Volume Title: The Regionalization of the World Economy

Volume Author/Editor: Jeffrey A. Frankel, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-25995-1

Volume URL: http://www.nber.org/books/fran98-1

Publication Date: January 1998

Chapter Title: Continental Trading Blocs: Are They Natural or Supernatural?

Chapter Author: Jeffrey A. Frankel, Ernesto Stein, Shang-Jin Wei

Chapter URL: http://www.nber.org/chapters/c7821

Chapter pages in book: (p. 91 - 120)

4 Continental Trading Blocs: Are They Natural or Supernatural?

Jeffrey A. Frankel, Ernesto Stein, and Shang-Jin Wei

4.1 Introduction

The world trading system seems to be moving, not just to a system of regional free trade areas (FTAs) but to a system of large continental groupings. In Europe, the European Union (formerly the European Community) removed internal barriers in 1992 and admitted three new members in 1994, bringing the total to fifteen. In December 1994, the leaders of Western Hemisphere countries met in Miami and agreed to form an FTA for the Americas. In East Asia, despite a relative paucity of explicit preferential trading arrangements, one also hears of the emergence of an implicit Japan-centered trade bloc.

This paper seeks to investigate three questions. According to bilateral trade data, is the world indeed breaking up into a small number of continental trade blocs? In theory, is a small number of continental blocs good or bad for world welfare? The answer to the second question is that it depends on parameter values. Thus we also make an attempt to examine the following question:

Jeffrey A. Frankel is professor of economics at the University of California, Berkeley, and a research associate of the National Bureau of Economic Research, where he is also director of the program in International Finance and Macroeconomics. After this project was completed, he took leave to serve on the Council of Economic Advisers. Ernesto Stein is a research economist at the Inter-American Development Bank. Shang-Jin Wei is associate professor of public policy at the Kennedy School of Government, Harvard University, a faculty research fellow of the National Bureau of Economic Research, a research associate of the Center for Pacific Basin Monetary and Economic Studies of the Federal Reserve Bank of San Francisco, and a Davidson Institute Research Fellow at the University of Michigan Business School.

The authors thank Benjamin Chui and Jungshik Kim for research assistance, and Warwick McKibbin, Gary Saxonhouse, and Alan Winters for supplying data. They would also like to thank for support the Center for International and Development Economics Research, funded at University of California, Berkeley, by the Ford Foundation; the Japan–United States Friendship Commission, a U.S. government agency; and the Institute for International Economics. This paper was written while Frankel was a professor at Berkeley and research associate of the NBER; it does not represent the views of the U.S. government.

for actual parameter values, is the current pattern of regionalization welfare-promoting or welfare-reducing?

Paul Krugman has helped to focus the recent debate on the welfare implications of regional trade blocs. He has, however, supplied equally clever arguments on both sides. In his first contribution (Krugman 1991a), he argued that a small number of trade blocs tends to be welfare-reducing, relative to the case when every country charges a common tariff against all other countries. In the latter case, the distortions introduced relative to global free trade are relatively small. For plausible parameter values, three regional blocs turned out to be the worst outcome for the world. This is a worrisome conclusion since this is precisely the direction toward which many observers think the world is moving.

Krugman's second contribution (1991b) included a simple argument for a diametrically opposite conclusion. If transportation costs are very high between continents, then regional trade blocs, if they are formed along continental lines, must be welfare-improving. The intuition is simple. With prohibitively high intercontinental transport costs, trade takes place mainly among countries on the same continents even in the absence of trade blocs. Therefore, the case for regional blocs that eliminate tariffs among countries in the same continents ("natural blocs" in Krugman's terminology) is the same as the standard case for global free trade. Such natural blocs are contrasted with "unnatural blocs," free trade agreements between individual countries in different continents, which are less likely to be welfare-improving.[1]

Each of these two arguments is valid within its own assumptions. The first argument assumes zero transport cost. The second argument relies on a prohibitively high intercontinental transport cost. The world is somewhere between these two extremes. In order to investigate the welfare implication of the current pattern of trade regionalization, we develop a more general model that can handle intermediate cases. We identify conditions under which trade blocs are welfare-reducing, even when formed along continental lines.

With the more general model, we also investigate whether a particular aspect of article XXIV of General Agreement on Tariffs and Trade (GATT) is sensible. Under this rule, countries that want to form a regional trade bloc must eliminate all barriers among themselves. (There are other restrictions as well: the members must not raise barriers against nonmembers.) In our model, however, we find that partial liberalization within a regional bloc is generally better than 100 percent liberalization, in contrast to the article XXIV provision.

The welfare implication of continental blocs depends on the values of some crucial parameters. In the final part of the paper, we make an attempt—perhaps best described as illustrative—to extract estimates of the real-world counterparts of these key parameters, particularly the magnitude of transport costs.

1. It should be noted that the idea of proximity as a desideratum for successful FTAs, on the grounds that it would minimize the amount of trade diversion, was not entirely new with Krugman. (See Balassa 1987, 44; and Wonnacott and Lutz 1989). The leading opponent to the idea is Bhagwati (1993).

The resulting estimates suggest that the current pattern of trade blocs features a degree of regional preferences that is likely to exceed what can be justified on "natural" geographic grounds. We propose the term "supernatural" for blocs that reduce economic welfare in this way.

Most of our conclusions regarding economic welfare presume worldwide symmetry. We look at the consequences of a worldwide regime that allows continental FTAs to form everywhere, not at the consequences of forming a single FTA with trade policies in the rest of the world taken as given. Many other possible considerations, in addition to asymmetry, are omitted from the analysis as well. For example, we focus only on the static economic effects.

4.2 Are Continental Trade Blocs Forming?

It may appear obvious that world trade is increasingly regionalized. A popular statistic to look at is intraregional trade as a percentage of the region's total trade. Table 4.1 shows that intraregional trade shares increased during the 1980s in each of three major parts of the world: from 54 percent to 60 percent in Europe, from 23 percent to 29 percent in East Asia, and (less strikingly) from 27 percent to 29 percent in the Western Hemisphere.

Such statistics are ill-suited, however, for the purpose of examining the degree of regional bias in trade policy. There are many factors that may contribute to the spatial distribution of trade. First, neighboring countries tend to trade more with each other. Second, large countries tend to trade more, and faster-growing economies tend to increase their mutual trade volume at a faster pace. Hence, the observed increasing share of intra-Asia trade, for example, could be entirely due to the region's above-average growth rate, with no explicit or implicit trade bloc in formation. A more useful way to detect the effectiveness of an existing trade bloc is to see whether trade among a given group of countries is unusually high after controlling for factors that naturally contribute to the volume of trade.

4.2.1 The Gravity Model of Bilateral Trade

The natural empirical framework for studying bilateral trade is the gravity model, which offers a systematic way to measure what patterns of bilateral trade are normal around the world. The simplest specification says that trade between two countries is proportionate to the product of their GDPs and inversely related to the distance between them, by analogy to the formula for

Table 4.1 **Intraregional Trade Shares (%)**

	1965	1970	1975	1980	1985	1987	1990
East Asia	.199	.198	.212	.229	.256	.263	.293
Western Hemisphere	.315	.311	.309	.272	.310	.279	.286
Western Europe	.502	.532	.524	.538	.548	.601	.602

gravitational attraction between two bodies. A dummy variable can be added to represent when both countries in a given pair belong to the same regional grouping. We can then see how much of the trade within each region can be attributed to a special regional effect.

The gravity model has a fairly long history and fits the data remarkably well empirically, though its theoretical foundations have hitherto been considered limited.[2] Earlier work by Anderson (1979) and other papers surveyed by Deardorff (1984, 503–6) have provided a partial foundation for the approach. Specifically, the idea that bilateral trade depends on the product of GDPs can be justified by the modern theory of trade under imperfect competition, as shown in recent work by Helpman (1987) and Helpman and Krugman (1985, section 1.5).[3]

Some recent papers that apply the gravity model to trade-bloc issues (e.g., Frankel 1993 and Frankel and Wei 1994, which focus on East Asia, among others) are incapable of distinguishing between regional biases reflecting discriminatory trade policies, and those that might derive from historical, political, cultural, and linguistic ties. In this paper, we include terms representing pairs of countries that speak a common language or have other historical ties.

The dependent variable is trade (exports plus imports), in log form, between pairs of countries in a given year. Our data source is the United Nations trade matrix, which covers sixty-three countries in our data set, so that there are 1,953 data points ($= 63 \times 62/2$) for a given year. In the regressions reported here, we use data for 1970, 1980, and 1990.

A large part of the apparent bias toward intraregional trade is certainly due to simple geographical proximity. Krugman (1991b) and Summers (1991) opine that *most* of it may be due to proximity. (At the other extreme, Bhagwati [1993] and Panagariya [1995, 9–10] assert that very little of the apparent bias is due to proximity, emphasizing rather regional trade arrangements that are already in place.) Surprisingly, empirical studies often neglect to measure this factor. Our measure is the log of distance between the two major cities (usually the capitals) of the respective countries. We also add a dummy *ADJACENT* variable to indicate when two countries share a common land border.

The other of the most important factors in explaining bilateral trade flows is one that we have already identified as the essence of the gravity model: the

2. The results of one extensive early project along these lines were reported in Tinbergen (1962, appendix 6, pp. 262–93) and Linnemann (1966). Recent empirical studies include Bergstrand (1989), Wang and Winters (1991), and Hamilton and Winters (1992).

3. It has long been considered that the classical Heckscher-Ohlin theory of comparative advantage does not have this property. Deardorff (chap. 1 in this volume), however, now argues that the gravity relationship can also be derived from the Heckscher-Ohlin model as well as the imperfect substitutes model. The proportionality between trade and the product of GDPs is also a property of our theoretical model introduced in section 4.3. Frankel, Stein, and Wei (1995) includes a more detailed exposition of how our model offers improved foundations for the gravity model, in that the effects of both size and distance are derived.

economic size of the two countries. In addition to the well-established effect of size on trade, there is reason to believe that GDP per capita has a positive effect on trade, for a given size: as countries become more developed, they tend to specialize more and to trade more. Finally, *LANG* is the dummy variable reflecting a common language.

The equation to be estimated is thus

$$\begin{aligned}\log(Trade_{ij}) = {} & \alpha + \beta_1 \log(GNP_i GNP_j) + \beta_2 \log(GNP/pop_i GNP/pop_j) \\ & + \beta_3 \log(DIST_{ij}) + \beta_4 (ADJACENT_{ij}) + \beta_5 (LANG_{ij}) \\ & + \gamma_1 (EC_{ij}) + \gamma_2 (WH_{ij}) + \gamma_3 (EA_{ij}) + u_{ij}. \end{aligned} \tag{1}$$

The last five explanatory factors are dummy variables. *EC*, *WH*, and *EA* are three of the dummy variables we use when testing the effects of membership in a common regional grouping, standing for European Community, Western Hemisphere, and East Asia, respectively. We will estimate average intraregional biases over 1970–90, and possible trend increases in the biases.

Table 4.2 reports pooled time-series-cross-section estimates of equation (1) that extend from 1970 to 1990. All gravity variables are highly significant statistically (>99 percent level). The coefficient on the log of distance is about −0.5, when the *ADJACENT* variable (which is also highly significant statistically) is included at the same time. This means that when the distance between two nonadjacent countries is higher by 1 percent, the trade between them falls by about 0.5 percent. We checked for possible nonlinearity in the log-distance term, as it could conceivably be the cause of any apparent bias toward intraregional trade that is left after controlling linearly for distance, but this did not seem to be an issue.[4] We should note that physical shipping costs may not be the most important component of costs associated with distance. What we call transport costs would better be understood as transactions costs, encompassing not just each physical transportation of goods, but also costs of communications and the idea that each country tends to have a better understanding of its close neighbors and their institutions.

The estimated coefficient on GNP per capita is 0.23, indicating that richer countries do trade more.[5] The estimated coefficient for the log of the product of the two countries' GNPs is about 0.7, indicating that, although trade increases with size, it increases less than proportionately (holding GNP per cap-

4. The log of distance appears to be sufficient; the level and square of distance add little. We have also tried distance measures that take into account the greater distances involved in sea voyages around obstacles like the Cape of Good Hope and Cape Horn (the data generously supplied by Wang and Winters), with little effect on the results.

5. We have also tried an estimate for 1991 using GNPs adjusted by purchasing power parity, in place of using exchange rates to translate GNPs. The coefficient on GNP/capita is 0.35, with little qualitative change in the estimates otherwise. (The sum of the coefficients on GNP and GNP/capita is about 1. Thus there is another, approximate way to describe the results: openness [as measured by trade/GNP] falls by 0.23 percent for every 1 percent increase in size [as measured by population].)

Table 4.2 **Gravity Estimation of Continental Trade Blocs (total trade 1970–90)**

Dependent Variable: log ($Trade_{ij}$)		
Intercept	−9.70*	−9.78*
	(0.27)	(0.27)
1980 dummy	−1.01*	−1.06*
	(0.05)	(0.05)
1990 dummy	−1.29*	−1.37*
	(0.06)	(0.06)
GNP	0.72*	0.73*
	(0.01)	(0.01)
Per capita GNP	0.23*	0.23*
	(0.01)	(0.01)
Distance	−0.51*	−0.51*
	(0.02)	(0.02)
Adjacency	0.72*	0.72*
	(0.10)	(0.09)
Common language	0.47*	0.47*
	(0.05)	(0.05)
EC bloc	0.31*	0.24*
	(0.06)	(0.09)
East Asia bloc	2.12*	2.26
	(0.09)	(0.18)
Western Hemisphere bloc	0.31*	−0.32*
	(0.08)	(0.10)
EC*Trend		0.006
		(0.006)
East Asia*Trend		−0.013
		(0.012)
Western Hemisphere*Trend		0.063*
		(0.009)
Observations	4555	4555
Standard error of regression	1.15	1.14
Adjusted R^2	0.76	0.76

Notes: Standard errors are in parentheses. All variables except the intercepts and dummy variables are in logs. Trend = Year − 1970.

*Significant at the 1% level.

ita constant). This presumably reflects the widely known pattern that small economies tend to be more open to international trade than larger, more diversified economies.

The linguistic dummy, *LANG,* represents pairs of countries that share a common language or had colonial links earlier in the century. We allowed for English, Spanish, Chinese, Arabic, French, German, Japanese, Dutch, and Portuguese. Two countries sharing linguistic/colonial links tend to trade roughly 60 percent more than they would otherwise (exp (0.47) = 1.60). We tested whether some of the major languages were more important than the others, and found little in the way of significant differences.

4.2.2 Estimation of Trade-Bloc Effects

If there were nothing to the notion of trading blocs, then the five basic variables in table 4.2 might soak up most of the explanatory power. There would be little left to attribute to a dummy variable representing whether two trading partners are both located in the same region. In this case the level and trend in intraregional trade would be due solely to the proximity of the countries, and to their rates of overall economic growth.

We found, however, that all three regional dummies are statistically significant. The coefficient for the EC dummy is 0.31. This means that over the period 1970–90, two EC countries traded 36 percent (exp (0.31) = 1.36) more than two otherwise similar countries. As is indicated by the EC–time-trend interactive term in the second column, the within-EC bias increases at the rate of about 0.6 percent a year, although this trend is not statistically significant. Averaged over time, the intra–Western Hemisphere trade bias is of the same order of magnitude as the EC. But it started out low in 1970 (in fact, negative), and increased over time at the rate of 6 percent a year. This relatively rapid trend increase is statistically significant. By 1990, two Western Hemisphere countries traded 150 percent more than two otherwise-similar countries (exp $[-0.323 + (20 \times 0.063)] = 2.55$). The East Asian grouping exhibits the highest intraregional bias. Intra–East Asian trade, however, once we have controlled for the gravity variables, shows no significant increase in bias over the period. If anything, this bias diminished over time, rather than rising, as often assumed.[6]

In table 4.3, we employ a dummy for all of Western Europe, instead of considering the EC bloc alone. The motivation is to treat the European continent on a par with the other two.[7] The intraregional bias for Western Europe is slightly smaller than that for the EC. The coefficients for other variables in the specification are essentially unchanged.

We have conducted more robustness checks in addition to those that have been mentioned earlier. Chapter 7 of this volume tests the effect of bilateral exchange rate variability on trade, and the effect of the degree of openness (more trade with all countries, as distinct from more trade just with other members of the grouping). Other extensions that we tried include allowing a role for differences in factor endowments as additional regressors, a correction for

6. In other words, the rapid growth of East Asian economies is in itself sufficient to explain the increase in the intraregional trade share evident in table 4.1. This finding, that intraregional trade bias in Asia did not rise in the 1980s as often assumed, confirms Frankel (1993), Petri (1993), Saxonhouse (1993), and Anderson and Norheim (1993).

7. Some readers are less interested in the effects of three continental blocs than of explicit FTAs, like the EC and NAFTA. The effects of such subregional blocs are estimated in Frankel, Stein, and Wei (1995). (The results suggest that regionalization in the Western Hemisphere during 1980–90 is concentrated in the Andean Pact and Mercosur groupings. The NAFTA grouping is not statistically significant in this sample period.)

Table 4.3 **Gravity Estimation, Western Europe instead of European Community (total trade 1970–90)**

Dependent Variable: log ($Trade_{ij}$)		
Intercept	−9.77*	−9.89*
	(0.28)	(0.28)
1980 dummy	−1.00*	−1.04*
	(0.05)	(0.05)
1990 dummy	−1.28*	−1.33*
	(0.06)	(0.06)
GNP	0.73*	0.73*
	(0.01)	(0.01)
Per capita GNP	0.22*	0.23*
	(0.01)	(0.01)
Distance	−0.50*	−0.50*
	(0.02)	(0.02)
Adjacency	0.72*	0.72*
	(0.10)	(0.09)
Common language	0.47*	0.48*
	(0.05)	(0.05)
Western Europe bloc	0.21*	0.35*
	(0.06)	(0.08)
East Asia bloc	2.12*	2.29*
	(0.09)	(0.19)
Western Hemisphere bloc	0.32*	−0.29*
	(0.09)	(0.10)
Western Europe*Trend		−0.015*
		(0.004)
East Asia*Trend		−0.015
		(0.012)
Western Hemisphere*Trend		0.061*
		(0.009)
Observations	4555	4555
Standard error of regression	1.15	1.14
Adjusted R^2	0.76	0.76

Notes: Standard errors are in parentheses. All variables except intercepts and dummy variables are in logs. Trend = Year − 1970.

*Significant at the 1% level.

heteroscedasticity related to the size of the countries (weighted least squares), and the inclusion of country pairs recorded with zero trade. The answers to the question of interest here, the estimates of the intraregional biases, are fairly robust to these variations.

The gravity model results thus show that, on average over 1970–90, the three regions did exhibit inward trade bias. The next question is whether these biases constitute an undesirable threat to the world trading system.

4.3 The Theory of Bilateral Trade with Imperfect Substitutes and Transport Costs

We now attempt to settle the Krugman versus Krugman controversy regarding the desirability of trading blocs. We construct in this section a more general model that can handle the intermediate realistic case where transport costs between continents are less than infinite, while greater than zero. Section 4.4 derives the implications of this model for trading blocs. Section 4.5 aims to match up the theory of sections 4.3 and 4.4 with section 4.2's empirical estimates of the effects of transport costs and regional trading arrangements on the volume of bilateral trade, in order to evaluate the welfare implications of regionalization of the world economy.

4.3.1 The Differentiated Products Model

We employ a monopolistic competition model similar to Krugman (1980), who in turn followed Dixit-Stiglitz. A representative consumer has the utility function

$$U = \sum_i c_i^{\theta}; \qquad 0 < \theta < 1, \tag{2}$$

where c_i is the consumption of the ith variety. This utility function results in preference for variety by consumers. The higher the parameter θ, the lower the love for variety. In the limit of perfect substitutability, $\theta = 1$. In the limit of complete love for variety, consumers care only about the number of varieties consumed, and not at all about the quantity: $\theta = 0$.

Labor is the only factor of production. The total national supply of labor is L. Increasing returns are introduced by assuming a fixed cost and a constant marginal cost in the production of each of the varieties. We assume that individual firms maximize profits, and free entry assures a zero-profit equilibrium. Under these simple assumptions, the scale of output of each variety does not depend on the size of the economy. Rather, it is the number of varieties n that increases when the size of the economy (L) increases:

$$n = \frac{L(1 - \theta)}{\alpha}, \tag{3}$$

where α is the parameter representing the fixed costs of production of a new variety. Notice that in the case of very low substitutability (as θ tends to 0), the number of varieties produced approaches L/α, and an infinitely small amount of each of them is produced, since consumers care only about the number of varieties available.[8]

To see the gains from international trade, which arise here from the opportu-

8. Details of derivations are given in Stein and Frankel (1994).

nity to consume a greater variety of goods, we assume that countries have similar tastes and technologies. If we have two countries of equal size, allowing for unfettered trade will double the number of available varieties in each country and thus raise utility. The gains from trade have nothing to do with differences in factor endowments or technology.

Deardorff and Stern (1994) question the realism of this setup. In their view, the Krugman result that a few large blocs 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). In a model based on comparative advantage, they show how FTAs formed by a few dissimilar countries can yield welfare levels that are very close to those under free trade. They suggest that the emphasis on love for variety in the monopolistic competition model overstates the trade-diversion effect, leading to an overly pessimistic view of FTAs. An unfortunate limitation of their model is the assumption of prohibitive tariffs between countries that do not belong to the same bloc. As Haveman (1992) shows, the addition of optimal tariffs to a comparative-advantage-type model results in welfare effects similar to those of Krugman's product-variety model.

We believe that the product-variety approach is relevant, since it helps explain an important and increasing portion of world trade, in particular in manufactures. This is reflected in the increase in intraindustry trade as a proportion of total trade. Ideally, one would like to come up with a model where gains from trade are explained both by comparative advantage and by increase in variety. (This avenue of research is explored by Spilimbergo and Stein in chap. 5 of this volume.) In the meantime, we think both the comparative advantage and the product variety models help to illustrate some of the effects of the formation of trading blocs, and are useful in their own right.

4.3.2 Introduction of Transport Costs and Tariffs

We will think of the world as being divided into a number of continents (C), each of them equidistant from one another. Each of these continents is composed of a number of countries (N). The transportation system we assume within each continent is a hub-and-spoke network. In each continent there is a hub, through which all trade involving that continent must pass. Each hub has N spokes, all assumed of equal length, connecting it to the N countries in the continent.

Transport costs will be assumed to be of Samuelson's iceberg type, which means that only a fraction of the shipped good arrives; the rest is lost along the way. The cost of transport through two spokes will be represented as a, while that of transport from hub to hub (across the ocean) is given by b, where $0 \leq a, b \leq 1$. Trade involving two countries on the same continent will have to be transported from the exporting country to the hub, and from the hub to the importing country. This involves two spokes, and so the fraction of a good

shipped that arrives at the market is $1 - a$. Similarly, the fraction of a good that arrives in the case of trade between countries in different continents, which involves two spokes and a hub-to-hub section, is $(1 - a)(1 - b)$.

When a consumer buys a foreign good, the government levies an ad valorem tariff t. We assume that the tariff is levied on the c.i.f. price.[9] The level of tariffs is exogenous and assumed to be uniform across countries, representing the most-favored-nation (MFN) principle, until we are ready to examine FTAs. Tariff receipts are returned to the consumer as a lump-sum transfer.

For simplicity, we will assume that the countries are equal in size. The symmetry of the model now assures that producers' prices are the same in every country. The same is true of the number of varieties and the quantity of each variety produced in every country.

Prices of home and foreign goods faced by home consumers are different due to transport costs and tariffs. If the producer prices in every country are p, then the price the domestic consumer will have to pay for every unit of foreign good consumed would be

$$(4) \qquad p_c = \frac{p(1 + t_c)}{1 - a}, \qquad p_{nc} = \frac{p(1 + t_{nc})}{(1 - a)(1 - b)},$$

where the subscript c refers to goods imported from within the continent, and nc otherwise (across continents). The corresponding tariffs imposed on the two types of foreign goods are t_c and t_{nc}.

It is useful to fix some more notation here. The situation in which there are no continental trade blocs and every country charges the same positive tariff on goods from all other countries (MFN) can be represented by $t_c = t_{nc} = t > 0$. Global free trade is described by $t_c = t_{nc} = 0$. Finally, the case of every continent forming an FTA is characterized by $t_c = 0$ but $t_{nc} = t > 0$.

The next step is to derive from the utility function the consumption of each foreign variety (both from neighbor countries and from countries in other continents), relative to the consumption of each home variety. We begin with the MFN case $t_c = t_{nc} = t > 0$.

For ease of exposition, we will index goods in such a way that the home country produces varieties $1, \ldots, n$; neighbors produce varieties $n + 1, \ldots, n + n^c$; and countries across the ocean produce varieties $n + n^c + 1, \ldots, n + n^c + n^{nc}$. The home consumer maximizes

$$(5) \qquad U = \sum_{i=1}^{n+n^c+n^{nc}} c_i^{\theta}$$

subject to the budget constraint

9. I.e., tariffs are a proportion of the value of the good including transport costs, in terms of the iceberg model. This assumption is simpler, and probably more realistic as well, than letting tariffs be levied on the f.o.b. price.

$$(6) \qquad \sum_{i=1}^{n} c_i p + \sum_{n+1}^{n+n^c} c_i p_{c,t} + \sum_{n+n^c+1}^{n+n^c+n^{nc}} c_i p_{nc,t} \leq w + T,$$

where w is the wage and T is the lump-sum transfer received by each consumer, which they regard as being fixed.

From the maximization problem of the consumers it is possible to derive the elasticity of demand for exports faced by the producers, which turns out to be $\varepsilon_x = 1/(1 - \theta)$, the same as the elasticity of domestic demand. The equality of these elasticities guarantees that the price resulting from the firm's profit maximization is the same as in the case of the closed economy. So are the quantity produced of each variety and the number of varieties n produced in each country. Transport costs and tariffs thus introduce no changes in these variables. But the key point is the effect on consumption patterns.

The first-order conditions for the consumer's problem yield the relative consumption of each variety:

$$(7) \qquad \frac{c_i^c}{c_i^h} = \left(\frac{p}{p_{c,t}}\right)^{1/(1-\theta)}$$

and

$$(8) \qquad \frac{c_i^{nc}}{c_i^h} = \left(\frac{p}{p_{nc,t}}\right)^{1/(1-\theta)},$$

where c_i^c and c_i^{nc} are the domestic consumer's consumption of foreign varieties, from countries within the continent and across the ocean, and c_i^h is the domestic consumer's consumption of the home varieties.

4.3.3 Welfare Implications of Trade Agreements

To evaluate world welfare, we derive the utility of a representative individual in any country. To determine the utility of the consumer, we need to know how much he or she consumes of each good, and introduce these values into the utility function. Equations (7) and (8) above give us the consumption of foreign relative to home varieties, so we need only to determine the consumption of each home variety, c_i^h. We do this by expressing the budget constraint in terms of c_i^h, and taking into account the redistribution of the tariff revenue to consumers.

If we normalize p to be 1, we can obtain, after some algebra,

$$(9) \qquad c_i^h = \frac{w/n}{1 + (N-1)\left(\frac{1}{p_c}\right)^{\varepsilon}\left(p_c - \frac{t}{1-a}\right) + (C-1)N\left(\frac{1}{p_{nc}}\right)^{\varepsilon}\left(p_{nc} - \frac{t}{(1-a)(1-b)}\right)}$$

where $\varepsilon = 1/(1 - \theta)$, and $w/n = \theta\alpha/L\beta(1 - \theta)$.[10] Once we have the consumption of domestic varieties, the consumption of foreign varieties can be obtained from the relative consumption equations (7) and (8). Replacing these in the utility function, we obtain the value of the utility of the representative individual:

$$(10) \quad U = c_i^{h\theta}\left[1 + (N-1)\left(\frac{1}{p_{c,t}}\right)^{\theta/(1-\theta)} + (C-1)N\left(\frac{1}{p_{nc,t}}\right)^{\theta/(1-\theta)}\right].$$

Given values for the parameters a, b, t, θ, N, and C, we can first obtain the value of c_i^h by plugging the price equations (4) into (9), substitute into (7), (8), and (10), and thus find the value of the utility of the representative individual, which is our measure of world welfare.

Equation (10) is the expression for utility in the absence of free trade agreements. It is straightforward to calculate utility under other arrangements in the same manner. When trading blocs are formed, we just introduce the new set of relative prices faced by the home consumers into their maximization problem, and we can obtain new values for utility in a similar way.

4.4 Welfare Effects of Continental Trade Blocs

We have presented a model that allows us to analyze the desirability of different trade arrangements from a world welfare perspective, as well as the changes associated with these different arrangements in terms of the bilateral volume of trade between countries. In the absence of transport costs, our model is reduced to Krugman's model (1991a). That is, we obtain a U-shaped welfare curve as a function of the number of trade blocs. For plausible values of parameters, a system of three or so trade blocs is the worst outcome. We now use the model *with* transport costs, to examine the desirability of trade blocs formed along continental lines.

In our first exercise, we explore the desirability of forming natural and unnatural trading blocs as a function of transport costs. In particular, in this application we look at FTAs, where the intrabloc tariffs are completely eliminated.

In our second exercise, we analyze the implications of what could be considered an intermediate degree of regionalization, a partial movement toward the creation of (natural) FTAs, and compare it to the outcome associated with a full movement in that direction. We allow for the formation of preferential trade agreements (PTAs) that differ from the FTAs in that the tariff level is reduced among partners, but not necessarily eliminated. Even though it is technically prohibited by article XXIV, many existing regional arrangements are in fact partial. We will show that a partial movement toward regional integration, as in the case of PTAs with preference below 100 percent, is usually superior

10. This last equality can be derived from equation (3) together with the profit maximization condition.

to a complete one, associated with FTAs.[11] At the same time, this application illustrates the need for a more complete characterization of trading blocs, one that goes beyond the natural/unnatural distinction.

Throughout, we consider only exercises involving symmetric formation of equal-sized blocs around the world. Deardorff and Stern (1994) and Srinivasan (1993) have taken exception to the symmetric logic of Krugman's bloc question. We, like Krugman, do not address here the asymmetric partial equilibrium exercise of examining the effects of forming a single bloc in one part of the world, particularly the effects on countries unfortunate enough to be left out of any bloc.[12] The motivation, as we see it, is to address the desirability of the international regime with respect to blocs worldwide, that is, article XXIV. It is of course true, however, that variation in GNPs across countries, if nothing else, renders the real world an inherently asymmetric place.

4.4.1 Transport Costs and Free Trade Agreements' Welfare Effects

In this application, we study how the effect of the formation of continental free trade agreements on welfare depends on intercontinental transport costs. Thus we are able to fill in the realistic intermediate case between Krugman's polar cases of zero and infinite intercontinental transport costs. We start with the simple case where the world consists of three continents comprising two countries each. Transport costs *within* continents, a, are for simplicity assumed to be zero in the simulations reported here.

Figure 4.1 shows the percentage change in welfare associated with the formation of trading blocs, both of the natural and unnatural type, for $\theta = 0.75$ and $t = 0.3$.[13] We can see that there is a critical level of intercontinental transport costs b, which governs the welfare effects. For the case of natural trading blocs, where each country forms a bloc with its neighbor, the critical value of b is 0.186: for values of b higher than this, the formation of continental FTAs will result in improvements in welfare. (Remember, in the limiting Krugman's case where $b = 1$, natural blocs are necessarily welfare-improving.) For lower values of b, continental FTAs would reduce welfare. (Remember the limit case where $b = 0$.) As noted in the introduction, we label such welfare-reducing arrangements "supernatural blocs," to indicate that intercontinental transport costs are not high enough to justify the formation of FTAs even along the lines of geographical proximity.[14]

11. Admittedly, the usual pattern in practice is not to reduce tariffs partway on all goods. Rather, agreements tend to exempt certain "sensitive" sectors from liberalization, which raises extra problems of distortions and rent-seeking behavior that are not considered here, but that constitute valid arguments in favor of article XXIV. (Less worrisome is the tendency for agreements to phase in tariff cuts gradually over time.)

12. This question is considered, however, by Stein (1994).

13. Some sensitivity analysis with respect to these and other parameters is reported in Frankel, Stein, and Wei (1995).

14. When tariffs are assumed to be levied on f.o.b. instead of c.i.f. values, the critical value of b is approximately 0.15.

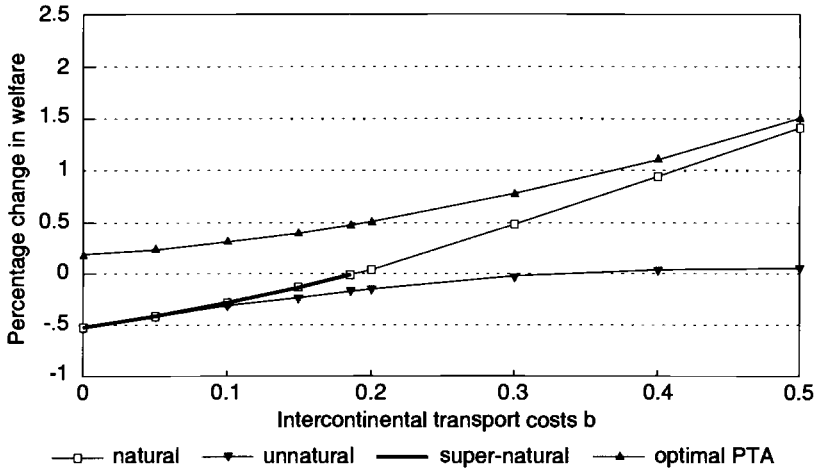

Fig. 4.1 Welfare effects of blocs: natural, unnatural, and supernatural
Notes: $\theta = 0.75$; $t = 0.3$; $a = 0$; $N = 2$; $C = 3$.

Unnatural trading blocs, where each country forms a bloc with one other country outside the continent, result in distinctly lower welfare for small values of b. Krugman's idea (1991a) that natural trade arrangements have a better chance of improving welfare than arrangements between unnatural partners is thus confirmed.

4.4.2 Allowing for Preferential Trade Agreements

In this application, we take another look at trading blocs of the "natural" kind (among neighbors), but we will allow for the formation of PTAs, that is, partial liberalization. To do this, we need to modify our model slightly. The tariff level between partners, instead of zero, will now be $(1 - k)t$, where $0 \leq k \leq 1$, and k is the degree of preference for intrabloc trade or the degree of intrabloc liberalization. The price of partner varieties faced by domestic consumers becomes

$$p_c = \frac{p\,[1 + (1 - k)t]}{1 - a}.$$

Until now we were considering only the two special cases of $k = 0$ (MFN, or the absence of trading blocs) and $k = 1$ (FTAs). Now we consider the whole range of possible levels of intrabloc preference. We will begin as in the previous application, with a world that consists of three continents, each formed by two countries.

What level of intrabloc preference maximizes welfare? Figure 4.2 shows the welfare level as a function of k, for $t = 0.3$, $\theta = 0.75$, $a = 0$, and several values

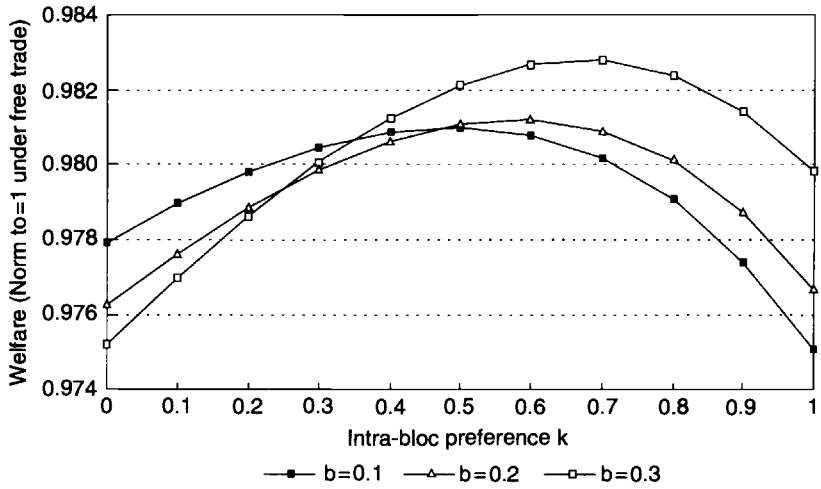

Fig. 4.2 Effects of preferential trade arrangements
Notes: $\theta = 0.75$; $t = 0.3$; $a = 0$; $C = 3$; $N = 2$.

of b.[15] This figure is closely related to figure 4.1. There we were comparing the welfare levels associated with the two extremes of $k = 0$ and $k = 1$ for every possible level of intercontinental transport cost b. For $b < 0.186$, figure 4.1 indicates that the formation of FTAs along natural regional lines is welfare-reducing (supernatural). In figure 4.2, this translates into a higher welfare level for the MFN or no-preference extreme ($k = 0$) relative to the opposite endpoint of full continental FTAs ($k = 1$) for $b = 0.1$.

The important thing to notice in figure 4.2 is that, for every level of intercontinental transport costs, the degree of intrabloc preference associated with maximum welfare is between 0 and 1. In general, PTAs with less than 100 percent preferences are superior to FTAs. This confirms a conjecture by Meade (1955). The key to this result is the diminishing marginal utility for the consumption of each variety. The intuition is easier to understand under zero transport costs, where trade policy does not affect total consumption. Under MFN, households will consume the same amount of every foreign variety, but a larger amount of the home varieties. Imagine that the formation of FTAs entails successive small reductions of intrabloc tariffs. With the first reduction, trade diversion has a small welfare effect, since there is a shift between varieties that were consumed in similar quantities. But trade creation effects are large, since home varieties (with smaller marginal utility) are replaced by member varieties (with larger marginal utility). Thus, a small reduction in intrabloc tariffs starting from MFN will improve welfare. The opposite is true for the last reduction of intrabloc tariffs. Under FTAs, the consumption of member and home varie-

15. For each set of parameter values in the figure (transport cost and θ), welfare is normalized to be 1 under free trade.

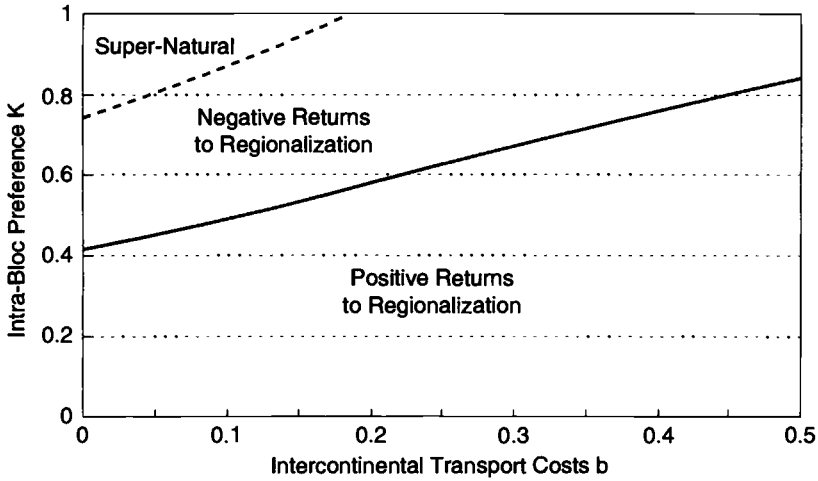

Fig. 4.3 Returns to regionalization
Notes: $\theta = 0.75$; $t = 0.3$; $a = 0$; $N = 2$; $C = 3$.

ties is the same (in the absence of transport costs), while the consumption of other foreign varieties is lower. In this case, the welfare effects of trade creation are negligible, while trade diversion has a larger effect, since varieties with larger marginal utility (those from other foreign countries) are replaced by varieties from member countries, which have smaller marginal utility.

Figure 4.3 provides another way of looking at this issue. For the set of parameters chosen, it represents all possible combinations of intercontinental transport cost b and intrabloc preference k. (As in the other graphs, we only show b up to 0.5 here, under the reasoning that transport costs higher than 50 percent are not plausible.) The solid line represents the level of intrabloc preference that maximizes welfare at each level of transport cost b. Below this line, there are positive returns to regionalization, that is, in this range increasing the degree of preference will result in higher welfare. Above this line, increases in the preference are welfare-reducing. We call this the area of negative returns to regionalization, *NRR*.

Within the NRR area, the dotted line represents, for every level of intercontinental transport cost, the intrabloc preference level that yields the same welfare as $k = 0$ (i.e., MFN). The trade arrangements that lie above this dotted boundary are the ones we call "supernatural" trading blocs. The term "natural" does not seem appropriate to describe trade arrangements that, even when formed along the lines of geographical proximity, represent a movement so deep toward regionalization that welfare is reduced compared to the no-bloc situation.[16]

16. The supernatural bloc area does not always exist. For certain values of the parameters—for example, the combination ($\theta = 0.85$, $t = 0.35$) in the stylized world of three two-country conti-

For $b = 0.10$, our base-case parameter values, and a world consisting of three continents of two countries each, NRR set in when preferences are 49.5 percent. Any greater degree of regional preference moves into the zone of NRR (figures 4.2 and 4.3). For this world, 87.6 percent preferences put the economy into the supernatural zone.

In reality, the status quo that should be compared to continental blocs is a world that contains more than two countries (or FTAs) per continent. We have repeated the experiment for the more realistic, if still stylized, case where the world consists of four continents of sixteen countries each. This sixty-four country setup has the virtue of corresponding roughly to our gravity-model data set examined in section 4.2. (We could get to four continents by adding the Mideast/Africa group of countries to the other three.) NRR set in sooner than before. If intercontinental transport costs are 0.2, then the world attains the welfare optimum as soon as intrabloc preferences reach 10.4 percent, and enters the supernatural zone when they reach 20.4 percent.[17] If intercontinental transport costs are only 0.1, then NRR set in even sooner. For a world consisting of four sixteen-country continents, the optimum degree of continental preferences is 7.6 percent, and the supernatural zone begins at 14.8 percent.

We now return to figure 4.1, to look at the welfare effects of trade agreements, this time not only allowing for less than 100 percent preferences, but also assuming the world trade system chooses the preference level optimally.[18] (We return, for the moment, to a hypothetical world of three continents consisting of two countries each.) The "optimal PTA" curve shows up as welfare-improving no matter what the intercontinental transport costs are. It makes a big difference if the preferences are set at the optimal level, rather than the 100 percent level that is called for in a true FTA.

This application, together with the last one, has provided some answers, within the limitations imposed by the structure of our model, to what Bhagwati (1993) calls the static-impact-effect question regarding the creation of trading blocs. Starting from the absence of trading blocs, a small movement in the direction of increased regionalization (by increasing intrabloc preference) is always a good thing. We can say that there are positive returns to regionalization up to the point of maximum welfare, and NRR thereafter. If intrabloc preferences are set at the optimal level, regionalism will have an immediate

nents—welfare under FTAs is higher than welfare under the MFN rule for every value of transport cost b. This eliminates the possibility of "supernatural" blocs. (An earlier footnote noted some sensitivity analysis with respect to the parameters. In general, the higher θ and t, the less likely blocs will be "supernatural." In addition, the higher θ and t, the higher the optimal preference level k for every level of transport cost b, which translates into a smaller area corresponding to negative returns to regionalization.)

17. If tariffs are levied on the f.o.b. value, the welfare optimum occurs when intrabloc preferences are as low as 27.0 percent, and enters the supernatural zone when they are 51.5 percent.

18. This "optimal" level is not the result of a Nash noncooperative equilibrium, where each bloc chooses the optimal preference level given the preference level chosen by the rest of the blocs (and given the tariff level t). It is just the preference level that maximizes welfare in a symmetric world, and can be interpreted as the cooperative solution (again, given the tariff level t).

positive effect on world welfare. If trade blocs are constrained to have 100 percent preferences (as in article XXIV in GATT), then world welfare could be made lower, assuming transport costs are relatively low and there is sufficient preference for variety.

From the purely static viewpoint of our model, 100 percent preference within a trade bloc is not optimal. Does this mean that GATT should eliminate article XXIV's requirement that FTAs stipulate complete liberalization? Not necessarily. We think that article XXIV can probably be rationalized in a dynamic political-economy framework. The ultimate goal is the achievement of global free trade. The requirement of article XXIV may be the best "dynamic time-path" (Bhagwati 1993) to achieve this goal. An explicit political-economy model that would be helpful for this issue is beyond the scope of this paper.[19]

4.5 Some Estimates of Intercontinental Costs to Evaluate the Extent of Regionalization

Where does the current pattern of trade regionalization lie in the welfare spaces mapped out above? To answer this question, it would be useful to obtain estimates of the key parameters, especially that of intercontinental transport costs, b. In this section, we make an audacious attempt to do this. We cannot claim any precision to the estimates, but hope that the exercise is instructive.

Perhaps the most natural place to look for an estimate of b is the difference between the c.i.f. value of a country's trade and its f.o.b. value as a percentage of its total trade. One disadvantage here is that the data are not comprehensively available on a bilateral basis. Another disadvantage of using aggregate c.i.f./f.o.b. numbers is that they depend on the composition of trade (which is in turn influenced by the true transport costs).

If we were willing to leave aside these deficiencies, we could proceed as follows. The ratio of total worldwide import values, including insurance and freight costs, to export values in 1990 was about 1.06.[20] Assume that 6 percent is a weighted average of intracontinental costs and intercontinental costs: $0.06 = ICS\,a + (1 - ICS)(a + b - ab)$, where ICS is the fraction of the world trade that is between countries on the same continents. In our sample, the ratio of intracontinental to total trade, ICS, is roughly 0.4. Without knowing the value of parameter a, we cannot get an exact estimate of b. But we can infer a rough upper bound: $b = (0.06 - ICSa)/[(1 - ICS)(1 - a)] - a/[1 - a] \leq 0.06/(1 - ICS) = 0.06/(1 - 0.4) = 0.10$. Even if the intercontinental transport cost takes this upper bound, we observe from figures 4.1 and 4.3 that supernatural trading blocs are a real danger.

However, there is reason to think that the 10 percent estimate could in fact

19. A country that joins an FTA may then experience an increase in political support for further steps toward liberalization. See the second half of Frankel and Wei (chap. 7 in this volume).

20. In 1980 it was 1.066 and in 1989, 1.053 (UNCTAD 1991, table 36).

be too low. Although the parameter b is labeled as "intercontinental transport cost" in the model, we noted earlier that it should represent all transaction costs pertaining to intercontinental trade (other than trade barriers imposed by governments). Certain costs such as those associated with personal contact between buyers and sellers are not captured well by the c.i.f./f.o.b. ratio. Recent literature on spillovers and geographic concentration suggests that the effects of proximity on stimulating production are much greater than mere transport costs. In the classic gravity model of world trade, Linnemann (1966) concluded that the effect of distance on trade comes from three channels rather than one: (1) transport costs, (2) the time element (involving concerns of perishability, adaptability to market conditions, irregularities in supply, and interest costs), and (3) "psychic" distance (which includes familiarity with laws, institutions, and habits).

If we were willing to assume that the observed tendency for countries to trade with neighbors was the result solely of these proximity-related aspects that we wish to measure, and not of preferential trading policies, the parameter b could be estimated in a simple way from the data on intraregional trade shares, within the confines of our theoretical model.[21] Such an estimate of b is, however, almost certainly *overstated.* We know from our gravity estimation that statistically significant tendencies toward regional trade preferences already exist, and thus explain part of the proclivity toward intraregional trade that shows up in table 4.1.

For this reason, our preferred approach is to infer the value of b based on estimates of elasticities in our gravity regression. This approach holds constant for the effects of regional trading arrangements already in existence, as well as the effect of per capita GNPs, common languages, and so forth.

We combine several pieces of information. First, in the algebra in section 4.3, the elasticity of demand, $\varepsilon_x = d\log(Trade)/d\log(P)$, is given by $1/(1-\theta)$. Second, in our sample, the mean distance between countries on the same continent is 2,896 kilometers, and on different continents is 11,776 kilometers—four times as great. If transport costs show up fully in the price facing the consumer, the percentage change in price associated with being located on a different continent is given by $(p_{nc}/p_c) - 1 = b/(1-b)$. This follows from equation (4). Third, this price effect, or $b/(1-b)$, should be approximately equal to

$$\frac{d\log(P)}{d\log(DIST)}\log(11{,}776/2{,}896) = \frac{d\log(Trade)/d\log(DIST)}{d\log(Trade)/d\log(P)}1.403$$
$$= [0.50(1-\theta)]\ 1.403.$$

21. Krugman (1991b) and Summers (1991), for example, use simple calculations to infer roughly the importance of distance in determining trading patterns, without explicitly allowing for the effect of existing trade preferences.

Choosing again our baseline value $\theta = 0.75$, our sample calculation suggests that intercontinental transport costs are roughly on the order of 15 percent ($= 0.175/1.175$). It is interesting to note that this estimate is indeed greater than 10 percent, the rough estimate implied by the c.i.f./f.o.b. data.[22] The estimate for b, together with our simulations, suggests that continental FTAs would put us firmly over the line into the supernatural zone.

What about PTAs? If taken at face value, the estimate of $b = 0.15$ together with figure 4.3 suggests that the optimal degree of preferences within a continental grouping, k^*, is roughly 55 percent in a stylized six-country world. In other words, intraregional trade barriers should be lowered to 45 percent of the level of worldwide barriers. When intrabloc preference proceeds past that point, it enters into the zone of negative returns to liberalization. For the more realistic sixty-four-country world, NRR set in as early as at 9 percent preferences and the supernatural zone at around 18 percent preferences.

The last step is to try to extract from our gravity estimates of section 4.2 a measure of k, the degree of preferences prevailing in existing regional trading blocs. If we take our point estimates in column 2 of table 4.2 at face value, they suggest that the EC in 1990 operated to increase trade among its members by roughly 43 percent ($\exp(0.245 + (20 \times 0.0057)) = 1.43$). Other parts of the world have varying intraregional biases. Let us ask the hypothetical question, what would the effect be on world economic welfare if the trading system settled down to an array of continental blocs that each had the same level of preferences as the EC?

Let the percentage effect on trade of bloc formation be represented by γ. Using our model in section 4.2, a bit of algebra reveals that the formation of a bloc with preferences of k lowers the prices of goods in intrabloc trade by $-tk/(1 + t)$. The ratio of the change in quantity to the change in price is equal to the elasticity of demand:

$$\frac{\gamma}{tk/(1 + t)} = \varepsilon_x = 1/(1 - \theta).$$

Solving for the parameter we wish to estimate,

$$k = \gamma(1 + t)(1 - \theta)/t.$$

Taking $\gamma = 0.43$ from the EC estimate, $\theta = 0.75$, and $t = 0.30$, the implied estimate of k is 0.47. In other words, in this illustrative calculation, EC preferences operate to reduce trade barriers by 47 percent for intrabloc trade. This parameter value is not far from the optimum for our three-continent six-country world, but lies in the supernatural zone for our more realistic four-continent sixty-four-country world. It follows, within the assumptions of our model, that if all continents followed the EC example, the regionalization of

22. It is also, as expected, lower than the estimate following the Krugman-Summers approach.

world trade would be excessive, in the sense that world economic welfare would be reduced relative to the MFN norm.

The tentative conclusion of this study is that some degree of preferences along natural continental lines, such as a Western Hemisphere PTA or enlargement of the EC into a European economic area, would be a good thing, but that the formation of FTAs where the preferences approach 100 percent would represent an excessive degree of regionalization of world trade, within the confines of our static economic model. The overall conclusion is that the world trading system is currently in danger of entering the zone of excessive regionalization.

The optimal path to liberalization apparently features a sharp departure from article XXIV. It entails reducing intracontinental barriers partway, for example by only an estimated 10 percent or so. The strategy of concentrating on reducing trade barriers at the multilateral level *before* (or at the same time as) liberalizing completely within any one continental trading arrangement appears under our assumptions to be preferable.

References

Anderson, James. 1979. A Theoretical Foundation for the Gravity Equation. *American Economic Review* 69 (March): 106–16.

Anderson, Kym, and Hege Norheim. 1993. History, Geography, and Regional Economic Integration. In Kym Anderson and Richard Blackhurst, eds., *Regionalism and the Global Trading System.* London: Harvester Wheatsheaf.

Balassa, Bela. 1987. Economic Integration. In *The New Palgrave Dictionary of Economics,* 43–47. London: Macmillan.

Bergstrand, Jeffrey. 1989. The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade. *Review of Economics and Statistics* 71 (February): 143–53.

———. 1993. Regionalism and Multilateralism: An Overview. In J. de Melo and A. Panagariya, eds., *New Dimensions in Regional Integration.* New York: Cambridge University Press.

Deardorff, Alan. 1984. Testing Trade Theories and Predicting Trade Flows. In R. Jones and P. Kenen, eds., *Handbook of International Economics,* 1:467–517. Amsterdam: Elsevier.

Deardorff, Alan, and Robert Stern. 1994. 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.

Frankel, Jeffrey. 1993. Is Japan Creating a Yen Bloc in East Asia and the Pacific? In Jeffrey Frankel and Miles Kahler, eds., *Regionalism and Rivalry: Japan and the U.S. in Pacific Asia.* Chicago: University of Chicago Press.

Frankel, Jeffrey, Ernesto Stein, and Shang-Jin Wei. 1995. Trading Blocs and the Americas: The Natural, the Unnatural, and the Super-Natural. Abridged version, *Journal of Development Economics* 47:61–96.

Frankel, Jeffrey, and Shang-Jin Wei. 1994. Yen Bloc or Dollar Bloc: Exchange Rate

Policies of the East Asian Economies. In T. Ito and A. Krueger, eds., *Macroeconomic Linkage*, 295–329. Chicago: University of Chicago Press.

Hamilton, Carl, and L. Alan Winters. 1992. Opening Up International Trade in Eastern Europe. *Economic Policy* 14:77–117.

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

Helpman, Elhanan. 1987. Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries. *Journal of the Japanese and International Economies* 1:62–81.

Helpman, Elhanan, and Paul Krugman. 1985. *Market Structure and Foreign Trade.* Cambridge: MIT Press.

Krugman, Paul. 1980. Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review* 70:950–59.

———. 1991a. Is Bilateralism Bad? In E. Helpman and A. Razin, eds., *International Trade and Trade Policy.* Cambridge: MIT Press.

———. 1991b. The Move toward Free Trade Zones. In Federal Reserve Bank of Kansas City, *Policy Implications of Trade and Currency Zones,* 7–42. Kansas City: Federal Reserve Bank.

Linnemann, Hans. 1966. *An Econometric Study of International Trade Flows.* Amsterdam: North-Holland.

Meade, James. 1955. *The Theory of Customs Unions.* Amsterdam: North-Holland.

Panagariya, Arvind. 1995. The Free Trade Area of the Americas: Good for Latin America? Center for International Economics Working Paper no. 12. College Park: University of Maryland.

Petri, Peter. 1993. The East Asian Trading Bloc: An Analytical History. In J. Frankel and M. Kahler, eds., *Regionalism and Rivalry: Japan and the U.S. in Pacific Asia.* Chicago: University of Chicago Press.

Saxonhouse, Gary. 1993. Pricing Strategies and Trading Blocks in East Asia. In Jeffrey Frankel and Miles Kahler, eds., *Regionalism and Rivalry: Japan and the U.S. in Pacific Asia,* 89–119. Chicago: University of Chicago Press.

Srinivasan, T. N. 1993. Comment. In J. de Melo and A. Panagariya, eds., *New Dimensions in Regional Integration.* New York: Cambridge University Press.

Stein, Ernesto. 1994. The Welfare Implications of Asymmetric Trading Blocs. In Essays on the Welfare Implications of Trading Blocs with Transportation Costs and Political Cycles of Inflation, Ph.D. diss., University of California, Berkeley.

Stein, Ernesto, and Jeffrey Frankel. 1994. The Welfare Implications of Trading Blocs in a Model with Transport Costs. Pacific Basin Working Paper no. PB94–03. San Francisco: Federal Reserve Bank, May.

Summers, Lawrence. 1991. Regionalism and the World Trading System. In Federal Reserve Bank of Kansas City, *Policy Implications of Trade and Currency Zones,* 295–302. Kansas City: Federal Reserve Bank.

Tinbergen, Jan. 1962. *Shaping the World Economy.* New York: Twentieth Century Fund.

UN Conference on Trade and Development (UNCTAD). 1991. *Review of Maritime Transport 1990.* New York: UNCTAD.

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

Wonnacott, Paul, and Mark Lutz. 1989. Is There a Case for Free Trade Areas? In Jeffrey Schott, ed., *Free Trade Areas and U.S. Trade Policy,* 59–84. Washington, DC: Institute for International Economics.

Comment

Paul Krugman

A warning to any economist who works in applied theory: your models may have real consequences. You may find yourself at the head of a large movement. You may say to yourself, "This is not my movement"; you may say to yourself, "My God, what have I done?" But there it is.

When I wrote down a stylized model of the consequences of regional trading blocs a number of years ago—a model that was intended to provide a language for discussing the competing hopes and fears regarding such blocs, rather than to be a basis for serious policy analysis—I did not expect to launch a subliterature. But in a way it should not be surprising that papers that take off from that original, almost tongue-in-cheek effort (either building on the original model or, with considerable justification, calling the whole approach into question) should have become a minor academic industry, and even have begun to have some influence on actual policy discussions. After all, a model—even a crude, small, somewhat silly model—often offers a far more sophisticated, insightful framework for discussion than scores of judicious, fact-laden, but model-free pontifications. But while worrying too much about realism can be a very bad thing—there is nothing worse than the would-be wise man who knows all the facts but has no sense of how they might fit together—it is also always a bit worrying when very simple models are made to bear too much weight. One always wants to stand back and ask whether the simplicity of the model is missing too much.

Now I found this paper by Frankel, Stein, and Wei completely admirable. Not only does it develop a very clever way to take account of the role of geography in influencing the impact of trade blocs; it makes ingenious use of econometric results to make the model, if not exactly empirical, at least constrained by statistical evidence. I would never have thought one could get so far with this general approach. Moreover, this paper is part of a very important wider project by the authors, which has turned the classic gravity approach to trade modeling into a tool of policy-oriented research in a way nobody had thought of before.

So I have no criticisms. I do, however, have two observations.

The first, minor concern is one of practical relevance. When I first wrote about trading blocs, it was an item of faith among many commentators that regional trade liberalization was the wave of the future—that the political success of the European Single Market and the Canada-U.S. Free Trade Agreement would soon be matched by a series of regional integrations. Meanwhile the multilateral system was regarded as being in desperate straits—those were the days when Lester Thurow's pronouncement that "the GATT is dead"

Paul Krugman is professor of economics at the Massachusetts Institute of Technology and a research associate of the National Bureau of Economic Research.

was taken very seriously by world leaders. But that was a long time ago. Since then, the Uruguay Round has been successfully passed, albeit in more modest form than originally hoped. And meanwhile, regional trade liberalization, while it has continued, has proved to be far more difficult politically than people had imagined. Maybe I can summarize this briefly by saying that a ninth GATT round these days seems more plausible to me than an extension of NAFTA to include Brazil or an extension of the European Economic Area to include Ukraine. If there is to be a world of regional trading blocs, it seems likely at this point that it will at the very least involve some distinction between the advanced-country cores and developing-country peripheries within each bloc. This, too, could be modeled; a crucial question would then be the division of gains within each bloc between core and periphery.

My second, more analytical concern is with the way Frankel, Stein, and Wei map the theoretical model onto the real world. In their conclusion that actual regionalization may well be "supernatural," they make use of stylized worlds in which there are several continents consisting of a number of countries—for example, a world of four continents of sixteen countries each. At first this seems more or less right—if one counts North and South America as a single unit, there are indeed four major inhabited continents, with close to a hundred national units among them. But there is a crucial assumption in the model that is not nearly true of the real world: that countries themselves are of equal economic size. In reality, of course, the size distribution of GDPs is highly unequal, and this surely makes a major difference when we try to model the effects of integration.

How should this be accommodated within the model? One answer would be to put the real size distribution of countries into the analysis. In a way, however, the whole point of this style of model is to assume away asymmetries among countries, so as to avoid the mind-numbing taxonomies of customs-union theory. This suggests that we might instead try to loosely deal with the unequal size of nations by using, not actual numbers, but some index of the number of "country-equivalents"—say the inverse of a Herfindahl index.

If we do this, the picture of the world is strikingly different. I recently made an estimate of the number of country-equivalents, using the inverse of the sum of squares of world GDP; while this index has risen over time, with the erosion of U.S. dominance, there are still fewer than ten country-equivalents in the world economy as a whole. Exactly how this observation should be mapped into the model is arguable, but I would suggest that a world of three continents with three "countries" each is in some sense closer to the truth than the version the authors actually use.

In the end, of course, one cannot avoid the asymmetries. Regional trade liberalization will not have the same effects on the United States and Mexico, on Germany and Estonia. It was always a distorting assumption to imagine that we could suppose otherwise; but this paper shows that the assumption has proved far more productive than I had any right to expect.

Comment T. N. Srinivasan

The revival of interest in regional preferential trading arrangements (RPTAs) in the late eighties after the failure of many such arrangements (except the European Community [EC] and the European Free Trade Association) in the past was in part triggered by the fear that the ongoing Uruguay Round of multilateral trade negotiations would fail and the global trading system would end up in a few trading blocs. Surprisingly, even after the successful conclusion of the Uruguay Round with the signing of its Final Act by countries in April 1994 and the establishing of the World Trade Organization in January 1995, the interest in RPTAs has not only not disappeared but has gathered further momentum. Countries in the Western Hemisphere, members of the Asia-Pacific Economic Cooperation (APEC) forum, the Association of Southeast Asian Nations (ASEAN), and others have agreed to remove all barriers to trade among themselves in the next two decades. Whether RPTAs are stepping stones or stumbling blocks in the path toward a liberal global trading system continues to be debated. It is being suggested that the proposed RPTAs differ from those in the past in that they go far beyond border measures and cover ostensibly trade-related but domestic policies as well. Nonetheless, a convincing case has yet to be made as to why removal of such policy distortions is feasible only through regional approaches.

The implication of the formation of a customs union (CU), free trade area (FTA), or more generally an RPTA for the welfare of the residents of member and nonmember nations has been a central analytical issue for economists, starting with Jacob Viner. A concern for the welfare of nonmembers can be seen in article XXIV of the General Agreement on Tariffs and Trade (GATT), 1947, on CUs and FTAs. It requires in part that barriers to trade of nonmembers should not go up whenever a CU or FTA is to be formed with a waiver from GATT's most-favored-nation (MFN) principle. One of the central propositions of the analytical literature is that associated with Kemp and Wan (1976), who showed (using a standard neoclassical general-equilibrium setup) that there exists a common tariff structure for a CU consisting of any arbitrary number of countries, which ensures that no consumer outside the union is made worse off and at least one consumer within the union is better off compared to the preunion global equilibrium, as long as lump-sum transfers among consumers *within* the union are feasible. Thus a path ending in global free trade consisting of successive enlargements of CUs, each enlargement being Pareto noninferior to the preceding one, exists in the Kemp-Wan world. Of course, whether a CU in the real world satisfies the Kemp-Wan condition is another question. Two more recent propositions established by Krugman (1991a, 1991b) for a world

T. N. Srinivasan is the Samuel C. Park, Jr., Professor of Economics at Yale University.

The author thanks Edward Leamer and Philip Levy for their valuable comments on an earlier draft.

of many identical countries are that, in the absence of transport costs, the number of trading blocs that *minimizes* global welfare is two or three; however, in a world of identical countries located in continents separated by almost prohibitive transport costs, continental trading blocs are welfare-improving because such blocs create trade among members without diverting trade away from nonmembers.

The paper of Frankel, Stein, and Wei is in two, in effect *unrelated,* parts, the first consisting of a simple econometric analysis of actual bilateral trade flows, and the second of numerical analysis from a stylized model that is a generalization of the Krugman setup. In the first part, bilateral *total* trade flows among sixty-three countries for 1970, 1980, and 1990 are put through the well-known gravity model but with a few additional explanatory variables (apart from dummies for the years 1980 and 1990) such as dummies for adjacency of the countries, membership in regional country groupings (EC or Western Europe, East Asia, and Western Hemisphere) and, in one version, separate time trends for each country group as well. The authors find that all three regional dummies are positive and statistically significant, so that any two countries within any one of the groups traded significantly more with each other than with two otherwise similar countries that are not within the same region. The authors view this as indicating a pronounced "regional bias" in trade. Since distance, adjacency, and so forth have been controlled for and, except for the EC, *the other two regions were not formally a trading bloc* during the period covered by the data set, I would argue that the coefficient of regional dummies, like that of any other dummy (or should I say "dumb") variable, merely assigns a quantitative magnitude for ignorance! In other words, for reasons that are not known and hence not captured in the explanatory variables, such as GNP, distance, and so forth, two countries within a region trade more with each other by a percentage indicated by the magnitude of the coefficient of the dummy.[1] The effect of the formation of a trade bloc is not the coefficient of the regional dummy, but the change in it in separate regressions *before* and *after* the region formally becomes a trade bloc.[2]

Interestingly, a time trend interacting with the regional dummy is significantly positive in the case of Western Hemisphere and significantly negative in the case of Western Europe. In the Western Hemisphere, however, the positive trend is so strong that it offsets the *negative coefficient* of the regional dummy (i.e., a *bias* away from trade within the region) in about five years, whereas in

1. The authors conclude from the fact that EC dummy is 0.31, for example, that any two EC countries trade 36 percent (exp (0.31) = 1.36) more than two otherwise similar countries. Strictly speaking, since the estimated coefficient for large samples is normally distributed, its exponential is log-normally distributed. Now, if X is log-normally distributed, i.e., if $\log X$ is normally distributed with mean μ and variance σ^2, then $EX = E[\exp \log X] = e^{\mu+(1/2)\sigma^2}$. Frankel, Stein, and Wei have ignored the $\frac{1}{2}\sigma^2$ term and as such underestimate the "regional bias"! I should also add that, since measured GNP, GNP per capita, and distance contain errors relative to their true values, the estimated regression coefficients are biased. The authors do not recognize this.

2. I owe this observation to Edward Leamer.

the case of Western Europe, a positive coefficient of regional dummy (i.e., a bias in favor of regional trade) is almost eliminated by the negative trend. The authors do not comment on this. Be that as it may, the short empirical section based on gravity model does not help in understanding the reasons for regional bias, if any, or its welfare implications for each region or those outside the region. As such, it does not have much to contribute to explaining the current enthusiasm for RPTAs and to the debate about its consequences for multilateralism.

Let me turn to the stylized model and numerical analysis based on it. I cannot emphasize enough that the model is *extremely special:* countries are *identical,* an *identical* number (N) of them are located in each of a number (C) of continents. The transport cost (measured in terms of the Samuelson iceberg-melting coefficient) between *any* two countries within any continent is the *same* and that between *any* country in *any* continent and any country in *any other* continent is also the same. All countries produce varieties (with costless product differentiation) from a continuum of possible varieties using the *same* technology that requires a fixed cost in terms of labor (the only factor of production) as well as a constant marginal cost, again in terms of labor. Labor is identical in productivity everywhere and is inelastically supplied. Consumers anywhere have the same utility function that is the *sum* of the utility from consuming each variety, which is assumed to be a constant elasticity function of consumption. Market structure is of monopolistic competition of the large group with free entry.[3] Of course, almost by definition all models are stylized representations of a complex reality. But a robust model is one that simplifies inessential details of the real world. As I argue below, the Krugman model, while robust for analyzing some aspects of trade, is not so far the case that Frankel, Stein, and Wei make of it.

Initially all countries have the same tariffs on all imported varieties, whether from a country within or outside their continent. In all experiments, the authors consider only symmetric formation of equal-sized blocs around the world. In the first experiment, there are three continents with two countries each, with no transport cost within continents. "Natural" blocs are continental blocs, while "unnatural" ones involve countries in two different continents. The welfare effects depend on the level of intercontinental transport costs. If they are sufficiently low, FTAs along "natural" blocs are welfare-reducing. For even lower level of transport costs, unnatural blocs are welfare-reducing. This is consistent with Krugman's first proposition.

3. Presumably this model due to Krugman is by now so familiar that the authors do not state all its technology, taste, and market-structure assumptions! Because of this, one is likely to miss the facts that their normalization of number of varieties, price of each variety, and the wage rate to be unity implies relationships between technology and taste parameters and the labor endowment and is dependent on the extreme symmetry of the model. I am not sure that this normalization is innocuous, except, possibly, in the "symmetry" case; after all, in a model with scale economies, the size of the economy (here, the size of the labor endowment) matters. For example, if countries differ in size, normalizing all three variables to unity is not feasible.

In the second experiment, within each bloc there is only partial liberalization in the sense that the tariff applicable to imports from a country within the bloc is a proportion of the tariff applicable to imports from outside. Unsurprisingly it turns out that, given any level of intercontinental transport cost, the optimal margin of preference for imports within a "natural" bloc is strictly above zero so that a *preferential trading* arrangement is welfare superior to a *free trade* arrangement within a bloc. The authors view this finding as implying that the requirement of GATT article XXIV that *all* trade within a CU or FTA be free is inappropriate, though they recognize there are other political-economy considerations for the requirement. The "optimal" margin of preference is shown to be an increasing function of the intercontinental transport costs, so that welfare increases (respectively falls) as regional preference increases from any level below (respectively above) the optimal level. If the regional preference margin is increased sufficiently above the optimal level, welfare falls below its no-trading-bloc (i.e., regional preference margin of zero) level.

The authors make what they appropriately term an "audacious attempt" to obtain estimates of their key parameters in order to find out whether the current pattern of regionalization has gone beyond the welfare-improving range. Quite rightly they do not claim any precision to their analysis, while concluding that "within the assumptions of our model . . . if all continents followed the EC example, the regionalization of world trade would be excessive, in the sense that world welfare would be reduced relative to the MFN norm."

It should be obvious that the authors' model is extreme not only in its symmetry but also in the assumption that *all* trade is based on preference for variety, market structure is of the Chamberlinian monopolistic competition of the large group with free entry, and transport costs are of the iceberg-melting variety. Indeed, the last assumption (and the assumption that tariffs are ad valorem) ensures that the price *gross* of costs of transport and tariffs is proportional to the price *net* of such costs, so that the relevant demand price elasticities (as perceived by each producer) are unaffected. Because of unchanging price elasticities, producer decisions, in particular production-level producer prices and number of varieties produced, are *unaffected* by tariffs or transport costs. Certainly the extreme symmetry and other assumptions make it easy to derive analytically the global equilibrium and numerically simulate the comparative statics. However, whether the simulations are more than illustrative of the possibilities is arguable. I am not convinced that the simulations are enough to conclude, even tentatively, that "some degree of preferences along natural continental lines, such as a Western Hemisphere PTA or enlargement of the EC into a European economic area, would be a good thing, but that the formation of FTAs where the preferences approach 100 percent would represent an excessive degree of regionalization of world trade." To be fair, the authors refer to other works in which the symmetry assumption is relaxed to a limited extent and comparative advantage considerations are introduced, with some significant differences in their conclusion. But robust policy conclusions have to await less stylized and more realistic and empirically grounded analyses.

References

Kemp, M. C., and H. Y. Wan, Jr. 1976. An Elementary Proposition concerning the Formation of Customs Unions. *Journal of International Economics* 6, no. 1: 95–97.

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

———. 1991b. The Move toward Free Trade Zones. In Federal Reserve Bank of Kansas City, *Policy Implications of Trade and Currency Zones,* 7–42. Kansas City: Federal Reserve Bank.