Regional Trading Arrangements: Natural or Supernatural?

Jeffrey A. Frankel; Ernesto Stein; Shang-Jin Wei

The American Economic Review, Vol. 86, No. 2, Papers and Proceedings of the Hundredth and Eighth Annual Meeting of the American Economic Association San Francisco, CA, January 5-7, 1996 (May, 1996), 52-56.

IMPROVING THE DESIGN OF REGIONAL TRADE AGREEMENTS[†]

Regional Trading Arrangements: Natural or Supernatural?

By JEFFREY A. FRANKEL, ERNESTO STEIN, AND SHANG-JIN WEI*

The question of the desirability of regional trading arrangements (RTA's) poses a trade-off. Favorable effects come from the elimination of distortions in the relative price between domestic goods and the products of other members of the customs union. The potential unfavorable effects arise from the introduction of distortions in the relative price between the goods of members and nonmembers. In terms of classic customs-union theory, the trade-off is between trade creation and trade diversion.

The entire exercise assumes that the first-best solution of worldwide free trade is not attainable for political reasons. Thus the choice—between a status quo of nondiscriminatory most favored nation (MFN) tariffs and a move to preferential trading arrangements (PTA's)—is an exercise in the theory of the second best.

This paper summarizes recent research by the authors. We make an evaluation of the trade-off between trade creation and trade diversion operational by parameterizing it along a geographical dimension. The geographical dimension would seem indispensable in an analysis of "regional" trading arrangements, but in the past has been relatively neglected. Our key result is that the desirability of RTA's depends on whether the extent of regionalization exceeds an optimal level that is determined by the magnitude of transportation costs between regions.

Assume a world of C continents, each consisting of N nations. For concreteness, we can consider the case $C = 3$, thereby capturing fears that the world is heading toward a system of three trading blocs: Europe, the Americas, and Asia. We begin with the monopolistic-competition model of trade, characterized by increasing returns to scale in production and love for variety on the part of consumers. Paul Krugman (1991a) has shown in a model without transportation costs that economic welfare is diminished by a move from a system where a large number of individual countries post MFN tariffs, to a system of free trade areas (FTA's). For plausible parameter values, the welfare minimum is reached when there are three large blocs.

A three-bloc world is harmful for two reasons. First, each of the large blocs is tempted to exploit its monopoly power by raising tariffs to a greater extent than they would if acting as smaller blocs or individual countries. The second reason holds even if the blocs are constrained from raising their tariffs against outsiders, as they are under Article XXIV of the GATT (the provision that allows FTA's). The elimination of tariffs within blocs introduces more distortions than it eliminates. Negative effects of trade diversion outweigh positive effects of trade creation.

That verdict is entirely dependent on the assumption of no transport costs, in which case it does not matter whether the FTA members are located on the same continent. We now introduce transport costs. Krugman (1991b) has pointed out that if intercontinental transport costs are prohibitive, then consolidation into continental blocs suddenly becomes the optimal outcome. The intuition is immediate: there is no intercontinental trade to divert. He calls FTA's that are drawn along continental lines *natural*, to distinguish them from intercontinental FTA's, such as the old British imperial preferences, which he calls *unnatural*.

In fact, transport costs are not prohibitively high, of course. Intercontinental trade is large

[†] *Discussants:* Rudiger Dornbusch, Massachusetts Institute of Technology; Alan Winters, World Bank.

* Frankel: Department of Economics, 549 Evans Hall, University of California, Berkeley, CA 94720-3880; Stein: Inter-American Development Bank, Washington, DC; Wei: Kennedy School of Government, Harvard University.

and growing. Presumably the argument that natural FTA's are welfare-improving is meant to apply to the extent that transport costs are *relatively* high. But relative to what? We need to fill in the intermediate case in which intercontinental costs are neither zero (in which case a world of three continental blocs is bad), nor prohibitively high (in which case the three-bloc world is good), but somewhere in between.

Let shipping costs between continents be given by b, as a fraction of the value of the good shipped. There are also costs to shipping within continents, but the key point is that intercontinental trade incurs the added cost. We have several results:[1]

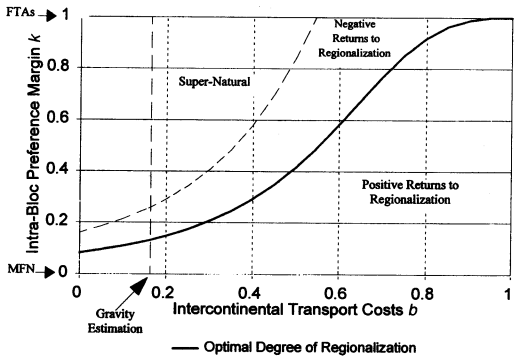

FIGURE 1. RETURNS TO REGIONALIZATION ($C = 3$, $N = 16$)

1. FTA's are likely to be detrimental over a moderate range of parameter values, even if they are drawn along natural continental lines. Specifically, simulation results show this outcome if $b < 0.18$. [This is for a case where $C = 3$, $N = 2$, elasticity of substitution = 4, and the external tariff rate is fixed at 0.3.]
2. Generalizing beyond pure FTA's to PTA's, we find that a small margin of preferences for continental neighbors is always beneficial.
3. The optimal margin of preferences depends on the parameters. For the case of our base set of parameters and $b = 0.15$, the optimal margin of preferences is 0.13. That assumes 16 countries on each of three continents. If individual units have already consolidated into two customs unions on each continent, then the optimal margin of preferences with respect to the other half of the continent is about 0.54.
4. If preferences exceed this optimal level, they enter the zone of negative returns to regionalization. If the PTA's continue to raise their margin of preferences, they will eventually reach what we call the supernatural zone, where welfare is not only suboptimal, but is actually lower than under the MFN status quo. For the case $b = 0.15$, we enter the supernatural zone at preferences of about 0.25 (assuming three continent-wide PTA's, among 16 countries each). If individual units have already consolidated into two large customs unions on each continent, then the supernatural zone begins at about 0.95. If preferences go as high as 100 percent, then we are back to the case of continental FTA's. Figure 1 illustrates the three zones, with b on the horizontal axis and the margin of preferences k on the vertical axis.
5. If each continent comprises many smaller PTA's instead of one large bloc, then the same results hold qualitatively, but the welfare impacts are smaller quantitatively.[2] That is, regional FTA's are still detrimental, but not as detrimental as if they were continent-wide; and regional PTA's with small margins of preference are still beneficial, but not as beneficial as if they were continent-wide.

Models such as these provide a framework for thinking about the welfare effects of RTA's. To get some idea where the world trading system in fact lies in Figure 1, we must turn to the data.

[1] The three-continent model is presented succinctly in Frankel et al. (1997). The theory is spelled out more completely in Stein (1994 Ch. 2). To our knowledge, Paul Wonnacott and Ronald Wonnacott (1981) was the earliest contribution to customs-union theory to refer to the importance of interregional transport costs in excess of intraregional costs.

[2] This is demonstrated in Frankel et al. (1995).

A useful tool for analyzing bilateral trade data is the gravity model. In its strictest form, the gravity equation says that trade between two countries is proportional to the product of their GNP's, and inversely related to the distance between them. It used to be said that, while the equation worked well empirically, it lacked theoretical foundations. By now, however, its foundations are relatively well established. The imperfect-substitutes model of trade described above, for example, will give the basic gravity relationship. We and others have also had empirical success adding to the equation: per capita incomes, land area, bilateral exchange-rate variability, and dummy variables for common borders and common languages. After allowing for these determinants of trade, one adds dummy variables to test for any bloc effects one wishes. Our estimates are based on trade among 63 countries between 1965 and 1992.

Gravity estimates find an effect of log distance on bilateral trade that is highly significant statistically. We substitute a typical coefficient estimate into the theoretical model, combined with the statistic that intercontinental trade covers a distance that is on average 4.0 times as great as intracontinental trade. The resulting estimate of b is about 0.16.

The estimates of bloc effects in the gravity model are much harder to pin down reliably than the effects of distance or the other variables. Nevertheless, we do find intrabloc biases that are generally significant statistically in: the European Community (EC), Mercosur, Andean Pact, Association of SouthEast Asian Nations (ASEAN), and Australia-New Zealand Closer Economic Relationship (ANZCER). When we test larger groupings that are not yet formal RTA's, we also find biases in Europe, the Americas, and Asia. These coefficients, when substituted into the theoretical model, generally correspond to margins of preference in the supernatural zone.

The theoretical model is highly stylized and leaves out many factors. These limitations do not eliminate its usefulness for helping one think about the role that geography plays in the trade-off between trade creation and trade diversion. It would be nice, however, to know whether the results are robust. Sensitivity to parameter values within the model is easily tested.[3] Relaxing fundamental assumptions takes more work. Perhaps the two highest priorities are generalizing the highly stylized model of trade and relaxing the assumption that the interbloc level of tariffs remains fixed.

Alan Deardorff and Robert Stern (1994) and T. N. Srinivasan (1993) question the realism of the Krugman model of trade based solely on goods as imperfect substitutes. In their view, the result that a few large FTA's are worse than many small ones can be attributed to excessive emphasis on the utility of consuming a large variety of goods that may differ only in brand name. They suggest that classical theories of comparative advantage would imply that welfare increases monotonically in the number of countries per bloc.

In reality, trade clearly arises for reasons of both comparative advantage and imperfect substitution. An appealing approach is to model industries as determined by comparative advantage—which is in turn determined by differences in factor endowments as in the traditional Heckscher-Ohlin model—but then to assume that consumers treat different varieties *within* a particular industry as imperfect substitutes. Thus industrialized countries produce automobiles rather than textiles because the former are capital-intensive and the latter labor-intensive, but American autos are imperfect substitutes for Japanese autos.

Antonio Spilimbergo and Stein (1997) have recently extended our results to allow for this mixture of comparative-advantage trade and imperfect-substitutes trade. They first look at the case of zero transportation costs. The Krugman (1991a) result once again emerges, provided consumers' love for variety is not too low: welfare reaches a minimum at three large blocs, versus larger numbers of smaller blocs. If the love for variety is very low, however, welfare rises monotonically as the number of blocs falls, justifying the skeptics. The conclusion offers an optimistic outlook for regionalism. When 60 countries combine into increasingly larger blocs, economic welfare is improved at every step of the way. This suggests that FTA's can be stepping stones

[3] For the sensitivity analysis, see Stein (1994 Ch. 2) and the appendix to Frankel (1996).

toward the ultimate goal of one bloc of 60 countries, also known as worldwide free trade.

Most interesting is what Spilimbergo and Stein (1997) find when they allow for intercontinental transport costs. Their simulations assume a world of four continents, with eight countries on each continent, four of them rich and four poor. Notwithstanding the introduction of differences in factor endowments as a determinant of trade, the results are qualitatively the same as before. Specifically, the three most important results continue to hold. (i) FTA's put the world into the supernatural zone (for a wide range of intercontinental costs, b). However, we are now able to see that the effect is quite different in rich countries than in poor countries. The latter are likely to become better off from a move to four continental blocs, even though the rich are worse off. (ii) PTA's can raise welfare, even for rich countries, provided the margin of preferences is not set too high. (iii) The optimal margin of preferences rises with intercontinental costs. Unless intercontinental costs exceed 0.25, however, the optimal margin of preferences is in the range of 26 percent to 34 percent. Anything above that level enters the zone of negative returns to regionalization, and anything over 65 percent enters the supernatural zone. Even quantitatively, these results are not very different from those we obtained in the model that ignored factor endowments.

Once we endogenize external tariffs, the problem changes more radically. A great many political-economy arguments have been made regarding regionalism, either to the effect that it can undermine general liberalization or to the effect that it can help build political momentum for multilateral liberalization. Which set of forces dominates? Are trade blocs stumbling blocks or building blocks for global free trade?[4]

We can get a rough idea which political forces have tended to dominate over the last 30 years by returning to the gravity model. We add a dummy variable to represent trade of bloc members with nonmembers. The results are mixed. Sometimes the coefficient is negative, indicating trade-diversion. This usually appears to be the case for the European Free Trade Area (EFTA), for example, and for NAFTA and ANZCER. Often, however, the coefficient is positive, suggesting that the bloc lowered its external barriers somewhat at the same time that it liberalized internally. This seems generally to describe ASEAN, the EC, the Andean group, and Mercosur.

The ultimate question for policy purposes is how the international trade rules might be optimally designed to ensure that regionalism is most likely to be welfare-improving. Our results have already cast some doubt on one provision of Article XXIV, which requires that an FTA eliminate internal barriers completely. We found that partial internal liberalization would be better.[5]

Some have proposed modifying Article XXIV to require that RTA's reduce barriers against nonmembers. In one proposal, the external tariff should be cut however much is necessary so that there is no trade diversion (John McMillan, 1993). Unfortunately, the degree of liberalization that members of an FTA must grant to outsiders under this criterion may be larger than a typical bloc is politically prepared to grant.

If, however, we are designing rules for a global trading regime, we must consider a situation in which all regional groupings might opt to form FTA's subject to the restrictions of the regime, not just one. It turns out that the degree of liberalization required for such a rule to be welfare-improving is more modest and attainable than the no-trade-diversion criterion (see Wei and Frankel, 1995). For example, consider the model with interbloc costs of 15 percent and intrabloc preferences of 50 percent (in a world of three 15-nation continental blocs). A simulation suggests that the McMillan restriction is rather severe: to prevent trade diversion, each PTA must liberalize externally by 85 percent as much as it liberalizes internally. If all three blocs are forming their trade policies in a simultaneous equilibrium, however, the criterion necessary to raise

[4] Frankel (1996 Ch. 10) offers a survey of the political-economy arguments.

[5] This ignores some arguments in favor of the provision, particularly that by raising the "hurdle" for approving FTA's, it discourages them altogether (see e.g., Jagdish Bhagwati, 1993).

economic welfare is more moderate: PTA's need only liberalize externally by 25 percent as much as internally. The case for RTA's then looks more promising.

REFERENCES

Bhagwati, Jagdish. "Regionalism and Multilateralism: An Overview," in Jaime de Melo and Arvind Panagariya, eds., *New dimensions in regional integration.* New York: Cambridge University Press, 1993, pp. 22–51.

Deardorff, Alan and Stern, Robert. "Multilateral Trade Negotiations and Preferential Trading Arrangements," in Alan Deardorff and Robert Stern, eds., *Analytical and negotiating issues in the global trading system.* Ann Arbor: University of Michigan Press, 1994, pp. 27–85.

Frankel, Jeffrey. *Regional trading blocs.* Washington, DC: Institute for International Economics, 1996 (forthcoming).

Frankel, Jeffrey; Stein, Ernesto and Wei, Shang-Jin. "Trading Blocs and the Americas: The Natural, the Unnatural, and the Super-Natural." *Journal of Development Economics,* June 1995, *47*(1), pp. 61–95.

_____. "Continental Trading Blocs: Are They Natural, or Super-natural?" in J. Frankel, ed., *The regionalization of the world economy.* Chicago: University of Chicago Press, 1997 (forthcoming).

Krugman, Paul. "Is Bilateralism Bad?" in E. Helpman and A. Razin, eds., *International trade and trade policy.* Cambridge, MA: MIT Press, 1991a, pp. 9–23.

_____. "The Move Toward Free Trade Zones," in *Policy implications of trade and currency zones,* Symposium Sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, WY, August 1991. Kansas City, MO: Federal Reserve Bank of Kansas City, 1991b, pp. 7–42.

McMillan, John. "Does Regional Integration Foster Open Trade? Economic Theory and GATT's Article XXIV," in Kym Anderson and Richard Blackhurst, eds., *Regional integration and the global trading system.* New York: Harvester Wheatsheaf, 1993, pp. 292–310.

Spilimbergo, Antonio and Stein, Ernesto. "The Welfare Implications of Trading Blocs Among Countries with Different Endowments," in J. Frankel, ed., *The regionalization of the world economy.* Chicago: University of Chicago Press, 1997 (forthcoming).

Srinivasan, T. N. "Regionalism vs. Multilateralism: Analytical Notes' Comment," in Jaime de Melo and Arvind Panagariya, eds., *New dimensions in regional integration.* New York: Cambridge University Press, 1993, pp. 84–89.

Stein, Ernesto. "Essays on the Welfare Implications of Trading Blocs with Transport Costs and on Political Cycles of Inflation." Ph.D. dissertation, University of California, Berkeley, 1994.

Wei, Shang-Jin and Frankel, Jeffrey. "Open Regionalism in a World of Continental Trade Blocs." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5272, September 1995.

Wonnacott, Paul and Wonnacott, Ronald. "Is Unilateral Tariff Reduction Preferable to a Customs Union? The Curious Case of the Missing Foreign Tariffs." *American Economic Review,* September 1981, *71*(4), pp. 704–14.