6

Extensions of the Empirical Analysis

The gravity equation we have estimated leaves out many factors. If an omitted factor is correlated with the existence of a regional trading arrangement, distance, or one of the other variables that we are interested in, we will get biased coefficient estimates. Thus it is important that we take into account as many other factors as we reasonably can. We begin in this chapter with a consideration of political and historical factors before moving on to the roles of investment, currencies, and many other factors. It is inevitable, of course, that we cannot take into account everything. Often the work of others will help us learn what we want to know. But in the end, we will see that the basic bottom line is unchanged: statistically significant bloc effects are arising in many parts of the world.

The Importance of Political Factors in Trade

An interesting trivia question illustrates the importance of political and historical factors in trade: who is Congo Zaire's biggest trading partner? Criteria of proximity or size will give the wrong answer. The country with which the former Belgian Congo conducts the most trade, even in absolute terms, is Belgium. We have already captured some of this effect, in that our language variable counts the links between former colonies and colonizers (although the Congo Zaire is missing from the set of countries considered in our estimates, for lack of recent data). The colonial link was severed in 1960. Even when the original reason for a high level of bilateral trade has disappeared, however, the stock of capital that firms

have invested in the form of marketing and distribution networks, brand-name loyalty among customers, and so forth lives on for many years thereafter. The word hysteresis is sometimes applied to this phenomenon, suggesting that the effect is considered to be permanent.

Political Alliances and Enmities

We noted earlier that trade between India and Pakistan is impeded by their historical animosity. Their trade is 70 percent lower than trade between two otherwise identical economies. We also noted in chapter 2 that Bhagwati (1992, 1993a) is suspicious of the claim that proximity is an important determinant of trade, citing the example of India and Pakistan, and that he is therefore skeptical of the notion of natural trading blocs. He asserts that the high levels of intraregional trade that are already observed in such areas as Europe must be the result of free trade areas (FTAs) and other preferential trade arrangements that are already in place. The issue becomes an important one for policy when other economists, such as Krugman and Summers, argue that proximity does promote trade and propose that regional trading arrangements be pursued on the grounds that it is *natural* for neighbors to trade with each other.[1]

The gravity-equation estimates convinced many of us some time ago that distance is in fact a very important determinant of trade. But special historical attractions or repulsions also matter, independent of distance. In South Asia (as well as the Mideast; see Fischer 1993, 434-36), it is true that neighbors do not necessarily trade more with each other. Historical enmity has indeed reduced trade between India and Pakistan. Perhaps the root of Jagdish Bhagwati's skepticism regarding the role of proximity in trade is that he has been heavily influenced by this one observation.

To repeat, the gravity model clearly shows that proximity is in general an important determinant of bilateral trade around the world, notwithstanding exceptions such as India-Pakistan, Iran-Iraq, Israel-Syria. Ideally, one would use a dummy variable to represent all pairs with a recent history of strong political or military conflict, especially including embar-

1. Panagariya (1995, 9-10) echoes Bhagwati's suspicions. He attacks Summers's argument that an FTA among natural bloc partners is less likely to be trade-diverting, with ''natural'' defined by a low level of trade with countries outside the group. To the extent that a low observed level of trade reflects natural barriers, such as distance, our results support Summers and Krugman's. Bhagwati and Panagariya (1995) are correct to point out, however, that high bilateral trade volume alone does not imply a natural partnership if it exists for reasons *other* than low transportation costs or geographical proximity. Summers (1991) seems to imply otherwise. The proposition that one criterion for a trade-creating, welfare-improving FTA is that a high proportion of trade be conducted among partners *ex ante* goes back to Lipsey (1960). But it seems evident that, to the extent such trade is the result of preferences already in place, it cannot be used to justify further preferences. These issues are addressed more formally in the next chapter.

goes and boycotts. This variable would in essence be the antithesis of the dummy variable for linguistic and colonial links. The distance and adjacency effects are so strong, however, that they show up as highly significant statistically even when no account is taken of the antagonist pairs.

The effects on bilateral trade of politico-military alliances, wars, colonial relationships, and other political factors have been extensively examined by Mansfield (1993), Mansfield and Bronson (1994, 1997), and Gowa and Mansfield (1993). Theoretically and empirically (within the gravity framework), they find, as one might expect, that trade is generally higher among countries that are allies and lower among countries that are actual or potential adversaries. If two countries are at war, trade usually falls off. The reduction runs as high as a highly significant 99 percent reduction [$=\exp(-4.90)$] in 1965. More typical is a highly significant 82 percent reduction [$=\exp(-1.74)$] in 1990.

The significance of distance and the other gravity variables tends to be unchanged. The log of distance still has a negative and highly significant effect in every cross-section, always in the range from -0.51 to -0.69 (estimated at five-year intervals, from 1950 to 1990). The formation of a preferential trade arrangement promotes trade between two countries by about 100 percent if they were not already military allies and by about 240 percent if they were. (The effects are different if the alliance involved either of two major powers: the Soviet Union and China.) Once again, we see that both proximity and existing preferences are important determinants of trade patterns.

The Example of the Canadian Provinces

What if two geographical units are not just politico-military allies but actually enter into a political union? We noted in chapter 4 that there would be great advantages from having data at the level of states or provinces within countries. We could then ascertain how common membership in a political union affects trade between two geographical entities. For example, it would help us predict increases in the volume of trade among the members of the European Union, in the event that they achieve full political federation. We have learned that when two geographical units share such links as a common language their bilateral trade is clearly boosted. It stands to reason that when two units share a common cultural heritage or legal system, their trade will be enhanced by even more.

Data are not generally available on trade among US states, Japanese prefectures, German Lander, British counties, or French departments. But there are data on trade undertaken by Canadian provinces, among each other and with major American states. They show a strong intranational bias to trade. Ontario exports three times as much to British Columbia as to California, even though the latter has 10 times as many people. (The figures are for 1988.)

McCallum (1995) has applied the gravity model to trade among the Canadian provinces and American states. The usual effects of size and distance show up. The fascinating result is the effect of a dummy variable to represent when two states or provinces lie in the same country. Two such provinces trade 22 times as much with each other as would a province and a state that are otherwise similar but lie on opposite sides of the border. Helliwell (1995) has updated this test. He finds that the intra-Canadian bias factor is 20.5 in 1988, 18.5 in 1989, and 24.8 in 1990. There are not enough data to read much meaning into the trend; one should settle for an average estimate of 21.[2] These results do not hold constant for provinces' adjacency, per capita incomes, or remoteness. It is possible that these omissions explain some of the intracountry bias, though this cannot provide much of the explanation.

The result is reminiscent of the striking finding, in Engel and Rogers's 1994 paper ''How Wide Is the Border?,'' that crossing the Canadian-US border adds as much to the relative price variability between two cities as does traversing a distance of 2,500 miles within either country. This tendency for Canadian provinces to trade with each other is all the more surprising because they tend to maintain trade barriers against each other, never having had the advantage of a constitution that, like the United States', makes trade policy the exclusive responsibility of the federal government. A notorious example is intra-Canadian trade barriers against beer.

Nevertheless, the McCallum-Helliwell and Engel-Rogers's results mean that there are powerful, unseen forces working to promote trade among the provinces. Reasons for the intra-Canadian bias in trade include the ease of doing business within the same legal system, an integrated media and advertising sector, nationwide store chains, and an East-West railroad network.[3] Helliwell and McCallum (1995) ''suspect that the answers lie in a whole host of educational and geographic ties based on migration and family ties and supported by networks of transportation, communication and education, along with portability of health care and pension rights—if not completely of beer.'' Presumably, the sources of intranational bias are even stronger for countries that do not share the cultural proximity and liberalized trade relations of Canada and the United States.

---

2. The Canadian bias for Quebec trade is even higher than for other provinces, at 26.0. This comes as a surprise in light of Quebec's insistence on the separateness of its culture. The explanation is that, even though Quebec's trade with other Canadian provinces is, as one expects, estimated to be lower than the gravity variables would predict, Quebec's trade with American states is lower still. Quarrelsome Quebec apparently is even more separate from the United States than it is from the rest of Canada.

3. According to Rousslang and To (1993), domestic costs of transporting, wholesaling, and retailing foreign-produced goods are in general so much greater than the analogous costs for domestic-produced goods as to add an estimated 12.7 percent to the total cost. By

This experiment is an instructive one for thinking about the likely effects of political integration or disintegration in other parts of the world. The Canadian federation of provinces or the American federation of states provide possible models for the European Union. The very high effects of political union that were estimated in the Canadian-US case tell us that trade among European countries would increase dramatically (as much as 21-fold) if the European Union attained the same degree of political integration that Canada and the United States have each achieved within their borders.[4] At the same time, however, the exercise tells us that the European Union has a very long way to go before attaining that degree of integration. The formation of a common market turns out to be a relatively small step by comparison.

Four Examples of the Breakup of Federations

Historical examples of political unions that split apart offer further insights into the effect of federation on trade and into the lags involved after a major change in political ties.

The Austro-Hungarian Empire

One interesting example is the breakup of the Hapsburgs' Austro-Hungarian Empire at the end of World War I. Predissolution data are of limited availability. (This is a problem even with some more recent examples, such as the separation of Pakistan and Bangladesh, the secession of Eritrea from Ethiopia, or the breakup of Yugoslavia.) A League of Nations study (Layton and Rist 1925) compared 1924 data on trade between Hungary and Austria-Czechoslovakia (i.e., the former Kingdom of Austria) with 1913 data on trade between the corresponding regions. Its findings imply that if the Austro-Hungarian union had been preserved, trade flows would have been two and a half times what they actually were in 1924. A more recent study by de Menil and Maurel (1994) applies the gravity methodology and finds that, five years after the breakup of the empire, trade among Austria, Hungary, and Czechoslovakia was still almost four times as large as what would be expected from the fundamental determinants. They conjecture that the prewar effects of the Austro-Hungarian Empire on trade must have been similar to the estimated effects of the British Empire, and on this basis they concur with the League of Nations' finding that the breakup reduced trade by more than half. The two statis-

comparison, they estimate the barriers imposed by tariffs and international transportation costs combined at only 7.2 percent.

4. Our results indicate that as most European countries still speak different languages than their neighbors, intra-European trade will fall a bit short of intra-US trade. Canada is a slightly closer standard of comparison in this regard.

tics together imply a total Austro-Hungarian prewar effect on trade of about eightfold. The authors also remark that the partial persistence of the effect into the 1920s must reflect ''a network of business and personal relations that it takes time to build, and that does not decay instantaneously.'' However, the concentration in trade among the constituents of the former empire, which was still highly significant in 1924-29, had apparently disappeared by 1930-32 (Maurel 1995).[5]

The Federation of Malaya

Another example is the 1965 breakup of the Federation of Malaya, which had only recently attained its independence from Great Britain, into the two sovereign countries of Singapore and Malaysia. Trade between the two countries grew subsequently, but only because each of the economies grew. Adjusting for their size in world markets, trade between Singapore and Malaysia fell by 2 percent in the year of dissolution (relative to 1964) and had fallen by 8 percent by 1967. The Malaysia-Singapore example illustrates again that political federation or dissolution matters. The effect is much smaller than in the Austro-Hungarian case, perhaps because the Federation of Malaya had only existed a short time while the Hapsburg Empire had been around for many centuries.

The Former Soviet Union

The most important breakup of our time is the dissolution of the Soviet Union. The data on trade among the republics of the former Soviet Union exist, prebreakup as well as postbreakup. Unfortunately, there are severe measurement problems, associated in particular with proper valuation of the trade. Furthermore, we must recognize that the thoroughly nonmarket nature of trade among former Soviet republics makes it a fundamentally different case from federations in the rest of the world. Between 1990-91 and 1992-93, the share of trade that the former Soviet republics conducted among themselves fell by more than half, as one would expect. Trade with the rest of the world also fell sharply, however. The most evident explanation of why all trade flows fell is that incomes fell sharply.[6] The simple trade-concentration ratio, which divides the intra-Soviet trade share by the Soviet share of world trade and which was very high to start with, more than doubled between 1990-91 and 1992-93.

---

5. The effect reappeared in 1933-38. The fluctuations in the 1930s are probably influenced by new shifts in regional relationships at least as much as by the gradual unwinding of the empire's legacy.

6. Interrepublic trade on the whole tended to fall much more rapidly than incomes from 1991 to 1992, but some of the republics' trade with Russia actually rose relative to their incomes (Stern 1994, 15-16, 19).

Table 6.1 Czech-Slovak trade, 1991-94[a]

	Trade totals (billions of dollars)					Trade ratios		
Year	World	Czech Republic	Slovak Republic	Bloc total	Intrabloc	Intra-regional	Simple intensity	Corrected intensity
1991	7,028.54	n.a.	n.a.	34.72	13.24	0.38	77.2	154.0
1992	7,518.67	n.a.	n.a.	35.37	12.07	0.34	72.5	144.7
1993	7,483.70	25.91	12.11	38.02	7.29	0.19	37.7	75.3
1994	8,535.50	29.02	13.41	42.43	9.16	0.22	43.4	86.6

n.a. = not applicable

a. Czechoslovakia officially dissolved on 1 January 1993.

Source: IMF, *Direction of Trade Statistics* and *International Financial Statistics*; *Czech Foreign Trade*, May 1993, 6, June 1993.

This figure runs counter to all the other examples. One likely possibility is that the former Soviet Union's share of world trade is overstated by the use of official exchange rates in 1990-91 and perhaps understated in 1992-93.[7] Another likely possibility is that a severe loss in trade credit and foreign exchange reserves impeded trade with the rest of the world more than it impeded trade among the republics, especially with Russia itself, much of which continues on a countertrade basis. A similar effect was observed over this period among Eastern European states: intraregional trade did not fall as rapidly after the dissolution of the Soviet bloc as the decline in incomes and the other determinants of the gravity model would lead one to expect (Brada 1993). Both within the former Soviet Union and within Eastern Europe, the barter-like arrangements the countries had long used with each other appear initially to have withstood the tremendous disruptions of the period better than did their fragile trade links with the rest of the world. In this case, we would expect to see a more normal pattern of trade to have begun to emerge subsequently—a pattern more oriented to the rest of the world. One measure, which holds the weight in world trade at 1990 levels to avoid valuation problems but allows trade-to-GDP ratios to vary, shows intra-Soviet concentration peaking in 1992 and then falling in 1993 and 1994. This supports the notion that trade may indeed be starting to return to more normal patterns.

The "Velvet Divorce" Between the Czech and Slovak Republics

Czechoslovakia officially divided into two republics on 1 January 1993, resuming the dissolution of the old Hapsburg Empire that had begun 75 years earlier. Data on trade between the Czech and Slovak halves exist both before and after the "velvet divorce." As table 6.1 shows, the ratio

7. An alternate measure based on implicit exchange rates, however, gives similar results for the concentration ratio while not even showing a decline in the simple intraregional trade ratio. A third measure, at domestic prices, is the only one available for 1987-91; it confirms the simple intraregional trade share at .7 in 1990—the same as the other two

of intra-Czechoslovak trade to total trade was a high .38 in 1991. Trade between the two republics declined slightly in 1992, as the dissolution was contemplated, and then fell sharply in 1993, when dissolution took place. Meanwhile, trade with the rest of the world was growing. Thus, the intra-Czechoslovak ratio fell sharply in 1993, to half its 1991 level. (It then recovered slightly in 1994.) The intensity ratios, which adjust for Czechoslovakia's share of world trade, tell the same story. The net effect of the dissolution, from 1991 to 1994, was a decline in trade between the Czech and Slovak regions of about 44 percent relative to what could be expected had the federation endured. Stated in reverse, Czech-Slovak trade in 1994 would have been higher by a factor of about 2.3 if the union had continued.

An Example of Reunification: Germany in 1990

While a number of countries have split up over the last three decades or so (the Malayan Federation, Pakistan, the Soviet Union, Yugoslavia, Ethiopia, Czechoslovakia), there is one major example of reunification: the reabsorption of the eastern Lander into West Germany in 1990 after the fall of the Berlin Wall. Data are available on trade undertaken by the two halves before and after. The reunification is an important experiment, because cultural and geographic proximity of the two Germanys dictate a high natural level of trade, and yet there was a severe unnatural barrier between the two until it was suddenly lifted. It is also a chance to observe an asymmetry, under which trade could respond more or less quickly when barriers are removed than when they are erected.

Tables 6.2a and b shows a sixfold increase in the level of intra-German merchandise trade from 1989 to 1994. The share of intra-German trade increased fourfold over this period. This change reflects not only the redrawing of national borders, but also the radical switch of the East German economy to a market system.

There were other major influences on East Germany's trade at this time. On the one hand, trade with all Western countries was opened up at the same time as trade with West Germany; on the other hand, a recession in East Germany may have depressed trade generally. Computing the trade concentration ratio is the simplest way of adjusting for these influences. It shows that intra-German trade increased fourfold between 1989 and 1994. The biggest increase comes in 1991, as expected, with steadily diminishing increments thereafter.

The fourfold positive effect of reunification on trade after four years is much greater than the negative effect that we estimated for the breakup of the Malayan Federation (or the effect for the Soviet Union, which we

measures—and reports that this statistic and the concentration ratio had been steady over the preceding four years.

Table 6.2a Intra-German goods trade, 1984-94[a]

Year	Trade totals (billions of dollars)			Trade ratios		
	World Trade[b]	Total German[b]	Intra-German[c]	Intra-regional	Simple intensity	Corrected intensity
1984	3,963	373	8	0.02	0.22	0.42
1985	3,958	371	8	0.02	0.23	0.43
1986	4,338	463	9	0.02	0.19	0.36
1987	5,083	560	12	0.02	0.19	0.36
1988	5,802	618	12	0.02	0.18	0.34
1989	6,371	655	12	0.02	0.18	0.34
1990	6,971	813	23	0.03	0.25	0.47
1991	7,126	860	39	0.05	0.38	0.71
1992	7,606	959	66	0.07	0.55	1.03
1993	7,574	838	64	0.08	0.69	1.31
1994	8,548	928	71	0.08	0.71	1.34

Table 6.2b Intra-German goods and services trade, 1984-94[a]

Year	Trade totals (billions of dollars)			Trade ratios		
	World Trade[b]	Total German[d]	Intra-German[e]	Intra-regional	Simple intensity	Corrected intensity
1984	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
1985	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
1986	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
1987	6,237	683	19	0.03	0.26	0.49
1988	7,093	745	20	0.03	0.25	0.48
1989	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
1990	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
1991	9,292	1,272	398	0.31	2.28	4.25
1992	9,866	1,443	508	0.35	2.41	4.46
1993	10,054	1,319	484	0.37	2.80	5.23
1994	11,238	1,448	531	0.37	2.85	5.33

n.a. = not available

a. Table 6.2a is likely to be more accurate than 6.2b because intra-German trade was estimated directly for the former and indirectly for the latter table. Goods and services estimates for 1991-94 were derived as West German trade with the world plus East German trade with the world (in real terms, multiplied by the cost of living index) minus unified German trade with the world.

b. Derived from world trade figures plus intra-German trade.

c. Figures divided by the deutsche mark/dollar exchange rate.

d. 1987 and 1988 figures from WTO, *International Trade Trends and Statistics 1995*.

e. 1987 and 1988 figures from *Bundesbank Monthly Report,* January 1990, 13-20.

Sources: WTO, *International Trade Trends and Statistics 1995* (for world trade); Bundesbank Annual Reports 1992-94, table 3 (for East German trade); *Bundesbank Monthly Report,* August 1995, 61, September 1995, 61, (for West German and unified German trade), January 1990, 13-20 (for intra-German estimates), March 1996, 66 (for cost of living index); *Statistiches Jahrbuch 1995*, 278 (for intra-German trade); IMF, *International Financial Statistics* (for deutsche mark/dollar exchange rates).

couldn't find at all). The effect is also somewhat greater than that of the dissolution of Czechoslovakia (a factor of 2.3) after the passage of three years. The large effect in the case of Germany reflects that the pre-1990 barriers consisted of far more than a simple political border and that the reunification constituted the restoration of the high level of integration that had existed until 45 years earlier. Thus, the German case, though suggestive, may be too special to answer definitively our questions regarding asymmetries between the erection of political barriers, in general, and their removal.

All these estimates, even the German example, suggest that, when compared with the much larger estimate for the Canadian federation, only a small fraction of the effect of a change in political status is felt within five years. One suspects it may take as long as a century to attain the full effect. This is an important lesson for the European Union: it should not expect political union to be followed within three to five years by anything like the full twentyfold increase in intraregional trade that the Canadian example suggests might eventually be possible.

Evidence from the Home-Country Bias to Trade

While the few available examples of intrafederation trade are fascinating, they leave one hungry for more information on the effect of national political boundaries. An alternative source of information lies in all countries' statistical tendency to undertake economic transactions domestically versus internationally. We saw in chapter 3 that countries are much less open than most people think: the ratio of trade to income is much smaller than it would be if economic transactions were in fact as easily undertaken with counterparties around the globe as with counterparties around the corner. Some of this difference is due to international trade barriers and to costs of doing business at a distance. But much of the difference remains even after taking these factors into account. What remains should be attributed to the apparently inevitable costs associated with doing business across national borders.

The intercept terms in our gravity model reflect countries' general openness. By inspecting them, we can compare openness across time or across countries.[8] To compare across countries, look at the intercepts in appendix tables B5.5a through d. They suggest, for example, that Chile and Singapore are unusually open countries, after adjusting for the gravity determinants. The estimates also suggest that India and Mexico had unusually low levels of openness in 1970 but have since moved closer to the US level. When one allows for the size of its population (not shown), the United States ranks as one of the most open countries.

8. To compare across time, look, for example, at the intercepts in tables 5.2 or 5.3, discussed in appendix C.

To put an interpretation on the absolute level of openness worldwide, however, requires a separate set of tests, as carried out by Wei (1996a). There are (at least) two bothersome new data problems in this approach that have not arisen in our international gravity equation, where we do not care about the absolute value of the constant term. First, the bilateral merchandise trade data and output data are not comparable because output includes services and bilateral data on trade in services are not comprehensively available. Wei thus subtracts output in the sectors of construction and other services from the national GNP data to make them comparable with the trade data. This means limiting the sample to 19 OECD countries; most other countries lack the necessary data. Second, we need some measure, comparable to the international distance data, of the average distance between buyer and seller for domestic transactions. Much of trade within Russia or Canada takes place over a longer distance than trade between European countries. Wei estimates the average distance of domestic transactions as one-quarter of the distance to the nearest neighbor.

He finds that the bias toward trading domestically is high and significant, though not as high as McCallum (1995) and Helliwell (1995) found for the Canadian federation. First, controlling only for country sizes and distance, to make the results comparable to the McCallum-Helliwell estimates of a twentyfold effect in the Canadian case, Wei finds that the home-country bias is an approximate factor of 10. When he holds constant for adjacency, common language, and remoteness, the home-country bias to trade falls sharply, to something like a factor of 2. This estimate is much closer to those that emerge from the historical episodes of federation breakups. It appears that part of the striking McCallum-Helliwell estimate may be attributable to the effects of adjacency and remoteness.[9] Still, the remaining intracountry bias is large and important.

Wei's second finding is that, over the sample period 1982-94, the downward trend in the home-country bias to trade is small in magnitude and significance. Globalization is not taking place as rapidly as many believe. Borders still matter a great deal.

The Importance of History

In this section, we pursue a longer-term historical perspective. There are a number of lessons for the present. Studies of the historical statistics

9. Perhaps as much as a factor of five [10/2 = 5]. This would apparently leave a factor of four to be attributed to intra-Canada bias [20/5 = 4]. But the only way to know would be to estimate the gravity model on provincial data, including variables for adjacency, remoteness, and common language. Adjacency refers to the fact that Canadian provinces share borders with each other somewhat more often than they do with American states. Remoteness refers to the fact that Canada is one of those countries that is located rather far from

illustrate the tendency for bilateral trade ties to change relatively slowly. Long-span historical studies usually use the simple measures, such as the intensity or concentration ratios, which have less demanding data requirements than the gravity equation.[10] Anderson and Norheim (1993a, 38), examining intraregional trade intensities from 1928 to 1990, find numbers above 1 and steadily rising for Europe, the Americas, and each of their halves (East versus West, and North versus South, respectively). Asia, Africa, and the Middle East each show even higher intraregional intensities, but without the clear long-term trends. Recall that an intensity measure above 1 shows intraregional concentration but cannot distinguish whether it is due to natural factors such as proximity or to policy factors such as preferential trade arrangements. The time-series comparison does bear on the latter question.

Colonial Links

Kleiman (1976) studies the lingering effects of the great colonial empires that were largely established in the 19th century. He finds that the bias in trade toward the colonial power around the time of independence (the early 1960s for many French and British colonies) was approximately a factor of 2 to 4 on average across all colonies and ex-colonies. The factor was 2 to 3 for trade between the United Kingdom and its colonies, somewhat higher for France, and higher still for Belgium, Italy, and Portugal (in 1960-62). He also finds that the bias, which he attributes to ''enforcement'' (imperialism), decays rapidly after independence: by about one-third for countries that have been independent four to six years and by about three-quarters for those that have been independent for two decades.

Examining British trade intensities going back to 1913 (and simple trade shares going back to the 1850s), Anderson and Norheim (1993) find that the concentration of trade with the countries that made up the British Empire peaked in 1938 and declined steadily after 1958, falling below the critical level of unity by 1984. The intensity of French trade with the countries that were French colonies peaked at a very high level in 1938 and declined steadily thereafter, although that figure remained substantially above 1 as recently as 1990. North America shows a gradual but fairly steady upward trend in its intraregional intensity throughout 1929-90. Latin America (excluding Mexico, which is classified as North American) shows an intensity during this period that has a greater average trend

the rest of the world, so that it is more dependent on trade with its immediate neighbor and with itself. The remoteness effect is examined later in this chapter.

10. Flandreau (1995), however, manages to estimate a simple gravity model for Europe for 1860-80. He finds that trade within the Latin Union and Scandinavian Union can be explained by proximity and need not be attributed to the effects of these early regional trading arrangements.

but suffered a major step backward between 1948 and 1958. The intensity of Japan's trade with Asia peaked at a very high level in 1938 and has been declining fairly steadily ever since. This is precisely the same result found by Petri (1993).

Historical estimates with the gravity equation show links and trends similar to the intensity coefficients. One approach is to add a separate variable, for example, representing the trade of Britain with the Commonwealth countries, but excluding trade among the latter. This differs from the standard specification of FTAs, where every member is presumed to be linked to every other member. The intent is to capture a special metropolitan relationship between the core country and its periphery. Linnemann (1966) found different coefficients for the two kinds of transactions for 1958-60 and even more so for the relationship between France and its former colonies. Wang and Winters (1991) and Hamilton and Winters (1992) find significant effects for British ex-colonial relationships (though not French) as late as 1984-86.

Blocs in the 1930s

The 1930s are generally considered to have exhibited regionalism at its worst. The Ottawa Agreements of 1932 strengthened existing tariff preferences among members of the British Commonwealth, which in effect increased tariffs on outsiders. From 1931 onward, there was also an identifiable sterling zone. The maintenance of fixed exchange rates among this group, in a world of increasingly variable exchange rates, may have helped promote trade within the group. The sterling bloc overlapped imperfectly with the Commonwealth.[11] Nevertheless, one can speak of a British bloc.

For another set of countries, those that stayed on the gold standard, trade may again have been promoted relative to the alternative. However, the widespread use of exchange controls and other trade barriers among these countries to protect the balance of payments obviously had the opposite effect. A bloc of Central and Southeastern European countries formed around Germany, particularly under the barter-like Schacht Agreements. An attempt to liberalize trade within a group of Benelux and Scandinavian countries was undermined by a split with regard to currency policies between the gold bloc and the silver bloc.

In the aftermath of the passage of the trade-decimating (though nondiscriminatory) Smoot-Hawley tariff in the United States, Secretary of State Cordell Hull negotiated liberalization agreements bilaterally under the 1934 Reciprocal Trade Agreements Act. Countries that pegged their currencies to the dollar were fewer and less important, however, than the

11. For example, Canada was in the Commonwealth but not the sterling zone, vice versa for three Scandinavian countries (Eichengreen and Irwin 1995).

membership of the gold and sterling blocs. The dollar peggers were chiefly Spanish-speaking countries that were subject to extensive American political influence.

Eichengreen and Irwin (1995, 1997) apply the gravity model to the interwar period, in part to show the importance of the long-lasting effects of past patterns of bilateral trade on current patterns. Eichengreen and Irwin (1995) find statistically significant intraregional bias for the British Commonwealth and Central European blocs in 1935 and 1938, which is as expected. Less expected is their finding that much of this effect is simply a continuation of significant bloc effects in 1928.(Intraregional trade in that year approximately doubles the level of 1930s trade with otherwise similar nonmember countries.) To be sure, the estimated effect of the British Commonwealth preferences does rise between 1928 and 1935, presumably as the result of the Ottawa Agreements.

According to results in Eichengreen and Frankel (1995), the Commonwealth bloc is slightly stronger than the German bloc in the 1930s, and the dollar bloc is not statistically significant. The only strong significant evidence of trade diversion is in the German bloc in 1938.

Eichengreen and Irwin test simultaneously for trade blocs and currency blocs in the 1930s with dummy variables. Their results point to a number of conclusions: worldwide trade was depressed in the 1930s not only by trade barriers but also by exchange rate variability, and trade was also increasingly regionalized in the 1930s by the sterling area, Commonwealth preferences, a Reichsmark bloc, a Benelux bloc, and a Western Hemisphere bloc. The authors choose to emphasize the evidence that these blocs were already beginning to affect trade flows in 1928, earlier than regionalization is conventionally dated.

Lags in Trade Patterns

We have noted the tendency for trading patterns to change relatively slowly, even when the change in regional trading arrangements or political links is sudden. Eichengreen and Irwin (1997) take the bull by the horns and include lagged values of bilateral trade in their gravity estimates for a sample of observations stretching from 1928 to 1965. They find, for example, that trade links among British colonies in 1954 and 1964 that might otherwise be attributed to Commonwealth preferences are in fact simply the lagged effects of trade flows of 1949, when the countries belonged to the British Empire. They use the term hysteresis to refer to effects that seem to linger long after the original reasons for the bilateral trade have vanished. They conclude that one should always include lagged variables in the gravity equation.

A possible alternative interpretation is that one should look out for missing variables. These variables are likely to change slowly over time

and so can produce significant coefficients on lagged trade. One possible missing variable that addresses the hysteresis issue directly is the accumulated stock of foreign direct investment (FDI). If firms of a mother country develop a large stock of capital in a colonial dependency, this investment might manifest itself in a heightened level of trade between the two countries long after the colonial relationship has been severed. By including bilateral FDI as an explanatory variable, which we do below, we hope to capture the effect directly. The measured stock of FDI may omit such intangibles as brand loyalty among consumers in the importing country or familiarity with local languages and customs in the exporting country. But bilateral FDI should capture much of the essence of this relationship.

The Effect of Cumulative Foreign Direct Investment

If a firm inherits a stock of retail outlets in another country, it is likely to export more to that country, other things being equal. Or if the firm inherits a stock of factories, plantations, or mines in another country, it is likely to import more from that country. Furthermore, subsidiaries are more likely than local firms to turn to the mother country for intermediate inputs. There is also the tendency that when an ''imperialist'' power builds the infrastructure in a country of the periphery, whether it is the railroad system in colonial Africa or the port system in modern Southeast Asia, it will do so in such a way as to facilitate trade links with itself. Thus, for all these reasons, one expects a positive association between FDI and trade.

Theoretically, there is also an effect that can go the other way. If high barriers to trade, such as tariffs or quotas, prevent firms from exporting into the protected market, they may undertake direct investment as a substitute, jumping the barriers by building factories to supply the protected market directly. The view that FDI and trade are substitutes rather than complements suggests that where a high stock of FDI is inherited from the past, for whatever reason, trade is likely to be lower rather than higher. Empirically, however, most studies have found that the positive association between FDI and trade dominates. They tend to be complements, not substitutes.

The proponents of the yen bloc hypothesis tend particularly to emphasize the role of FDI as a tool that Japan uses in East Asia to redirect the flow of economic transactions to itself. They point to the rapid growth of Japanese FDI in East Asia in the 1980s. These arguments often neglect the fact that the increase in Japanese FDI into East Asia was proportional to the increase in trade, which we have in turn described as natural in light of the rapid growth of the economies involved. Moreover, Japan's FDI into East Asia is smaller than its FDI into the United States and

Canada, which is more than twice what one would expect from North America's share of world trade. This pattern, evident in the official Japanese Ministry of Finance figures on the destination of FDI, applies even more strongly if one takes into account that the FDI statistics represent investment either approved by or reported to the government and thus greatly overstate the extent of actual Japanese investment in developing countries. More accurate balance of payments data from the Bank of Japan show a smaller percentage of investment going to Asia (Ramstetter 1991a, 8-9; 1991b, 95-96; Frankel 1993, 67-69).

Nevertheless, it is well worth exploring the possibility that past FDI influences the trade flows we have been analyzing. More formally, we wish to test the hypothesis that the stock of FDI that country *i* has placed in country *j* is an important determinant of exports from *i* to *j*. Thus we try adding cumulative bilateral FDI as a factor in the gravity model of bilateral trade.

Bilateral FDI data are only available on a limited basis. Many previous studies have looked at bilateral FDI between the United States and other countries, or between Japan and other countries.[12] But there are few worldwide estimates. We have for this purpose put together a relatively comprehensive and recent bilateral FDI data set.

There are serious problems of endogeneity: all the factors that affect bilateral trade could also be expected to affect bilateral FDI. The simplest way to try to address this problem is to use the lagged stock of FDI as an explanatory variable. Results for 1992 trade, determined as a function of the 1990 stock of FDI, are reported in table 6.3.[13] Bilateral FDI has an apparent high and very significant effect on bilateral trade. The point estimate is .17. In other words, a 1 percent increase in the stock of FDI leads to a subsequent .17 percent increase in exports.[14] The coefficients on distance, adjacency, and language decline slightly, but the gravity results are basically unaffected. The estimated coefficients for all bloc effects decline slightly, suggesting that the bloc effects earlier might have been appropriating a bit of the FDI effect. The bloc effects for the Asia Pacific Economic Cooperation forum (APEC), the Western Hemisphere (when holding constant for openness), and the European Community (when not holding constant for openness) are still significant.

---

12. Sachs and Shatz (1994, 45-49) show in a gravity model of US bilateral exports and imports that the bilateral activities of multinational corporations are a significant determinant.

13. Taken from table 4 of Wei and Frankel (1997). This test necessitates introducing the distinction between the importing country and the exporting country, which is discussed at length later in the chapter. The FDI tests are extended to 1990 and 1994, with a special emphasis on Southeast Asia, in Frankel & Wei (1996).

14. An extension of these tests finds a somewhat smaller effect in 1990 and a somewhat larger effect in 1994 (Frankel and Wei 1996). The effects are significant in all three years.

Table 6.3 Do exports follow FDI? (dependent variable: 1992 exports from country *i* to *j*)[a]

	Coefficient	With openness	With FDI	With FDI & openness
Intercept	−15.565**	−17.368**	−9.415**	−10.689**
	(1.356)	(1.612)	(1.552)	(1.617)
GNP_i	0.734**	0.747**	0.669**	0.6892**
	(0.043)	(0.042)	(0.039)	(0.038)
GNP_j	0.591**	0.604**	0.503**	0.513**
	(0.032)	(0.034)	(0.032)	(0.030)
$GNP/capita_i$	−0.191#	−0.119	−0.522**	−0.530**
	(0.105)	(0.103)	(0.117)	(0.110)
$GNP/capita_j$	0.096**	0.105**	0.074*	0.083*
	(0.036)	(0.034)	(0.035)	(0.033)
Distance	−0.591**	−0.556**	−0.525**	−0.478**
	(0.053)	(0.068)	(0.050)	(0.063)
Adjacency	0.500**	0.477**	0.421**	0.341**
	(0.125)	(0.129)	(0.115)	(0.116)
Common language	0.247*	0.355**	0.134	0.265**
	(0.110)	(0.108)	(0.097)	(0.090)
Stock of FDI $(1990)_{ij}$			0.142**	0.169**
			(0.021)	(0.019)
East Asia bloc	0.156	−0.029	0.194	−0.046
	(0.197)	(0.219)	(0.179)	(0.198)
APEC bloc	0.997**	1.059**	0.863**	0.783**
	(0.114)	(0.150)	(0.109)	(0.138)
Western Hemisphere bloc	0.246	0.711**	0.165	0.630**
	(0.157)	(0.196)	(0.147)	(0.178)
EEC bloc	0.323**	0.270**	0.231*	0.156
	(0.114)	(0.113)	(0.109)	(0.108)
East Asia openness		0.315*		0.397*
		(0.122)		(0.110)
APEC openness		−0.067		−0.127
		(0.169)		(0.148)
Western Hemisphere openness		−0.380**		−0.477**
		(0.116)		(0.107)
EEC openness		0.155		0.003
		(0.126)		(0.113)
Number of observations	347	347	347	347
Standard error of regression	0.677	0.638	0.639	0.581
Adjusted R^2	0.793	0.817	0.816	0.847

a. All variables except dummies are in natural logarithms. **, *, and # denote significance at the 99%, 95%, and 90% levels, respectively. Standard errors are heteroskedacity-consistent. Region 2 takes the value of one if both countries (*i* and *j*) in the pair are in the region. Region 1 takes the value of one if the pair includes a country in the region. Standard errors are in parentheses.

Source: Wei and Frankel (1997).

Simply lagging the stock of FDI is not, however, a truly satisfactory way of addressing endogeneity. Even in a time-series context, precedence does not ensure causality. Firms could undertake direct investment in rational anticipation of an expanding market for trade. In a cross-section context, the simultaneity problem is even worse. Trade could be correlated with FDI because trade causes investment rather than the other way around. Multinational firms acquire retail outlets abroad to help sell consumer goods, for example, or acquire mines in order to import minerals. How can potentially bidirectional causality be addressed? We have obtained what seems to be a good instrumental variable for bilateral FDI: a dummy variable indicating the existence of a bilateral tax treaty. This variable is a significant determinant of bilateral FDI, as a (''first-stage'') regression confirms. Thus it is a good candidate for an instrumental variable. When the tax treaty variable is used to isolate exogenous variation in the stock of FDI, we find that the effect of FDI on trade vanishes. This finding suggests that, even though bilateral FDI is highly correlated with bilateral trade, there is in fact no direct causal relationship. (An alternative possibility is that the endogeneity problem is even worse for the tax treaty variable than for FDI itself.)

One can estimate a gravity model of the determination of FDI analogous to the standard one for trade.[15] Results presented in Frankel and Wei (1996) are a start. They omit many of the possible determinants that have been identified in the literature on FDI. Nevertheless, the gravity framework has its attractions, notably a much larger number of observations than in the typical study of FDI.

To the extent that the motive for FDI is to sell into the local market, one might expect distance and transport costs to have, if anything, a positive effect on FDI, thus reversing a key plank of the gravity model. On the other hand, to the extent that the motive is exporting back to the source country, distance should have a negative effect, just as it does for trade. The same is true if distance matters because it breeds unfamiliarity with local culture.[16]

We find that the coefficient on distance is even more significant and negative than is the case in the gravity model of trade. Similarly, the coefficient on language (which also includes former colonial links) is extremely high and significant. In 1992, the existence of linguistic links raised the stock of FDI about ninefold [exp(2.24) = 9.4]. In this light, it is not surprising that adding FDI into the trade equation, in the results

15. Eaton and Tamura (1994, 1996) estimate bilateral gravity models for FDI. But they include only two source countries: the United States and Japan. They find that features of a country that are associated with more trade with the United States or Japan are also associated with more FDI from those countries.

16. Eaton and Tamura (1996) find in their gravity model that distance inhibits FDI much less than it inhibits trade.

reported above, deprived the language variable of its statistical significance. The coefficients on GDP are also highly significant.[17] To sum up the results on the determination of bilateral FDI, they seem to be similar to the determinants of bilateral trade, which explains the effect that lagged FDI had in the earlier trade equations in the preceding section.

A full exploration of the determinants of FDI, though relevant to the study of regionalization, is outside the scope of this book. Many of the existing regional economic arrangements include liberalized rules for FDI. A full model of the simultaneous determination of bilateral trade and investment would raise many interesting questions but will have to wait for future extensions.[18]

Rich and Poor Countries

It has been a property of our gravity analysis that rich countries like to trade more with other rich countries, rather than with poor countries. This would surprise trade theorists of the past, as international trade has always been thought to be motivated by countries' differences. According to the Heckscher-Ohlin theory of comparative advantage, poor countries produce goods intensive in unskilled labor and trade them to rich countries for goods intensive in capital and skilled labor.

The Linder Effect

We tried including bilateral absolute differences in GNP per capita figures. This specification is a way to test for Linder or Heckscher-Ohlin effects without having to disaggregate exports from imports. The question is whether trade flows are large among similar countries or dissimilar countries. The variable did not have the positive effect that one would expect if countries traded capital-intensive products for unskilled-labor-intensive products. Rather, it had a moderately significant negative effect, as the Linder hypothesis, that similar countries trade more than dissimilar ones, would predict. This undercuts support for Heckscher-Ohlin. Meanwhile, the positive coefficient on the product of per capita GNPs remains, as before, very highly significant, suggesting that the most powerful effect of the three is the relationship between development and trade (Frankel, Stein, and Wei 1995, table 3).

---

17. Of course, there is probably a bad misspecification in this equation in this regard. We have not yet included terms for GDP per capita, which would capture the fact that rich countries tend to be the source of FDI.

18. Wei (1996b) brings measures of tax rates, corruption, and wage rates into the study of bilateral FDI determination in this data set.

Factor Endowment Differentials

We also tried to better capture classic Heckscher-Ohlin effects. We estimated a gravity equation that included more direct measures of factor endowments: the two countries' differences in capital/labor ratios, educational attainment levels, and land/labor ratios. The data (for a subset of 656 of our 1,953 pairs of countries) were generously supplied by Gary Saxonhouse (1993). The results are not quite as bleak for Heckscher-Ohlin as was the negative coefficient on income differentials. There is a bit of support for some of these terms, specifically for capital/labor ratios and educational attainment in 1980. In the other years, however, the results are poor. The coefficients on the blocs and other variables are, in any case, little affected (Frankel, Stein, and Wei 1995, table 4; Frankel and Wei 1993b, table 5).

The Endogeneity of Incomes

We have treated both economic size and income per capita as exogenous variables in the gravity equation. Yet there are good reasons to suspect that they are endogenous, influenced by the level of trade. Economic theories, from the basic principle of comparative advantage to modern theories of technology transfer through trade, have long said that countries that are more open to trade will have higher or more rapidly increasing levels of real income. There is good empirical support for this effect as well.[19]

But trade is the dependent variable in our gravity equation. Thus the apparently significant effect of income on trade in our equation could be spurious, a correlation that is in truth attributable to the influence of trade on income. Not only would estimation of the income coefficient be biased in the event of such a simultaneity problem, but so would the estimation of the other coefficients. If, for example, FTAs are more likely to be successful among rich countries than poor ones, improper measurement of the effects of income on trade could also lead to improper measurement of the effects of FTAs.

While many variables determine countries' income levels, the most important quantifiable determinants are their levels of factor accumulation: labor force, stock of physical capital, and stock of human capital

19. Frankel and Romer (1996) and Frankel, Romer, and Cyrus (1995) give extensive references to the empirical literature. Those two papers also show that the statistical effect of openness on growth holds up well, even when allowing for the possible endogeneity of trade. Thus, while we are about to use exogenous variables from the growth equation (such as investment) to correct for simultaneity bias in the trade equation, the two earlier papers use exogenous variables from the trade equation (such as distance) to correct for simultaneity bias in the growth equation.

(education and technological know-how). It is plausible that these determinants of income are exogenous and thus good instrumental variables.[20] Table 6.4 reports along these lines the results of instrumental variables estimation of the gravity equation.[21] We follow the standard neoclassical growth model in listing the three factor accumulation variables as determinants of income in column IV1, and we follow the ''conditional convergence'' literature in adding 1960 income to the list in column IV2. The use of the instrumental variables technique sometimes reduces the coefficient on income or income per capita very slightly, but the differences are not statistically significant. The other coefficients are usually little changed as well, whether on other gravity variables, narrowly defined regional trading arrangements, or more broadly defined blocs. Evidently, the endogeneity of income makes little difference.

The Role of Currency Links

In addition to the strategy of enacting regional trading arrangements, governments sometimes seek to promote trade among a group of countries by first linking their currencies. The link might be relatively loose, as in the Exchange Rate Mechanism of the European Monetary System, or somewhat tighter, as in the fixed exchange rates that some Latin American countries have established against the dollar, or very tight, as in the common currency on which the members of the European Union have set their sights. The idea is that currency links reduce the costs of doing business with the partner country; these costs take the form of foreign exchange risk (including the cost of hedging such risk) and transaction costs. The reduction in costs should promote trade. This has certainly been a major motivation for proponents of Economic and Monetary Union (EMU) in Europe.

Some European leaders believe that currency links are not just a desirable supplement to a successful common market, but are actually a necessary component of it. Others have also read the Latin American experience as suggesting that exchange rate stability is a prerequisite for successful regional trading arrangements: extreme instability helped torpedo some of the preferential trade arrangements of the 1970s, and large swings in the real exchange rate between the Argentine peso and the Brazilian *cruzeiro* have recently put strains on Mercosur (e.g., Abreu and Bevilaqua 1995). The Mexican peso crisis of December 1994 (rightly or wrongly)

20. Wei (1996a) also allows for the endogeneity of income in the gravity equation, using simple population as an instrumental variable. Harrigan (1991) uses a more complete set of factor endowments as instrumental variables for income. Unfortunately, he omits any measure of distance from his trade calculations, which must have a major effect on the results.

21. Drawn from Cyrus (1996). Results for 1980 and 1985 are reported there as well.

Table 6.4 Correcting for simultaneous determination of income in the gravity equation, 1990 (dependent variable is the log of exports from country *i* to *j*)

	With five bloc terms			With European bloc terms only		
Variable	**OLS**	**IV1**	**IV2**	**OLS**	**IV1**	**IV2**
Constant	−26.087**	−26.045**	−25.534**	−24.596**	−26.485**	−23.691**
	(0.815)	(1.120)	(0.879)	(0.651)	(0.770)	(0.678)
log(GNP_i)	0.837**	0.839**	0.844**	0.873**	0.895**	0.873**
	(0.0222)	(0.0235)	(0.0225)	(0.0230)	(0.0240)	(0.0233)
log(GNP_j)	0.836**	0.834**	0.820**	0.872**	0.890**	0.852**
	(0.0219)	(0.0233)	(0.0222)	(0.0228)	(0.0237)	(0.0230)
log(GNP_i/pop_i)	0.370**	0.372**	0.366**	0.349**	0.442**	0.323**
	(0.0317)	(0.0448)	(0.0345)	(0.0295)	(0.0361)	(0.0311)
log(GNP_j/pop_j)	0.288**	0.282**	0.248**	0.263**	0.351**	0.203**
	(0.0315)	(0.0445)	(0.0342)	(0.0293)	(0.0359)	(0.0308)
log(Distance)	−0.725**	−0.726**	−0.734**	−0.966**	−0.991**	−0.954**
	(0.0579)	(0.0591)	(0.0582)	(0.0455)	(0.0460)	(0.0456)
Adjacency	0.607**	0.607**	0.602**	0.653**	0.624**	0.668**
	(0.167)	(0.167)	(0.167)	(0.172)	(0.173)	(0.172)
EEC bloc	−0.372*	−0.368†	−0.317†	−0.773**	−1.013**	−0.658**
	(0.174)	(0.191)	(0.177)	(0.177)	(0.185)	(0.179)
EFTA bloc	0.129	0.133	0.192	−0.175	−0.427	−0.0561
	(0.289)	(0.303)	(0.292)	(0.301)	(0.308)	(0.303)
Western Hemisphere bloc	1.424**	1.423**	1.421**			
	(0.147)	(0.147)	(0.147)			
East Asia bloc	0.573*	0.569*	0.518†			
	(0.271)	(0.281)	(0.273)			
APEC bloc	1.165**	1.170**	1.230**			
	(0.188)	(0.209)	(0.192)			
EEC openness	0.420**	0.425**	0.485**	0.0495	−0.154†	0.147†
	(0.0960)	(0.131)	(0.103)	(0.0790)	(0.0907)	(0.0816)
EFTA openness	−0.536**	−0.531**	−0.465**	−0.840**	−1.074**	−0.727**
	(0.112)	(0.149)	(0.119)	(0.0975)	(0.110)	(0.100)
Western Hemisphere openness	−0.395**	−0.393**	−0.375**			
	(0.0811)	(0.0854)	(0.0820)			
East Asia openness	0.940**	0.936**	0.894**			
	(0.112)	(0.128)	(0.115)			
APEC openness	−0.411**	−0.406*	−0.337**			
	(0.121)	(0.159)	(0.128)			
Number of observations	2,597	2,597	2,597	2,597	2,597	2,597
Standard error of regression	1.447	1.447	1.448	1.540	1.548	1.543
R^2	0.762	0.762	0.762	0.730	0.729	0.729
Adjusted R^2	0.761	0.761	0.760	0.729	0.728	0.728

** = 99%; * = 95%; † = 90%
OLS = ordinary least squares; IV = instrumental variable

Source: Cyrus (1996).

soured some Americans and Mexicans alike on the North American Free Trade Agreement (NAFTA) and its expansion. On the other hand, exchange rate stability does not seem to have been necessary to Canadian-US trade links. The exchange rate between the Canadian and US dollars has floated more cleanly and over a higher fraction of the postwar period than any other bilateral exchange rate in the world.

We can readily investigate the extent to which the bilateral stabilization of exchange rates within major regional groups has contributed to the regionalization of trade. The question can be decomposed into two: to what extent have bilateral exchange rates in fact been stabilized within the major regional groups, and to what extent does reduced bilateral exchange rate variability promote trade?

As regards the first question, the unsurprising finding is that the values of currencies of European countries are tied far more closely to the value of the mark than to other major currencies, and the values of Western Hemisphere currencies tend to be tied to the value of the dollar. The finding that surprises some people, however, is that out of nine currencies in East Asia, only a few show a significant or growing link to the yen, and all show that links to the dollar are much stronger than to the yen. Although only Hong Kong pegs formally to the dollar, no currency pegs formally to the yen. Those that have at times adopted loose basket pegs (Malaysia, Singapore, and Thailand) give far more weight in the basket to the dollar (about .8) than to the yen (about .1). Each of the Asian countries is more properly classed in a dollar bloc than in a yen bloc. It thus appears that there are not three currency blocs in the world, but two: a mark bloc in Europe and a dollar bloc in the Pacific.

The next question is whether a policy of stabilizing bilateral exchange rates among a set of countries will help promote trade among the partners in question. There have been quite a few time-series studies of the effect of exchange rate uncertainty on trade overall but fewer cross-section studies of bilateral trade. Three exceptions are Thursby and Thursby (1987), De Grauwe (1988)—which look only at a group of industrialized countries—and Brada and Mendez (1988). We have reexamined the question using our broader data set, covering 63 countries. We add an additional variable to our gravity model of bilateral trade to capture the effect of exchange rate variability alongside the other variables. A problem of simultaneous causality should be noted at the outset: if exchange rate variability shows up with an apparent negative effect on the volume of bilateral trade, the correlation could be as easily due to the government's deliberate efforts to stabilize the currency vis-a-vis a valued trading partner as to the effects of stabilization on trade.

Volatility is defined to be the standard deviation of the first difference of the logarithmic exchange rate. We started with the volatility of nominal exchange rates and embedded this term in our gravity equation. We also

tried the volatility of real exchange rates. Most coefficients are similar to those we have encountered in the estimates without exchange rate variability: the Western Hemisphere, East Asia, APEC, and the European Community all show statistically significant bloc effects.

The ordinary least squares (OLS) results show a negative effect of exchange rate volatility (whether nominal or real) on bilateral trade that is highly significant in 1965, 1970, and 1975, as well as 1980. Only in 1985 and 1990 does the negative effect disappear (indeed, it turns positive).

By way of illustration, these point estimates can be used for some sample calculations. They suggest that if the level of EC real exchange rate variability that prevailed in 1980, a standard deviation of 2 percent, had been eliminated altogether, the volume of intra-EC trade would have increased by 14.2 percent. This OLS estimate should be regarded very much as an upper bound. For one thing, the 1980 point estimate of the effect of exchange rate volatility is the largest of all the years. In the earlier observations, the magnitude of the estimated effect is one-fifth to one-half the size.

Interpretations of the OLS estimates are in any case complicated by the likelihood of simultaneity bias in the regressions. Governments may choose deliberately to stabilize bilateral exchange rates with their major trading partners. This has certainly been the case in Europe. Hence, there could be a strong observed correlation between trade patterns and currency links even if exchange rate volatility does not depress trade.

To address this problem, we use the method of instrumental variable estimation, with the standard deviation of relative money supply as our instrument for the volatility of exchange rates. The argument in favor of this choice of instrument is as follows. Relative money supplies and bilateral exchange rates are highly correlated in theory (they are directly linked under the monetary theory of exchange rate determination) and in our data as well, but monetary policies are less likely than exchange rate policies to be set in response to bilateral trade patterns.

The instrumental variable results show the same sign pattern across the years as the OLS estimates, but the negative effect is statistically significant only in 1965. The coefficient for 1980 is a completely insignificant 0.28. Even if the point estimate is taken at face value, it would imply that the elimination of exchange rate variability worldwide would increase trade by only 0.9 percent (0.28×3.22). These weak results when correcting for simultaneity are the second strike against the hypothesis that the stabilization of exchange rates within the European Union or other groups has been an important factor in promoting intraregional trade.[22]

22. The estimates are taken from Frankel and Wei (1997, tables 4, 5, and 6). Similar results are reported in the earlier Frankel and Wei papers.

To summarize, these results are generally consistent with the hypothesis that real exchange rate volatility depressed bilateral trade a bit in the 1960s and 1970s. But the evidence for a negative trade effect, which starts out relatively strong in 1965, diminishes steadily in the 1970s and 1980s, especially if one takes due account of simultaneity. Even as exchange rates became more volatile, the effect of any given level of volatility fell. Presumably, importers and exporters learned to cope with uncertainty. The proliferation of currency options, forward contracts, and other hedging instruments over this period may explain why the effect, which appears once to have been there, has more recently disappeared.

In any case, the exchange rate variability term does little to reduce the strength of the trade bloc effects among such relevant groups as the European Community, the Americas, or APEC.

Bilateral Exports and Imports: The Roles of Relative Prices and Relative Distance

In the interest of ensuring that the results from our gravity model are as robust as possible, we have considered a number of further extensions. Those in this section concern the distinction between the importing country and the exporting country.

Disaggregating Imports from Exports

We have focused primarily on estimation results in which imports and exports are aggregated. Our goal has been to uncover the deep underlying structure of bilateral trade patterns. We think that exports plus imports (say, divided by 2, to arrive at the average) are a good measure of this underlying trade structure. It is common in gravity models to treat imports and exports separately, however, without constraints that force the coefficients on domestic and foreign income, or income per capita, to be equal. Most studies of the model in this form find income coefficients that are only slightly different on the import side and the export side. Linnemann, for example, found that the estimated income elasticities were different for the two, but that constraining them to be equal despite this produced little change in the results. If one concludes that the coefficients are equal, then it is legitimate to impose this constraint. Coefficients on distance and the other bilateral variables are necessarily equal for imports and exports.[23]

23. For a given country, proximity or common language could conceivably have a bigger effect on imports than exports. When we use distance to measure transport costs, for example, we abstract from the fact that sailing time or flying time from east to west may be different than from west to east. With our general formulation, however, where the variable to be explained is simply T_{ij}, no such distinctions are possible. It seems far-fetched, in any case, that this could matter.

Thus it is in turn legitimate to add imports and exports together, to simplify the estimation and the presentation, as we have done.[24]

A further motive for adding imports and exports together, aside from simplicity, is that to treat them separately requires that one enter the realm of macroeconomics. As soon as one considers imports and exports separately, one has to recognize that their difference is the trade balance. The trade balance is not determined by the deep underlying structure of trade. Rather, it is the difference between saving and investment and so is determined by such macroeconomic factors as the real exchange rate, business cycles, government spending, investment opportunities, and intertemporal optimization by savers.

Exploring the realm of macroeconomics has many rewards. Robust econometric results do not tend to be among these rewards. Statistical estimates often change wildly every time one extends the sample period, modifies the specification slightly, or applies a different technique.[25] Until this point in our study, the hope has been that when a country undergoes a real appreciation, the positive effect on its imports and the negative effect on its exports approximately cancel out. This would justify exports and imports aggregating into a single equation and omitting a term for the real exchange rate.

These arguments notwithstanding, if the income coefficients were in truth different for the importing country and the exporting country, then it would be no good to pretend otherwise. Aggregating the two equations together would not be legitimate. As a further check on the robustness of our results, we have now tried the estimation on disaggregated imports and exports.

We begin by relaxing the constraint that the income coefficients are equal in the importing and exporting country, but without the real exchange rate term. The results are reported in tables 6.5a and B6.5a for 1965-75 and tables 6.5b and B6.5b for 1980-90 (table 6.3 reported the estimates for 1992). In table B6.5a we use the Summers-Heston measure of real GDP. As others have found, the income coefficient is a little greater for the importing country than the exporting country. This coefficient is consistent with the old-fashioned Keynesian idea of a demand-determined

24. It is, moreover, legitimate to take logs. We have thus simply added log(2) to the righthand side of the equation that describes the log of exports and the log of imports.

25. Without thinking of themselves as having entered the realm of macroeconomics, a few authors have sought to estimate relative price terms in their bilateral trade equation on the grounds that these terms appear in a gravity-type equation derived from a general equilibrium theory. In fact, however, relative price variation is in practice so heavily dominated by exchange rate variation that there is really no way of escaping that this is a macroeconomic phenomenon. This is especially true because conventional indices measure prices only relative to a base year, preventing comparison on a pure cross-sectional dimension. In any case, these studies tend to get unsatisfactory results for the relative price terms.

Table 6.5a Gravity model estimations (with bloc effects) with exports and imports estimated separately, 1965-75
(dependent variable is the log of country *i*'s exports to *j*)[a]

Variable	1965	1970	1975
GDP_i	0.798**	0.779**	0.826**
	(0.023)	(0.023)	(0.022)
GDP_j	0.736**	0.795**	0.844**
	(0.023)	(0.026)	(0.024)
$GDP/capita_i$	0.603**	0.799**	0.787**
	(0.038)	(0.038)	(0.036)
$GDP/capita_j$	0.539**	0.647**	0.565**
	(0.041)	(0.044)	(0.039)
Distance	−0.743**	−0.831**	−0.922**
	(0.037)	(0.037)	(0.035)
Adjacency	0.179	0.260	0.202
	(0.134)	(0.165)	(0.147)
Common language	0.482**	0.706**	0.567**
	(0.088)	(0.089)	(0.092)
East Asia bloc	0.872*	1.424**	0.912**
	(0.406)	(0.264)	(0.218)
APEC bloc	0.308*	0.658**	0.936**
	(0.150)	(0.131)	(0.124)
Western Hemisphere bloc	−0.346**	−0.436**	−0.210#
	(0.118)	(0.117)	(0.116)
EC bloc	−0.195*	−0.346**	−0.264**
	(0.092)	(0.095)	(0.086)
Number of observations	1,953	2,354	2,767
Standard error of regression	1.267	1.497	1.541
Adjusted R^2	0.688	0.690	0.687
Log likelihood	−3228	−4284	−5116

** = significance at 99% level; * = 95%; # = 90%.

a. The GDP per capita variables are from the Summers-Heston data set.

marginal propensity to import, expressed in elasticity form. But the difference between the import elasticity and export elasticity is small and often insignificant statistically. The difference in coefficients on GNP per capita is larger and more significant, with the exporting country always having the higher value. This might suggest that, of the various explanations for the role of per capita income in determining trade, the tendency for poorer countries to engage in more import protection does not seem to be the dominant one. One must always recall, however, that averaged over the very long run, a country's exports equal its imports.

The magnitudes of the estimated coefficients on per capita incomes are higher than in the earlier estimates. The same is true of the coefficient on distance. It still has a trend that, if anything, rises rather than falls. This is particularly true in the 1970s. (Over 1980-90 it is fairly steady at .9.) The estimated coefficient on the adjacency variable now has a strong

Table 6.5b Gravity model estimations (with bloc effects) with exports and imports estimated separately, 1980-90
(dependent variable is the log of country *i*'s exports to *j*)

Variable	1980	1985	1987	1990
GDP_i	0.809**	0.827**	0.822**	0.862**
	(0.022)	(0.022)	(0.022)	(0.023)
GDP_j	0.841**	0.870**	0.842**	0.866**
	(0.024)	(0.023)	(0.021)	(0.023)
$GDP/capita_i$	0.834**	0.931**	0.999**	0.995**
	(0.037)	(0.037)	(0.036)	(0.039)
$GDP/capita_j$	0.726**	0.714**	0.832**	0.799**
	(0.041)	(0.041)	(0.043)	(0.039)
Distance	−0.907**	−0.859**	−0.896**	−0.935**
	(0.037)	(0.036)	(0.036)	(0.038)
Adjacency	0.301*	0.512**	0.443**	0.450**
	(0.146)	(0.145)	(0.136)	(0.130)
Common language	0.471**	0.516**	0.532**	0.456**
	(0.099)	(0.092)	(0.090)	(0.100)
East Asia bloc	0.818**	0.747**	0.943**	0.618**
	(0.207)	(0.189)	(0.180)	(0.206)
APEC bloc	1.098**	0.958**	0.933**	0.853**
	(0.121)	(0.108)	(0.105)	(0.112)
Western Hemisphere bloc	−0.017**	−0.109**	0.004	0.282*
	(0.117)	(0.128)	(0.119)	(0.112)
EC bloc	−0.088	0.064	0.138#	0.064
	(0.080)	(0.083)	(0.085)	(0.089)
Number of observations	2,824	2,823	3,109	3,070
Standard error of regression	1.605	1.602	1.653	1.667
Adjusted R^2	0.679	0.682	0.688	0.695
Log likelihood	−5337	−5330	−5968	−5920

** = significance at 99%; * = 95%; # = 90%.

upward trend from year to year. Evidently, the costs of doing business with close neighbors have been falling more rapidly than the costs of doing business with distant ones.

What are the effects on the bloc variables? As before, the strongest effects belong to East Asia and APEC. As before, there is no evidence of these two bloc effects increasing over time.

Next we added the real exchange rate, the relative price of *i*'s goods in terms of *j*'s (The source is the Summers-Heston measure of the real exchange rate.) The results are not reported here. The coefficient on the real exchange rate fluctuates in sign and is usually insignificant statistically. This is similar to what others have found. Boisso and Ferrantino (1993) also had little success with a relative price term. Bergstrand (1985, 1989) emphasized the importance of prices in theory, but the estimated coefficients were again usually not significant statistically. We will not concern ourselves with the real exchange rate term further.

Remoteness

In this section, we estimate a modified version of the gravity model. To explain bilateral trade between a country and a specific trading partner, we incorporate the distance of each country from its average trading partners, which we call remoteness, in addition to the direct bilateral distance. This extension of the basic gravity formulation is based in part on a new formulation in Deardorff (1997).[26]

The remoteness variable measures how far an exporting (or importing) country is from all other countries. It is a measure of ''overall distance.'' An exporter's remoteness is defined as its average distance from its trading partners, using partners' GNPs as the weights. An importer's remoteness is defined analogously. The hypothesis is that the remoteness of an exporter from the rest of the world has a *positive* effect on bilateral trade volume in the equation (conditional on bilateral distances).

An example will illustrate the intuition. The distance between Australia and New Zealand is the same as the distance between Spain and Poland. Spain and Poland have lots of other natural trading partners close at hand, but Australia and New Zealand do not. One might thus expect the antipodean pair, who have less in the way of alternatives, to trade more with each other, holding other influences constant, than the European pair. The idea is that it is not just the absolute level of bilateral distance that matters, but also bilateral distance expressed *relative* to the distances of each of the pair from their *other* partners. As a check for robustness, we have also computed the remoteness of exporters and importers as the equally weighted distance from their trade partners, with little effect on the results (Frankel and Wei 1997, table 2).

Results are reported in appendix table B6.6. The coefficient on the exporter's remoteness is always positive and is statistically significant three out of four times over 1970-92. Other things being equal, if country Z is further from the rest of the world than country S by 1 percent, then Z's exports to a common third country, say A, will be higher than that of country S by 0.3 to 0.6 percent. Another way of stating this result is to break down the negative coefficient on bilateral distance (.9) into two roughly equal effects: an effect of bilateral distance *relative* to the average distance of the exporter (.3 to .6) plus an *absolute* distance effect (.6 to .3).

The coefficient on the importer's remoteness is consistently negative, surprisingly, and is statistically significant at the 5 percent level. Apparently, if Z's average distance from the world is greater than S's by 1 percent, then Z's imports from M, other things equal, are less than S's by an estimated 0.4 to 0.8 percent. We have not figured out why this might be.

---

26. Also based on suggestions from Jacques Polak (1996) and Ed Leamer and on a theoretical formulation in Stein (1995).

In appendix table B6.6, we confine ourselves to testing formal FTAs. The bloc variable is defined the same as in earlier tables. (The dummy variable labeled ''nonbloc,'' however, is defined as trade between a member of the group and a nonmember.[27]) The regional arrangements tested were the European Community, the European Free Trade Association (EFTA), the Canada-US FTA, Mercosur, the Andean Group, and the Association of Southeast Asian Nations (ASEAN). Year-by-year point estimates of the EC bloc effect and its dynamics tell an interesting story. In terms of levels, within-EC trade has always been less than the prediction of the gravity variables. In terms of trend, however, the degree of within-EC bias has clearly risen. This suggests that while the European countries were more open to trade than many countries, for historical or other reasons, their trade pattern exhibits evidence of increasing bias among members and increasing trade diversion away from member countries. Trade among the EFTA countries seems less than what one would have expected based on their economic, geographic, and linguistic links, although the difference is not statistically significant. The Canada-US effect is insignificant. The four South American countries that constitute Mercosur traded more intensely among themselves than the gravity model would have predicted. There appears to have been an increase in the intragroup trade intensity in the 1970s and 1980s. The Andean Group also appears to show a certain degree of intragroup trade bias. The intra-ASEAN trade bias is positive and significant in every year. However, the bias appears to diminish in the last decade, from 2.85 in 1980 to 1.80 in 1992.

Econometric Extensions

Finally, we consider a few extensions concerning the functional form of the equation and the nature of the error term. While these inquiries are a bit more on the technical side, exploring them will reassure us about the robustness of our basic results.

Interactive Effects: Are Blocs More Effective When the Members Are Close Together?

One might wonder whether there are interactive effects between bloc membership and the basic gravity variables: distance and adjacency. Perhaps membership in a bloc has a greater effect when trading partners are

27. Its coefficient has the same interpretation as the coefficient on openness in earlier tables. However, the coefficient on the bloc variable in table B6.6 exceeds the bloc coefficient in the equation specification used before by the amount of the openness coefficient. (This follows from the definitions of the dummy variables: openness = nonbloc + bloc.)

neighbors than when they are far removed, an effect that is greater than the sum of the two separate effects of bloc membership and distance.[28]

We looked at these interaction terms as a robustness check. The coefficient on distance paired with the bloc effect was usually negative and highly significant statistically in three years out of seven over the 1965-91 period (1965, 1970, and 1991). The finding suggests that the combined effect of proximity and bloc membership *is* greater than the sum of the two individual effects. (The negative effect of distance is, of course, a positive effect of proximity.) The coefficient on adjacency paired with the bloc effect fluctuated in sign and was generally not statistically significant.

The results are reported in appendix tables B6.7a and B6.7b. (The former does not distinguish the effects of the individual blocs; the latter does.) The dummy variable for regional blocs in almost all cases had a positive coefficient and was highly significant in four years out of seven. This result suggests that formal preferential trade arrangements in general promote trade, even after taking into account the continental effects.[29]

Zero-Valued Entries

For some country pairs, the data entry is zero, normally due to levels of trade that are too small to be recorded. There are 245 missing values in our data set in 1985, for example. These are generally countries that, by virtue of small size and remoteness, would be expected to have little trade with each other. It is not always possible to ascertain whether their trade is literally zero or is merely very small and has been rounded down to zero. Either way, these zero pairs present a problem for econometric estimation of the gravity model. The reason is that the standard procedure is to take the logarithms of the original multiplicative gravity equation so as to be able to estimate it in linear form. But one cannot take the log of zero.

There are three approaches that have been most commonly taken in the literature. First, one can simply omit the zero pairs from the data set. This is the strategy followed by Brada and Mendez (1985), Bikker (1987), and others. It is also the strategy we have followed in most of our results. One must be concerned, however, that the exclusion of these data points might bias the results. One could argue that a sample-selection problem arises inevitably from leaving very small countries out of the data set

28. Brada and Mendez (1985) find that the bloc effects interact with other effects, that preferential trade arrangements are more likely to boost trade if the members are close together and have relatively high per capita incomes. Mansfield and Bronson (1997) find that bloc effects interact with the effect of common membership in regional geographical groupings, as well as in multilateral alliances.

29. The regional bloc dummy variable in tables B6.7a and b is the sum of the narrowly defined FTAs—e.g., ASEAN, Mercosur—not the continental blocs.

altogether. Even the few studies that include a broader set of countries than our set of 63 nevertheless leave out the smallest countries. No one includes Andorra or Tuvalu, for example.[30]

Second, one can substitute arbitrary small numbers for the zeroes, such as $1,000 or some other minimum unit. This is the technique used by Linnemann (1966) and Wang and Winters (1991), among others. (They found that inclusion of the missing values made little substantive difference to the results.) The virtue of this strategy is that it allows the computer to run the linear OLS regression. The obvious drawback is that it is ad hoc. A less obvious drawback, perhaps, is that the log of a small positive number, though not negative infinity, is still a negative number that is very large (in absolute value). OLS regression, in effect, gives larger weight to extreme values, whether large or small. For this reason, the zero pairs might then receive too large a weight in the estimates. This is related to the subject of weighted least squares, which is addressed below.

Third, one can express the dependent variable, bilateral trade, in levels rather than logs (the semilog formulation), and then use Tobit to estimate. Tobit is a technique that estimates separate parameters to determine whether an observation is nonzero and then to determine what the coefficients are, conditional on the observation being nonzero. This is the technique used by Biessen (1991), Havrylyshyn and Pritchett (1991), and Eaton and Tamura (1994), among others.

A fourth novel approach, employed by Eichengreen and Irwin (1995, 1997), is to express the dependent variable as the log of $(1 + TRADE_{ij})$. Their logic is as follows. On the one hand, when $TRADE_{ij}$ is large, the dependent variable is approximately equal to the usual one—namely $\ln(TRADE_{ij})$—so that the coefficients can be interpreted as elasticities in the usual way. On the other hand, when $TRADE_{ij}$ is small, the dependent variable is approximately equal to $TRADE_{ij}$ itself, which can be interpreted appropriately. Unfortunately, this solution entails some inelegant econometric complications.

Our preferred way of dealing with the problem of zero observations is a simple robustness check. The only reason that the zeroes cannot be included is that the log of zero is minus infinity. Why let a pesky detail like that stop us? We tried running the equation in multiplicative, instead of log-linear, form so as to allow the inclusion of pairs of countries that are reported as undertaking zero trade. When the sizes of the countries are close to (or equal to) zero, the predicted level of bilateral trade will also be close to (or equal to) zero, exactly as it should be. Once the equation has been estimated in nonlinear form, we can test whether the results are sensitive to the exclusion of the zero observations. We find that the inclu-

30. Boisso and Ferrantino (1996) use a data sample of over 260,000 export-import pairs. This sample only goes up to 1985, however. The price of including so many countries is that more recent data are not available.

sion or omission of such countries in the multiplicative specification makes little difference to the results. This finding offers some assurance that the omission of the zero observations from our standard log linear regressions does not lead us far astray.[31]

Heteroskedasticity

The OLS regression technique attaches the same importance to a pair of large countries as to a pair of small countries. It figures that each piece of information is useful. We might figure, however, that the information contained in the number for US-Canada trade, or for Japan-Germany trade, is far more valuable than the information contained in data for Ecuador-Peru or Iceland-Kuwait trade. A good way to think about this is that a country such as the United States is made up of hundreds or thousands of economic units, each of which is no bigger than Iceland or Kuwait. Hypothetically, if we had the trade data for each of these units, their accuracy would likely be not very much better or worse than the data for Iceland or Kuwait. When we add up the US trade data into an aggregate number, it is far more reliable than it would be for a smaller unit: measurement errors tend to cancel out because of the law of large numbers. This is why the US trade numbers are more informative than the Iceland trade numbers. Technically, the problem is heteroskedasticity (larger error terms for some observations than others). The appropriate correction in this case is the technique of weighted least squares, with weights based on the size of the countries. As with so many of the other econometric extensions that we have tried, the correction turns out to make little difference.[32]

Conclusions from the Gravity Model

This completes our econometric extensions of the gravity model. We have confirmed that statistically significant regional trading arrangements are indeed springing up in a number of places, both at the level of formal arrangements and at the broader level of informal continentwide groups. The Andean Group, Mercosur, and ASEAN remain effective preferential trade arrangements. At the continental level, so do the European Union,

31. The results were reported in Frankel and Wei (1993a, appendix tables A2 and A3; 1993b, appendix tables A2, A3). The use of the multiplicative form itself changes the results some, however.

32. Reported in Frankel and Wei (1993b, appendix table A1). Techniques that will correct for heteroskedasticity without requiring any information regarding the structure of the variance matrix have become fashionable among econometricians. When one has a good idea as to the source of the heteroskedasticity, however, as in this case, it seems much better to apply that *a priori* knowledge.

the Americas, and East Asia or the Pacific. Perhaps this chapter's most interesting findings, in their own right, were the trade effects of historical-political links and of bilateral FDI. The absence of systematic factor-endowment effects is also memorable. The larger message, in any case, is that chapter 5's findings of significant regional effects have now been subjected to a great many perturbations and robustness checks and, on the whole, have stood up well.

The next question is whether the regionalization trend constitutes an undesirable threat to the world trading system. We will be returning to the gravity estimates toward the end of the book, when we evaluate the regionalization of trade policy according to what can be justified on the grounds of proximity and the other natural determinants that we have analyzed here.