishna, and A. Panagariya. 1999. *Trading blocs: Alternative approaches to analyzing preferential trade agreements*. Cambridge, MA: MIT Press.
- Boisso, D., and M. Ferrantino. 1997. Economic distance, cultural distance, and openness in international trade: Empirical puzzles. *Journal of Economic Integration* 12 (4): 456–84.
- Bond, E., and C. Syropoulos. 1996. The size of trading blocks, market power and world welfare effects. *Journal of International Economics* 40 (3–4): 411–37.
- Clarete, R., C. Edmonds, and J. Wallack. 2003. Asian regionalism and its effects on trade in the 1980s and 1990s. *Journal of Asian Economics* 14 (1): 91–129.
- Clark, X., and J. Tavares. 2000. A quantitative approach using the gravity equation.

- Development Discussion Paper no. 748. Cambridge, MA: Harvard Institute for International Development, Harvard University.
- Corden, M. 1972. Economies of scale and customs union theory. *Journal of Political Economy* 80 (3): 465–75.
- Deardorff, A. V., and R. Stern. 1994. Multilateral trade negotiations and preferential trading arrangements. In *Analytical and negotiating issues in the global trading system*, ed. A. V. Deardorff and R. Stern, 53–85. Ann Arbor: University of Michigan Press.
- Deardorff, A. V. 1998. Determinants of bilateral trade flows: Does gravity work in a neoclassical world. In *The regionalization of the world economy*, ed. J. A. Frankel, 23–28. Chicago: University of Chicago Press.
- De Rosa, D. A. 1998. Regional integration arrangements: Static economic theory, quantitative findings and policy guidelines. World Bank. Mimeograph. <http://www.worldbank.org/research/projects/regional.htm>.
- Duttagupta, R., and A. Panagariya. 2003. Free trade areas and rules of origin: Economics and politics. IMF Working Paper no. WP/03/229. Washington, DC: International Monetary Fund.
- Egger, P. 2001. European exports and outward foreign direct investment: A dynamic panel data approach. *Weltwirtschaftliches Archiv* 137 (3): 427–49.
- Eichengreen, B., and D. A. Irwin. 1995. Trade blocs, currency blocs and the reorientation of world trade in the 1930s. *Journal of International Economics* 38 (1–2): 1–24.
- Ethier, W. 1998a. The new regionalism. *Economic Journal* 108 (449): 1149–61.
- . 1998b. Regionalism in a multilateral world. *Journal of Political Economy* 106 (6): 1214–45.
- . 1999. Multilateral roads to regionalism. In *International trade policy and the Pacific Rim*, ed. J. Piggott and A. Woodland, 131–52. New York: St. Martin's.
- . 2001. The new regionalism in the Americas: A theoretical framework. *North American Journal of Economics and Finance* 12 (2): 159–72.
- Evenett, S. J., and W. Keller. 1998. On theories explaining the success of the gravity equation. *Journal of Political Economy* 110 (2): 281–316.
- Feenstra, R. C., J. R. Markusen, and A. K. Rose. 2001. Using the gravity equation to differentiate among alternative theories of trade. *Canadian Journal of Economics* 34 (2): 430–47.
- Fink, C., and C. A. Primo Braga. 1999. How stronger protection of intellectual property rights affects international trade flows. World Bank Working Paper no. 2051. Washington, DC: World Bank. <http://www.worldbank.org/html/dec/Publications/Workpapers/wps2000series/wps2051/wps2051.pdf>.
- Frankel, J. 1997. *Regional trading blocs in the world economic system*. Washington, DC: Institute for International Economics.
- Frankel, J., E. Stein, and S. Wei. 1995. Trading blocs and the Americas: The natural, the unnatural, and the super-natural. *Journal of Development Economics* 47 (1): 61–95.
- Freund, C. 2000. Different paths to free trade: The gains from regionalism. *Quarterly Journal of Economics* 115 (4): 1317–41.
- Fukao, K., T. Okubo, and R. M. Stern. 2003. An econometric analysis of trade diversion under NAFTA. *North American Journal of Economics and Finance* 14: 3–24.
- Gehrels, F. 1957. Customs union from a single-country viewpoint. *Review of Economic Studies* 24 (1): 61–64.
- Gilbert, J., R. Scollay, and B. Bora. 2001. Assessing regional trading arrangements

- in the Asia-Pacific. Policy Issues in International Trade and Commodities Study Series no. 15. Geneva: United Nations Conference on Trade and Development.
- Greene, W. 2002. The behaviour of fixed effects estimator in non-linear models. New York University, Department of Economics, Stern School of Business. Mimeograph.
- Haveman, J., and D. Hummels. 1998. Trade creation and trade diversion: New empirical results. *Journal of Transnational Management Development* 3 (2): 47–72.
- Helpman, E. 1987. Imperfect competition and international trade: Evidence from fourteen industrial countries. *Journal of the Japanese and International Economies* 1 (1): 62–81.
- Helpman, E., and P. Krugman. 1985. *Market structure and international trade*. Cambridge, MA: MIT Press.
- Hertel, T., T. Walmsley, and K. Itakura. 2001. Dynamic effects of the “new age” free trade agreement between Japan and Singapore. *Journal of Economic Integration* 16 (4): 446–84.
- Johnson, H. 1960. The economic theory of customs union. *Pakistan Economic Journal* 10 (1): 14–32.
- Ju, J., and K. Krishna. 1998. Firm behaviour and market access in a free trade area with rules of origin. NBER Working Paper no. 6857. Cambridge, MA: National Bureau of Economic Research, December.
- Krishna, K., and A. O. Krueger. 1995. Implementing free trade areas: Rules of origin and hidden protection. NBER Working Paper no. 4983. Cambridge, MA: National Bureau of Economic Research.
- Krishna, P. 1998. Regionalism and multilateralism: A political economy approach. *Quarterly Journal of Economics* 113 (1): 227–51.
- Krueger, A. 1999a. Free trade agreements as protectionist devices: Rules of origin. In *Trade theory and econometrics: Essays in honour of John S. Chipman*, ed. R. M. James, C. M. James, and R. Raymond, 91–102. New York: Routledge.
- . 1999b. Trade creation and trade diversion under NAFTA. NBER Working Paper no. 7429. Cambridge, MA: National Bureau of Economic Research, December.
- Krugman, P. 1993. Regionalism versus multilateralism: Analytical notes. In *New dimensions in regional integration*, ed. J. de Melo and A. Panagariya, 58–79. Cambridge, UK: Cambridge University Press.
- Levy, P. 1997. A political-economic analysis of free-trade agreements. *American Economic Review* 87 (4): 506–19.
- Li, Q. 2000. Institutional rules of regional trade blocs and their impact on international trade. In *The political consequences of regional trade blocks*, ed. R. Switky and B. Kerremans, 85–118. London: Ashgate.
- Lipsey, R. 1957. Mr. Gehrels on customs unions. *Review of Economic Studies* 24 (2): 211–14.
- . 1958. *The theory of customs unions: A general equilibrium analysis*. PhD diss., University of London.
- Markusen, J. R. 1983. Factor movements and commodity trade as complements. *Journal of International Economics* 14 (3–4): 341–56.
- Matyas, L. 1997. Proper econometric specification of the gravity model. *World Economy* 21 (3): 363–68.
- . 1998. The gravity model: Some econometric consideration. *World Economy* 20 (3): 397–401.
- McLaren, J. 2002. A theory of insidious regionalism. *Quarterly Journal of Economics* 117 (2): 571–608.

- Michaely, M. 1976. The assumptions of Jacob Viner's theory of customs unions. *Journal of International Economics* 6 (1): 75–93.
- Mundell, R. 1964. Tariff preferences and the terms of trade. *Manchester School of Economic and Social Studies* 32:1–13.
- Panagariya, A. 1999. The regionalism debate: An overview. *World Economy* 22 (4): 477–511.
- . 2000. Preferential trade liberalization: The traditional theory and new developments. *Journal of Economic Literature* 38 (2): 287–331.
- Panagariya, A., and R. Duttagupta. 2002. The “gains” from preferential trade liberalisation in the CGE models: Where do they come from? University of Maryland. Mimeograph. <http://www.bsos.umd.edu/econ/panagariya/apecon/techpaper.htm>.
- Pomfret, R. 1997. *The economics of regional trading arrangements*. Oxford, UK: Clarendon Press.
- Riezman, R. 1979. A 3×3 model of customs unions. *Journal of International Economics* 9 (3): 341–54.
- Robinson, S., and K. Thierfelder. 2002. Trade liberalisation and regional integration: The search for large numbers. *Australian Journal of Agricultural and Resource Economics* 46 (4): 585–604.
- Scollay, R., and J. Gilbert. 2000. Measuring the gains from APEC trade liberalisation: An overview of CGE assessments. *World Economy* 23 (3): 175–97.
- Soloaga, I., and L. A. Winters. 2001. Regionalism in the nineties: What effect on trade? *North American Journal of Economics and Finance* 12 (1): 1–29.
- Viner, J. 1950. *The customs union issue*. New York: Carnegie Endowment for International Peace.
- Winters, L. 1997. *Assessing regional integration arrangements*. Washington, DC: World Bank, Development Research Group.
- World Bank. 2000. *Trade blocs*. Washington, DC: World Bank.
- World Trade Flows (WTF). 1997. *World trade flows database 1970–1992*. Cambridge, MA: National Bureau of Economic Research.
- . 2000. *World trade flows 1980–1997* database. Davis: Institute of Governmental Affairs, University of California.
- World Trade Organization (WTO). 1995. *Regional trading arrangements and the world trading system*. Geneva: WTO.
- . 2002. *Work of the Committee on Regional Trade Agreements*. http://www.wto.org/english/tratop_e/region_e/regcom_e.htm.
- Zissimos, B., and D. Vines. 2000. Is the WTO's Article XXIV a free trade barrier? CSGR Working Paper no. 49/00. Warwick, UK: Centre for the Study of Globalisation and Regionalisation, University of Warwick.

Comment Bih Jane Liu

The number of preferential trading arrangements (PTAs) has grown dramatically over the last decade or so. How PTAs affect the trade and investment among member countries and between member and nonmember countries thus becomes a very important issue nowadays. This paper, focus-

ing on this important issue, is very comprehensive and well written. It not only surveys the related theoretical and empirical literature thoroughly, but it also has done a great job in collecting the data and doing the econometric analysis. By using a gravity model to examine eighteen PTAs over the period of 1970–1997, this paper shows that most of the PTAs, including the European Union (EU), the North American Free Trade Agreement (NAFTA), and the Southern Common Market (MERCOSUR), have diverted more trade from nonmember countries than they have created among member countries. In addition, it shows that for most of the agreements where non-trade provisions have affected foreign direct investment (FDI), the PTAs effects have been investment creation rather than diversion.

Although I agree with most of the points that the authors made in this paper, as a discussant I would like to raise some questions.

The first question is related to model specification. This paper uses PTA variables and many country-specific variables to capture the trade relationship among countries. Thus the usual problem of omitted-variable bias in gravity models could be avoided to a certain degree. However, the complicated heterogeneity relationship among member countries and between member and nonmember countries still could not be completely captured by these two kinds of variables, for example, if there are three countries, *A*, *B*, and *C*, in the same trading bloc. *A* and *B* are natural partners, but *A* and *C* are not. Apparently, the trade creation and diversion effects between *A* and *B* would be different from that between *A* and *C*. In this case the PTA variables could not be used to differentiate the different trade effect between *A* and *B* and *A* and *C*. The question then is whether the features of heterogeneity can be captured by the variables such as GDP, real exchange rate, distance, country dummies, linguistic similarity, and so on? To me, the answer seems unclear. Moreover, to derive trade creation or trade diversion effects, the inclusion of PTA dummies in the gravity equation may not be sufficient. This is because PTA dummies reveal little about cost of production across different countries. Some variables related to wage or cost of production and/or the cross terms of these variables with PTA dummies may be needed to understand whether trade from lower-cost PTA members can be created and whether trade is diverted from lower-cost PTA nonmembers to higher-cost PTA members after a trading bloc is formed.

The second question has something to do with the endogeneity problem. We all know that GDP, GDP per capita, and the real exchange rate are endogenous variables that should be determined simultaneously with exports (the dependent variable) within the model. However, in this paper these endogenous variables are used as the right-hand-side variables in the gravity equation (1). As a result, the error term in equation (1) will be correlated with them, and hence the estimators will be biased. To avoid the endogeneity problem, I would suggest that the authors use lag variables or to run simultaneous equations.

This paper utilizes the panel data for the period of 1970–1997 to examine eighteen PTAs. While it achieves maximum degree of freedom, I am wondering whether the results thus obtained will suffer from subsample instability and heteroskedasticity. This is because the relationship between exports and country-specific factors may vary across different trading blocs. A sensitivity analysis may serve to make sure that the results are robust even for subsamples or subperiod samples.

Past studies showed that most PTAs are trade creating in net terms. By contrast, the results from this paper suggest that most PTAs do not create additional trade, either for members or for nonmembers of the agreements. I suspect that it has something to do with the induced investment effect. If PTAs have significant investment creation effect as shown in this paper, then this new investment may be either a substitute for or a complement of trade and will hence affect trade flows indirectly. In the case where investment creation does substitute for intraregional trade, we may observe negative trade creation as shown in this paper. To show whether negative trade effect is indeed caused by investment, I would suggest that the authors run trade and investment equations simultaneously with each variable appearing in the other equation. By doing so, we are able to better understand the interrelationship between trade and investment under PTAs.

One last point is that the wave of globalization together with a reduction in transport costs and an improvement in communication and technologies all make outward direct investment and foreign outsourcing become more prevalent in the 1990s. This in turn changes the nature of international trade from trade in final goods to trade in intermediate inputs. Because the empirical data in this paper cover the time period from 1970 through 1997, how this change in trade structure should be taken into account in the study of the impacts of PTAs will also deserve special attention.

Comment Erlinda Medalla

First of all, let me say that I really enjoyed reading Philippa Dee and Jyothi Gali's paper. The primary reason I chose to discuss this paper is because I wanted to get reacquainted with the topic and get on track with the literature. And on this, the paper does an excellent job.

The paper is very provocative.

I cannot fault the logic (how arguments are developed).

Erlinda Medalla is a senior research fellow of the Philippine Institute for Development Studies.

But to me, there are some jarring notes and conclusions. And I admit there is still a lot I don't completely understand.

My comments come from these two reactions I had, which could be interrelated. Conclusions jar in my mind because I probably did not completely understand the matter.

Hopefully, my comments will help the paper become more understandable to people less technical in orientation and to people who have not done work and are not familiar with the intricacies of the topic.

The Member Liberalization Index

I am not clear on how it is actually estimated. My understanding is that it reflects the coverage of the provisions of the preferential trading arrangement (PTA) (whether or not they are used).

The paper mentions that a high index indicates that the PTA is more liberal to members in the particular area covered by the provisions of the PTA. In nonmerchandise trade, if this is among the provisions of the PTA and its share in total trade is x percent, then I take it that this is the value it takes. (If this is close enough, then my subsequent reading and understanding of the analysis would not be too erroneous.) Nonetheless, I think the authors need to explain this further.

The Gravity Model

I appreciate the distinction made between ex ante CGE analysis and the ex post technique of the gravity model. I think both have merits.

The merit of the gravity model, as mentioned, is that it is ex post and examines actual PTA. The shortcoming is that it only looks at links between the PTA formation and trade outcomes.

My own bias is that simply looking at trade outcomes is very limiting. Judging the benefits merely based on trade effects (whether trade creating or trade diverting) is not enough. As the paper says, there is weak linkage between trade and welfare. And this is precisely what we want to look at when we want to assess if a PTA is beneficial or not.

This is one of the jarring notes I was referring to earlier. Looking at trade diversion and trade creation is based mainly on a static model explained in the appendix. It is not only static, it is also just a (very) partial analysis.

When domestic production contracts because of the PTA, the gain is not only the reduction in the former deadweight loss. The additional gain is the resources this releases (maybe not all because of sunk costs) to relatively more profitable activities. The resulting less-distorted relative prices (as, in effect, the relative prices facing producers with the higher tariff for all trading partners result in higher relative domestic prices of importables) with PTA would mean a better resource allocation, at least from the point of view of the individual member country.

Nonetheless, it is always good to look at a number of approaches but keep in mind their particular limitations.

The Specifications of the Model

It was a bit difficult to figure out. I had to read it several times to understand why it is that the dependent variable is exports from country i to j . Was there another specification that uses imports of country i from country j ? If not, I thought the authors also wanted to measure import diversion? Or are they all mixed together (y is really either export or import). Then I figure that we are dealing with a comprehensive set of data, such that a bilateral export to one is bilateral imports of another? It would help if there is more explanation of this at the outset.

In the explanatory variables, why use the sum of GDP and then the absolute difference rather than use directly separate GDPs of the partner countries? Connected to this, should there be an overall activity level variable, for example, world GDP (or GDP of the ROW)?

The three dummy variables used for each PTA have not been easy to follow and visualize because of the large sample of countries. The first dummy is easier to understand. It takes on the value of the MLI for exports where the bilateral flow is between countries with a PTA, say for exports from Australia to New Zealand, with Closer Economic Relations (CER). Hence, CER1 is the corresponding MLI. In addition, I understand CER2 and CER3 are zero. (Am I right?) For exports involving two countries outside this particular PTA (say exports from Thailand to the Philippines, not members of CER), the three dummies are zero (all CER dummies are zero), but if they have their own PTA (e.g., AFTA), there needs to be an additional three dummies, the first of which will be MLI (AFTA1 dummy value is MLI). Furthermore, for exports from Australia (of CER) to, say, Thailand (of AFTA), then AFTA2 is MLI related to AFTA. And finally, for exports from Thailand to Australia, CER3 is MLI related to CER. (Did I understand it correctly?) This is really complicated, and I congratulate the effort there must have been in dealing with such a huge amount of observations involving numerous variables.

I needed to understand this clearly in order to understand the meaning of the coefficients of the three dummies and why they capture intrabloc trade and import and export diversion, which brings me to my next comment. What is the impact of using MLI rather than zeroes and ones? And in separating exports and imports, does this not introduce an upward bias on trade diversion? (Because standard deviation for the sum adds up.) In the first place, is this the interpretation, simply summing up the coefficients to get the net trade effects? I deduced this from the explanation in the case of Frankel (1997), who uses only two dummies. Is my interpretation correct?

The paper made, toward the end, a qualification on the results about the

intra-PTA effect having large standard error, compared to the extra-PTA effect, which is estimated more accurately. Doesn't this pose a serious limitation? How robust will the estimate be on the net?

Finally, the paper later mentioned (when discussing investments) that because of model convergence problems, the effects of new age provisions on exports to and imports from nonmembers could not be analyzed separately. Does the author have some a priori notions about the direction of effects?

On Investments

I have similar misgivings on this as in the previous analysis on trade outcomes.

Instead of exports, it uses stock of outward investment. Isn't flow the more appropriate variable? (I thought flow is easier to measure than stock.) Accepting this, why outward, not inward, FDI? I would think that for a member country, this is what is important to find out. Again, I may just be missing something here. They are just two sides of the same coin. But using inward FDI would have had more intuitive appeal. It would help if the authors explain why (even in a footnote), as it was not readily obvious to me and possibly to other readers as well (especially nontechnical ones who I think this paper is important to reach).

I tend to share Ethier's (2001) arguments that there is likely "beachhead" response from nonmember countries on FDI. I also agree that some investments could be lost from tariff-jumping-motivated investments. However, in this age of reduced trade barriers, how much of these are left?

Finally, whether there is investment diversion or not, I think the receiving country would always gain (although there may be a reduction in global welfare).

On the whole, I would like to congratulate the authors for such an excellent paper. I learned a lot from it, and it has inspired me to study it more. It was well thought out and well written.

References

- Ethier, W. 2001. The new regionalism in the Americas: A theoretical framework. *North American Journal of Economics and Finance* 12 (2): 159–72.
- Frankel, J. 1997. *Regional trading blocs in the world economic system*. Washington, DC: Institute for International Economics.

