NBER WORKING PAPER SERIES

REGIONAL TRADING ARRANGEMENTS: NATURAL OR SUPER-NATURAL?

Jeffrey A. Frankel
Ernesto Stein
Shang-Jin Wei

Working Paper 5431

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1996

Presented at the American Economic Association meetings, San Francisco, January 6, 1996. Forthcoming, American Economic Review Papers and Proceedings, May 1996. This paper is part of NBER's research programs in International Finance and Macroeconomics, and International Trade and Investment. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper 5431
January 1996

REGIONAL TRADING ARRANGEMENTS: NATURAL OR SUPER-NATURAL?

ABSTRACT

This paper summarizes recent research by the authors on the effects of free trade areas (FTAs). Within our model, which emphasizes inter-continental transport costs, several conclusions arise. (1) FTAs are likely to be detrimental over a moderate range of parameter values, even if drawn along natural regional lines. (2) A small margin of preferences for neighbors is beneficial. (3) Optimal preferences depend on the parameters, particularly on transport costs. (4) If preferences are raised further, they enter the zone of negative returns to regionalization, and eventually the super-natural zone, where welfare is lower than under the MFN status quo. Estimates from the gravity model suggest the world system may already be in the super-natural zone. The core model leaves out many factors. But we have pursued a variety of extensions by now. Perhaps the two most important are generalizing the highly stylized model of trade (to include factor endowments), and relaxing the assumption that the inter-bloc level of tariffs remains fixed. In the latter case, allowing tariffs to be endogenous yields a much more optimistic outlook for the effects of FTAs.

Jeffrey A. Frankel
Department of Economics
549 Evans Hall
University of California
Berkeley, CA 94720-3880
and NBER

Ernesto Stein
Stop W0304
Inter-American
Development Bank
1300 New York Ave., NW
Washington, DC 20577

Shang-Jin Wei
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER

The question of the desirability of regional trading arrangements (RTAs) 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 non-members. In the 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 non-discriminatory MFN (Most Favored Nation) tariffs and a move to PTAs -- is an exercise in the Theory of the Second Best.

This paper summarizes some 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 RTAs 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 East Asia. We begin with the monopolistic-competition model of trade, characterized by increasing returns to scale in production and a love for variety on the part of consumers. 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 FTAs (Free Trade Areas). Welfare is further diminished by a consolidation of FTAs into a few large blocs. 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 as individual countries. They do so in a vain effort to shift the terms of trade in its own advantage -- vain because the other blocs are doing the same. 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 deviations from MFN for the purpose of FTAs). The elimination of tariffs within blocs introduces more distortions, in the decision whether to purchase the goods of non-member countries or of fellow-members, than it eliminates, in the decision whether to purchase the goods of fellow members or of domestic producers. Negative effects of trade-diversion outweigh positive effects of trade-creation.

That judgment is entirely dependent on the assumption of no transport costs. Indeed, so far, it does not matter whether the FTA members are located on the same continents. Let us now introduce transport costs. Krugman (1991b) has pointed out that if inter-continental transport costs are infinite, then consolidation into continental blocs suddenly becomes the optimal outcome. The intuition is immediate: if transport costs are prohibitively high, then there is no inter-continental trade to divert. FTAs give countries the benefit of free trade

within the only universe that matters to them, their continents. He calls FTAs that are drawn along continental lines *natural*, to distinguish them from inter-continental FTAs, which he calls *unnatural*, such as the old British imperial preferences.

Transport costs are not, in fact, prohibitively high, of course. Inter-continental trade is large and growing. Presumably the argument that natural FTAs 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 (or other costs to doing business) 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 inter-continental trade incurs the added cost b. We have several results.[1]

(1) FTAs 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 $b<.18$. [This is for a case where $C=3$, $N=2$, elasticity of substitution = 4, and external tariff-rate is fixed at .3.]

(2) Generalizing beyond pure FTAs to Preferential Trading Arrangements (PTAs), we find that a small margin of preferences for continental neighbors is always beneficial. Intuitively, the gain from lessening the wedge between domestic and fellow-member prices is first-order, while the loss from introducing a new wedge between member and non-member prices is second-order.

(3) The optimal margin of preferences depends on the parameters. For the case of our base set of parameters and $b=.15$, the optimal margin of preferences is .13. That assumes 16 countries on each continent of three. 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 .54.

(4) If preferences exceed this optimal level, they enter the zone of negative returns to regionalization. If the PTAs continue to raise their margin of preferences, they will eventually reach what we call the super-natural zone, where welfare is not only sub-optimal, but is actually lower than under the MFN status quo. For the case $b=.15$, we enter the super-natural zone at preferences of about .25 (assuming 3 continent-wide PTAs, among 16 countries each). If individual units have already consolidated into two large customs unions on each continent, then the super-natural zone begins at about .95. If preferences go as high as 100 percent, then we are back to the case of continental FTAs. As noted, these tend to be detrimental, that is, to lie in the super-natural zone (for a moderately wide range of parameters). 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 PTAs instead of one large bloc, then the same results hold qualitatively, but the welfare impacts are smaller quantitatively.[2] That is, regional FTAs are still detrimental, but not as detrimental as if they were continent-wide; and regional PTAs 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 RTAs. To get some idea where in Figure 1 the world trading system in fact lies, we must turn to the data.

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 GDPs, 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 areas, dummy variables for common borders, common languages and landlockedness, and a measure of bilateral exchange rate variability. 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.[3] 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 intra-continental trade. The resulting estimate of b is about .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 intra-

bloc biases that are generally significant statistically in: the EC (European Community), Mercosur, Andean Pact, ASEAN (Association of SouthEast Asian Nations) and ANZCER (Australia-New Zealand Closer Economic Relationship). When we test larger groupings that are not yet formal RTAs, 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 super-natural zone.

The theoretical model is highly stylized, and leaves many factors out. 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.[4] 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 inter-bloc level of tariffs remains fixed.

Deardorff and Stern (1994) and 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 FTAs 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 the location of production, i.e., brand name. They suggest that classical theories of comparative advantage would imply that welfare increases monotonically in the number of countries per bloc.

Haveman (1992) studies FTAs using a model where trade arises from comparative

advantage rather than from product differentiation (without transport costs). He gets a similar result as Krugman (1991a): expected world welfare is minimized in a world of only two customs unions.

In reality, trade clearly arises for reasons both of comparative advantage and of 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 autos from Japan.

Spilimbergo and Stein (1995) have recently extended the Frankel-Stein-Wei results to allow for this mixture of comparative-advantage trade and imperfect-substitutes trade. They first look at the case where transportation costs are zero (the traditional assumption). 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 12 blocs (of 5 countries each), and then combine into 6 blocs (of 10 each), followed by 3 blocs (of 20 each), economic welfare is improved at every step of the way. This suggests that FTAs can be stepping stones toward the ultimate goal of one bloc of 60 countries, also known as worldwide free trade.

Most interesting is what Spilimbergo and Stein find when they allow for inter-

continental transport costs. Their simulations assume a world of four continents, with 8 countries on each continent, four of them rich and four of them 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. (1) FTAs put the world into the supernatural zone (for a wide range of inter-continental costs, *b*). However, we are now able to see that the effect is quite different in rich countries than poor countries. The latter are likely to be better off from a move to four continental blocs, even though the rich are worse off. (2) Preferential trading arrangement can raise welfare, even for rich countries, provided the margin of preferences is not set too high. (3) The optimal margin of preferences rises with the level of intercontinental costs. Unless intercontinental costs exceed .25 however, the optimal margin of preferences is in the range of 26% to 34%. Anything above that level enters the zone of negative returns to regionalization, and anything over 65 per cent enters the super-natural 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?[5]

We can get a rough idea which political forces have tended to dominate over the last

thirty years, by returning to the gravity model. We add a dummy variable to represent trade of bloc members with countries outside their blocs. The results are mixed. Sometimes the coefficient is negative, indicating trade-diversion, which could result either from the enactment of intra-bloc preferences with an unchanged initial level of external tariffs, or from a tendency to raise external tariffs. This usually appears to be the case for EFTA (European Free Trade Area), for example, and 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 to describe ASEAN in particular, and often the EC, the Andean group, and Mercosur as well.

The ultimate question for policy purposes is how the international trade rules might be optimally designed to insure 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.[6]

Some have proposed modifying Article XXIV to require that RTAs reduce barriers against non-members. In one proposal, the external tariff should be cut however much is necessary so that there is no trade-diversion (McMillan, 1993). The logic is that, under this restriction, non-members will not suffer an adverse shift in their terms of trade, so that the formation of the FTA will not harm them while it helps members. 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 FTAs 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 that would be required of a single grouping acting in isolation.[7] This is all the more true if internal preferences are partial. Consider the model, for example, with inter-bloc costs of 15 percent and intra-bloc 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 so great a liberalization were politically possible, one wonders, why would the negotiation of worldwide free trade not be possible? If all three blocs are forming their trade policies in a simultaneous equilibrium, however, the criterion necessary to raise economic welfare is more moderate: PTAs need only liberalize externally by 25 percent as much as internally. The case for RTAs looks more promising.

References

Bhagwati, Jagdish. "Regionalism and Multilateralism: An Overview," in Jaime de Melo and Arvind Panagariya, New Dimensions in Regional Integration, New York: Cambridge University Press, 1993.

Deardorff, Alan, and Robert Stern. "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, D.C.: Institute for International Economics, 1996, forthcominog.

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

_____. "Continental Trading Blocs: Are They Natural, or SuperNatural?" Forthcoming in J.Frankel, ed., The Regionalization of the World Economy, Chicago: University of Chicago Press, 1996.

Haveman, Jon. On the Consequences of Recent Changes in the Global Trading Environment, Ph.D. thesis, University of Michigan, 1992.

Krugman, Paul. "Is Bilateralism Bad?" in E.Helpman and A.Razin, eds., International Trade and Trade Policy. Cambridge, MA: MIT Press, 1991a.

_____. "The Move Toward Free Trade Zones," in Policy Implications of Trade and Currency Zones, A Symposium Sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 1991b, pp. 7-42.

McMillan, John. "Does Regional Integration Foster Open Trade? Economic Theory and GATT's Article XXIV," in Regional Integration and the Global Trading System, edited by Kym Anderson and Richard Blackhurst, New York: Harvester Wheatsheaf, 1993.

Spilimbergo, Antonio, and Ernesto Stein. "The Welfare Implications of Trading Blocs Among Countries with Different Endowments," forthcoming in J.Frankel, ed., The Regionalization of the World Economy, Chicago: University of Chicago Press, 1996.

Srinivasan, T.N. "'Regionalism vs. Multilateralism: Analytical Notes' Comment," in Jaime de Melo and Arvind Panagariya, New Dimensions in Regional Integration New York: Cambridge University Press, 1993.

Stein, Ernesto. "The Welfare Implications of Asymmetric Trading Blocs," Chapter 2 in Essays on the Welfare Implications of Trading Blocs with Transport Costs and on Political Cycles of Inflation, *Ph.D. thesis*, University of California, Berkeley, 1994.

Wang, Zhen Kun, and L.Alan Winters. "The Trading Potential of Eastern Europe," Centre for Economic Policy Research Discussion Paper No. 610, November 1991, London, UK.

Wei, Shang-Jin, and Jeffrey Frankel. "Open Regionalism in a World of Continental Trade Blocs," NBER Working Paper 5272, September 1995.

Wonnacott, Paul, and Ronald Wonnacott. "Is Unilateral Tariff Reduction Preferable to a Customs Union? The Curious Case of the Missing Foreign Tariffs," American Economic Review 71, no.4, 1981, pp. 704-714.

Endnotes

1. The three-continent model is presented succinctly in Frankel, Stein and Wei (1996). The theory is spelled out more completely in Stein (1994). To our knowledge, Wonnacott and Wonnacott (1981) was the earliest contribution to customs union theory to refer to the importance of inter-regional transport costs in excess of intra-regional costs.

2. Frankel, Stein and Wei (1995).

3. One study that does a particularly careful job of measuring bilateral distance is Wang and Winters (1991).

4. As Stein (1994) and the appendix to Frankel (1996).

5. Chapter 10 of Frankel (1996) offers a survey of the political economy arguments.

6. This ignores some arguments in favor of the provision, particularly that by raising the "hurdle" for approving FTAs, it discourages them altogether. See, e.g., Bhagwati, 1993.

7. Wei and Frankel (1995).

Figure 1: Returns to Regionalization

(C=3; N=16)